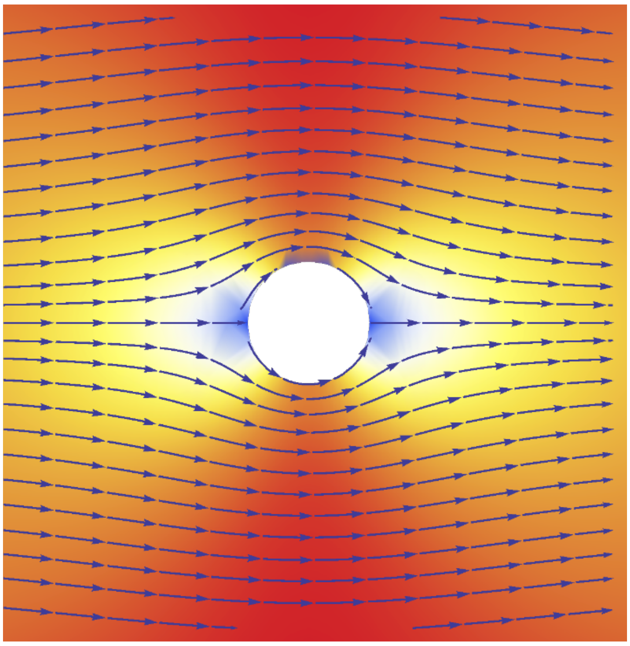
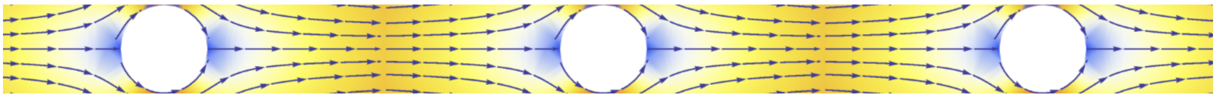
HYDROTRONICS

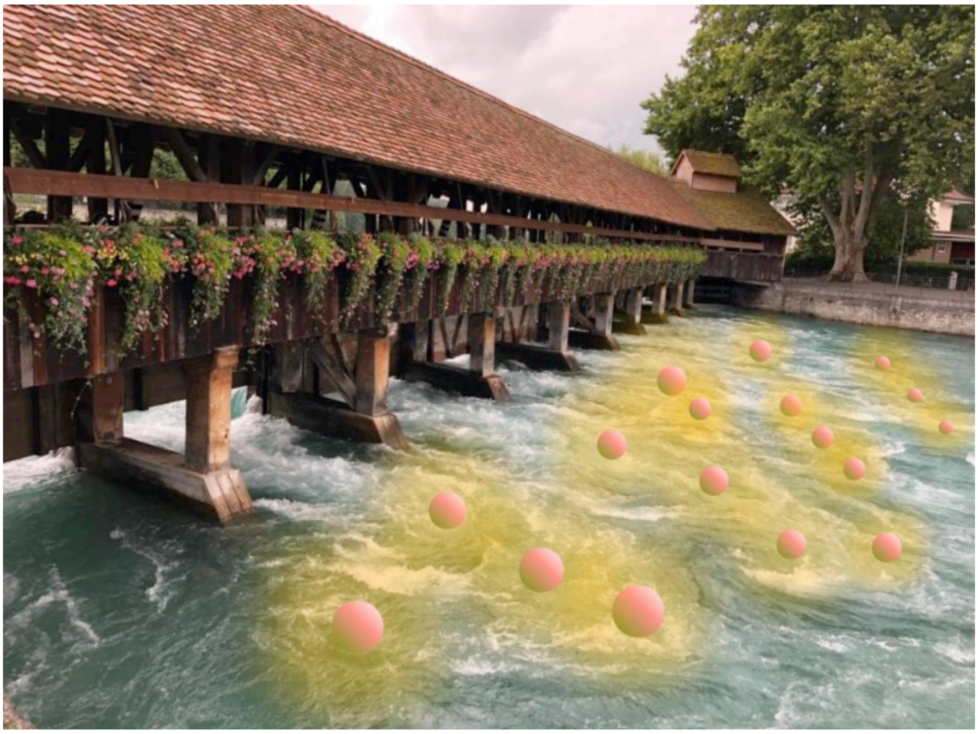


Electron hydrodynamics and beyond



Textbook based on the results of the HYDROTRONICS project

Table of Contents



Hydrodynamics is the oldest and most successful theoretical formalism describing the collective motion of strongly interacting particles. Hydrodynamics has been used to treat a wide range of physical systems, from water to interstellar matter. Advances in the nanostructuring of ultrapure materials over the past few years have given rise to the field of hydrodynamic electronics. Several creative experiments in nanoelectronics have demonstrated that electrons in solids can behave hydrodynamically. In particular, the observation of so-called "super-ballistic" transport demonstrated that the collective flow of charge carriers can be more effective than the usual diffuse, or ballistic, motion of individual particles. The HYDROTRONICS Consortium was founded to deepen our understanding of collective behavior of electrons in novel materials

Introduction to the HYDROTRONICS textbook[1]

Our society would be unimaginable without the control of hydropower. The beginnings of human history date back to advanced civilizations that developed along rivers and estuaries. Since then, humans have controlled water flows to obtain drinking water, develop irrigation systems, and construct artificial waterways for the transport of people and goods. Early examples of sophisticated irrigation systems in Mesopotamia date back to the 6th millennium BC. Hydraulics, the study of the flow behaviour of water, is one of the oldest branches of science and technology.

The theoretical foundation of hydrology and hydraulics is fluid mechanics. The first recorded treatise on fluid mechanics dates back to Archimedes of Syracuse, who studied buoyancy and formulated the law now known as Archimedes' principle around 250 BC. Centuries after Archimedes, the science flourished during the Renaissance in Europe. The invention of the mercury barometer by Torricelli in 1643 is an important milestone in our modern society. Modern barometers, used in medicine, meteorology, and aviation, continue to shape our lives today.

---

[1] Based on extract translated from J. Schmalian, “Hydrodynamische Elektronik”, Yearbook of the Heidelberg Academy of Sciences (2021) [in German]

*Shaduf to raise the water above the Nile level for irrigation purposes in Egypt [Water **7** 5031 (2015)]*

Jean-Pol Grandmont, Braine-le-Château, Belgium (12th century)

Over approximately two hundred years after Torricelli, fluid mechanics was formulated by the most brilliant minds in the history of science, including Newton, Euler, Bernoulli, Poisson, Pascal, and Lagrange. This development reached a climax with the formulation of the Navier-Stokes equations in the mid-19th century. These equations are still intensively studied. The analysis of the existence and regularity of solutions These equations are on the list of the seven most important open problems in mathematics formulated by the Clay Mathematics Institute in Cambridge, Massachusetts. The institute has offered a prize of one million US dollars for each solution.

Beyond their original purpose of understanding water flow in rivers and canals, the Navier-Stokes equations are currently used to describe blood flow in veins, understand ocean currents, describe airflow around cars and airplane wings, and quantitatively analyse the spread of environmental pollution. Even the simulation of air or fluid motion in video games is based on solving these equations. In combination with Maxwell's equations, they are used in the field of magnetohydrodynamics to understand plasmas, liquid metals, and electrolytes. There are applications in geophysics, astrophysics, engineering, and cancer research.

“Poiseuille River” – Svinafellsjokull gletcher, Iceland

Over the past decade, there have been a number of fantastic experimental investigations in which hydrodynamic electronics have been observed [1-7, 23, 24]. These involve ultrapure materials such as $PdCoO_2$ or WP_2, where the sample geometry can be tailored through elegant microstructuring, as well as particularly high-quality graphene on hexagonal boron nitride.

In $PdCoO_2$ [1] and WP_2 [2], but especially in graphene [23,24], a flow behaviour of the electrons was observed that follows the Hagen-Poiseuille law of fluid mechanics. Let l_{imp} be the distance that an electron travels between collisions that violate the momentum conservation. For a sample geometry smaller than l_{imp}, the resistivity of a constriction decreases with its transverse dimension $w$as $\rho \propto w^{-2}$. The proportionality factor depends on the viscosity η of the electron fluid. Such behavior was first theoretically treated for hydrodynamic electron flow by Radii Gurzhi [8]. Ref. [9] is the first experiment in which hydrodynamic Effects have been observed in semiconductor systems. The extension of electronic hydrodynamics to the magnetic field has recently been discussed [10].

A clear manifestation of the hydrodynamic behaviour of charge carriers in graphene is the observation of viscous transport. In an elegant experiment, a negative resistance was determined using four-wire measurements in Ref. [5]. In particular, the authors observed a negative voltage drop near current injection contacts and in an intermediate temperature range (approximately between 100 and 200 K). This remarkable observation can be directly explained by swirl patterns in graphene, i.e., flow fields with vorticity [14].

Closely related to this result is the observation of viscous electron flow in a geometry with constrictions [6]. In such experiments, the sample size is typically smaller than the mean free path of the material, so that charge carriers move ballistically. Accordingly, the system is particularly conductive. This results in an upper limit for the permissible value of conductivity [15]. Remarkably, the viscous flow shows super-ballistic conductivity. The ballistic limit was clearly exceeded. This result is attributed to the cooperative behaviour of the viscous electron fluid. The fluid organizes itself into streams of different velocities. Near edges and defects of the constriction, the flow slows down, whereas the majority of charge carriers flow much faster within the constriction.

Another spectacular manifestation of the hydrodynamic behaviour in graphene is the increased thermal conductivity κ, without a corresponding increase in electrical conductivity σ[4]. In Fermi liquids, at low temperatures, the ratio known as the Lorenz number takes on $L = \kappa/(\sigma T)$the universal value $L_0 = (\pi^2/3)(k_B/e)^2$. Besides a universal number, only the Boltzmann constant $k_B$and the electron charge are taken $e$into account. In undoped graphene, a strong increase in the Lorentz number up to $L \approx 20L_0$[4] was measured.

Inspired by these results, since 2019, the HYDTROTRONICS consortium has undertaken a wide range of research activities to investigate the details of electronic hydrodynamics in various physical systems. The findings, ranging from the discovery of new solutions to the hydrodynamic equations to novel regimes of transport and nonlinearities, as well as the physics of twisted materials, are collected in this textbook.

References:

[1] P.J.W. Moll, P. Kushwaha, N. Nandi, B. Schmidt & AP Mackenzie, Science **351**, 1061 (2016).

[2] J. Gooth, F. Menges, C. Shekhar, V. Süß, N. Kumar, Y. Sun, U. Drechsler, R. Zierold, C. Felser & B. Gotsmann, Nature Communications **9**, 4093 (2018).

[3] M. Titov, R.V. Gorbachev, B.N. Narozhny, T. Tudorovskiy, M. Schütt, P.M. Ostrovsky, I.V. Gornyi, A.D. Mirlin, M.I. Katsnelson, K.S. Novoselov, A.K. Geim & LA Ponomarenko, Phys. Rev. Lett. **111**, 166601 (2013).

[4] J. Crossno, J.K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T.A. Ohki & K.C. Fong, Science **351**, 1058 (2016).

[5] D.A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G.H. Auton, E. Khestanova1, K.S. Novoselov, I.V. Grigorieva, L.A. Ponomarenko, A.K. Geim & M. Polini, Science **351**, 1055 (2016).

[6] R. Krishna Kumar, D.A. Bandurin, F.M.D. Pellegrino, Y. Cao, A. Principi, H. Guo, G.H. Auton, M. Ben Shalom, L.A. Ponomarenko, G. Falkovich, I.V. Grigorieva, LS Levitov, M. Polini & A.K. Geim, Nature Physics **13**, 1182 (2017).

[7] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, MS Foster & P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[8] RN Gurzhi, Usp. Fiz. Nauk **94**, 689 (1968) [Sov. Phys. Usp. **11**, 255 (1968)].

[9] M.J.M. de Jong & L.W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[10] P.S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016); T. Scaffidi, N. Nandi, B. Schmidt, A.P. Mackenzie & J.E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[11] B.N. Narozhny, I.V. Gornyi, A.D. Mirlin & J. Schmalian, Annals of Physics **529**, 1700043 (2017).

[12] U. Briskot, M. Schütt , I.V. Gornyi, M. Titov, B.N. Narozhny & A.D. Mirlin, Phys. Rev. B **92**, 115426 (2015).

[13] J.M. Link, B.N. Narozhny, E.I. Kiselev & J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018),

[14] L.S. Levitov & G. Falkovich , Nature Phys. **12**, 672 (2016).

[15] Yu. V. Sharvin, J. Exp. Theor. Phys. (USSR) **48**, 984 (1965).

[16] D. Vollhardt & P. Wölfle, *The Superfluid Phases of Helium 3* (Taylor & Francis Ltd, 1990).

[17] A.V. Andreev, S.A. Kivelson, and B. Spivak, Phys. Rev. Lett. **106**, 256804 (2011).

[18] A.D. Mirlin & J. Schmalian, Physics Journal **15**, 18 (2016).

[19] A. Kashuba, Phys. Rev. B **78**, 085415 (2008); L. Fritz, J. Schmalian, M. Müller & S. Sachdev, Phys. Rev. B **78**, 085416 (2008)

[20] M. Müller, J. Schmalian & L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[21] P.K. Kovtun, D.T. Son & A.O. Starinets, Phys. Rev. Lett. **94**, 111601 (2005).

[22] B.N. Narozhny & A. Levchenko, Rev. Mod. Phys. **88**, 025003 (2016).

[23] J.A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D.J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A.K. Geim & S. Ilani, Nature **576**, 75 (2019).

[24] M.J.H. Ku, T.X. Zhou, Q. Li, Y.J. Shin, J.K. Shi, C. Burch, L.E. Anderson, A.T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M.M. Fogler, P. Kim, A. Yacoby & R.L. Walsworth, Nature **583**, 537 (2020).

1. Charge and heat transport in the hydrodynamic regime

1.1 Non-local hydrodynamic transport phenomena

We experience now a rare moment of intense interaction between the fields of solids and fluids. This is due to the appearance of new high-mobility materials where charge carriers exchange momentum among themselves faster than with the lattice. In this regime, their collective motion is a viscous fluid flow. Ideas from fluid mechanics can then be employed to solve problems of nanoscale electronics: in particular, decrease resistance below the ballistic limit and make the current flow against the electric field.

1.1.1 Force-free dissipative flows

As shown in Ref. [1], no less remarkable is what electronics can do for fluid mechanics. 150 years after Stokes, it can reveal new fundamental phenomena in laminar flows never predicted or observed before. The reason is that electronics brings a new setting not regularly considered in low-Reynolds hydrodynamics: the presence of equipotential (metallic) electrodes serving as flow sources embedded inside the fluid. Ref. [1] has shown that the conditions on the electric potential (pressure) imposed by sources could be in conflict with those of a viscous flow, which leads to anomalies at the boundaries and novel flow properties. In particular, a nontrivial analysis of the current distribution and work done in different parts of the flow is needed to ensure the energy budget.

In the simplest case, when one considers a viscous radial flow, they discover that the viscous force vanishes identically *i.e.*, the Laplacian of velocity is zero. At the same time, the energy dissipation and heat release are everywhere nonzero, because the velocity gradients are nonzero. The electric field is thus expelled from the bulk and is concentrated in the boundary layer in a viscous flow. The potential jump is proportional to the viscosity, that is to the mean free path. That means that the electric field inside the ballistic layer is independent of the mean free path. When one goes deep into the hydrodynamic regime (say, by increasing the temperature in graphene), the mean free path shrinks but the electric field stays finite. Stokes encountered a similar phenomenon of bulk dissipation equal to the surface work in his analysis of the decay of water waves: The flow in the bulk is potential, while the viscous forces only perform work on the surface.

How is the above picture of field expulsion modified for an electrode of an arbitrary shape? A purely Ohmic flow in 2d can be found via conformal mapping that deforms one electrode into another and also transforms streamlines and potential contours. The naive reasoning outlined above suggests that, when viscosity is present, the transformation of the potential might still be possible due to conformal invariance of the Laplacian. However, this conformal equivalence does not hold for viscous flows, since the field distribution depends nontrivially upon the shape of the electrode; in particular, for nonsymmetric electrodes the flow, in general, is not potential and the electric field partially penetrates the fluid.

While both Ohmic and viscous flows are inherently dissipative, there is a dramatic difference in the spatial distribution of the work done to compensate this dissipation. In Ohmic flows, the momentum and the energy losses are locally compensated by an electric field proportional to the current at every point. On the contrary, momentum is diffusively conserved by viscous flows, while the energy is lost everywhere there is a velocity gradient. The electrical work compensating the viscous energy loss can be partially or even fully done on the flow boundaries.

To summarise, electronic fluids bring into hydrodynamics a new setting: equipotential flow sources embedded inside the fluid. The nonlocal relation between the current and electric field due to momentum-conserving interparticle collisions leads to a total or partial field expulsion from such flows. This results in freely flowing currents in the bulk and a boundary jump in the electric potential at current-injecting electrodes. Despite the fact that the spherical flow in the bulk is force, the velocity gradients are nonzero everywhere and so is the dissipation and heat release. One then needs to analyse carefully the energy budget, including the work done on the electrodes. A new type of boundary conditions was derived in Ref. [1] appropriate for this case. In the same reference, the current and work distribution was then analysed for free flows, and the dependence of the field expulsion on the geometry of the electrode was discussed. The phenomenon was linked to the breakdown of conformal invariance.

1.1.2 Para-hydrodynamic transport and vortices

An even less trivial distribution of current, momentum diffusion and work correspond to an ultimate fluid-mechanical phenomenon – vortices. In Ref. [2] the macroscopic classical vortices were for the first time directly observed experimentally in electronic flows and described theoretically. Vortices are the hallmarks of hydrodynamic flow. The visualization of whirlpools in an electron fluid was achieved by using a nanoscale scanning superconducting quantum interference device on a tip, which brought the image of the current distribution in a circular chamber connected through a small aperture to a current-carrying strip in the high-purity type II Weyl semimetal WTe_2. In this geometry, the Gurzhi momentum diffusion length and the size of the aperture determine the vortex stability phase diagram. Vortices are present only for small apertures, whereas the flow is laminar (non-vortical) for larger apertures. Near the vortical-to-laminar transition, the splitting of a single vortex into two vortices was observed. This behaviour is expected only in the hydrodynamic regime and is not anticipated for ballistic transport.

Perhaps the most unexpected finding is the fluid-like momentum transport characterized by a Gurzhi length, which is much smaller than the one estimated from the bulk microscopic parameters. To gain insight into the origin of this surprising behaviour, it is pivotal to recall the general derivation of the Ohm–Stokes law. Kinetic theory links momentum relaxation to the decay rate γ_1 of the first angular harmonic of the nonequilibrium electron momentum distribution, whereas the kinematic viscosity is expressed through the decay rate γ_2 of the second harmonic of the momentum distribution. Importantly, this expression is valid for any microscopic momentum scattering mechanism; γ_1, γ_2 can originate from several sources, including impurity scattering, phonons and electron–electron collisions.

However, spatial diffusion of momentum can originate from effects unrelated to electron–electron scattering. One appealing alternative mechanism that could result in both diffusion and relaxation of electron momenta is related to the finite thickness d of the sample. The resistivity of WTe_2 flakes is strongly dependent on thickness, a behaviour attributed to surface oxidation. Our transport measurements show that the conductivity of the flakes is one to two orders of magnitude lower than that of the bulk crystals, indicating a large enhancement of γ_1 induced by the momentum-relaxing scattering off the surfaces. However, enhancement of γ_1 alone, such that $\gamma_1 > \gamma_2$, would of course lead to ohmic transport with no vorticity, in contrast to the observed hydrodynamic flow, which requires $\gamma_2 > \gamma_1$. Thus, an enhancement in γ_1 implies an enhancement in γ_2. Indeed, the theory developed in Ref. [2] shows that small-angle scattering results in $\gamma_2 \cong 4\gamma_1$, giving rise to para-hydrodynamic transport.

The emerging picture is therefore as follows. For fully specular surface scattering, the transport is ballistic. Small-angle scattering at the surfaces results in the enhancement of momentum relaxation and of lateral momentum diffusion. In this para-hydrodynamic regime, momentum diffusion occurs not through the usual momentum-conserving electron–electron scattering but rather through close-to-specular scattering of individual particles at the top and bottom surfaces. The surface-induced para-hydrodynamics presents a unique opportunity to explore hydrodynamic phenomena in a wide range of high-mobility materials without the hard-to-achieve strong bulk electron–electron interactions.

This surface-induced para-hydrodynamics, which mimics many aspects of conventional hydrodynamics including vortices, opens new possibilities for exploring and using electron fluidics in high-mobility electron systems.

1.1.3 Hydrodynamic flows in ultrapure graphene at charge neutrality

The diversity of quasiparticle spectra in solids allows for hydrodynamic flows in electronic systems that have not been previously encountered in conventional fluids. In particular, intrinsic graphene exhibits unconventional behaviour where the energy flow is hydrodynamic (and is described by a velocity field obeying the Navier-Stokes-like equation), while the charge flow is Ohmic [3]. The resulting flow profile depends on sample geometry.

In the more conventional channel geometry (modelling a Hall bar) we find that the hallmark Poiseuille flow cannot be driven by the electric field irrespective of boundary conditions at the channel edges. The resulting electric current is independent of viscosity. Nevertheless, one can observe nonuniform current densities similarly to the case of the well-known ballistic-diffusive crossover. The standard diffusive behaviour with a uniform current density across the channel is achieved under the assumption of specular scattering at the channel boundaries. It can be made inhomogeneous by applying an external magnetic field. In this case the current profile is anti-Poiseuille (with a minimum current density in the centre of the channel) [4]. See also 2.1.2 for a more detailed discussion.

On the other hand, in the circular Corbino geometry the electric current is nonuniform even in the simplest Drude picture. At charge neutrality, the non-hydrodynamic electric current is accompanied by the hydrodynamic energy flow and hence by viscous dissipation and energy relaxation [5]. The two currents are induced by the same source providing the total dissipated energy and determining the voltage drop in the sample. Given that the voltage drop in the bulk is fixed by the electric current density, the additional voltage drop due to viscous dissipation must take form of a potential jump at the sample-lead interface (with an excess electric field induced in the thin Knudsen layer around the interface). This potential jump is distinct from the usual contact resistance and is a function of the system size.

An external magnetic field couples the charge and energy currents forcing them to flow in orthogonal directions. In contrast to the Hall bar geometry, the lateral energy current flows freely around the Corbino disk without accumulating quasiparticles at any point. As a result, the device resistance as a function of magnetic field remains positive and parabolic. Furthermore, the magnetoresistance of the Corbino device exhibits a crossover from the "hydrodynamic" (viscosity dominated) to the "bulk" (disorder-limited) behaviour with increasing system size (compared to the Gurzhi length). In the clean limit the magnetoresistance remains finite and is determined by viscosity, offering a way to measure the viscosity coefficient in neutral graphene. The "contact magnetoresistance" induced through the dissipation jump is present but is typically weaker than the bulk contribution [6].

The experimental imaging of magneto-hydrodynamic Corbino flows (including the potential and work distribution in ultrapure graphene samples) was reported in Ref. [7] (see also [11]).

The decoupling of charge and energy currents in intrinsic graphene may be traced back to the specific feature of graphene band structure where two bands touch at the Dirac points. At finite temperatures, both bands contain mobile carriers leading to the two-component nature of the electronic system. Given the exact particle–hole symmetry at neutrality, this system is "compensated" and hence there is no classical Hall effect. The bulk Hall conductivity vanishes, while the longitudinal conductivity is unaffected by the magnetic field. The nonlocal response instead depends strongly on the applied field. This effect was studied in Ref. [8] capturing the main qualitative features of giant nonlocality observed experimentally in Ref. [9].

1.1.4 Collective excitations in ultrapure graphene in the hydrodynamic regime

One of the most intriguing aspects [3] of the hydrodynamic approach to electronic transport is collective behaviour. Viscous effects and weak disorder lead to damping of the hydrodynamic sound mode. The latter effect does not appear in traditional hydrodynamics since in electronic systems dissipation due to "external" scattering (e.g., disorder and electron–phonon scattering) appears already in the description of an "ideal" (*i.e.*, inviscid) electronic fluid. The sound modes have to be distinguished from plasmonic excitations. Typically, the two types of collective modes are defined in different parameter regimes: hydrodynamics is applicable at small momenta, $\boldsymbol{q\ell_{hydro}} \ll \mathbf{1}$, while plasmons are nonequilibrium excitations that belong to higher momenta. In graphene, the possibility of discussing momenta exceeding $\mathbf{1}/\boldsymbol{\ell_{hydro}}$, is afforded by the collinear scattering singularity which leads to the existence of two parametrically different length scales and hence of an intermediate momentum range, $\boldsymbol{\ell_{hydro}^{-1}} \ll \boldsymbol{q} \ll \boldsymbol{\ell_{coll}^{-1}}$. Here a linear response theory can be used to find the collective modes. Remarkably, the macroscopic equations of this theory coincide with the linearized hydrodynamic equations such that the resulting dispersions should be valid in the hydrodynamic regime as well.

We find that the plasmon mode as established by the macroscopic theory should be contrasted with the diffusive charge mode and not the sound mode of electronic hydrodynamics. The plasmon and the sound belong largely to different frequency regimes, but most importantly, stem from the two different, decoupled sectors of the theory (the sound mode can also be obtained from the linear response theory hence one can extend its region of applicability beyond the hydrodynamic regime). The latter fact is the reason why the plasmon dispersion is independent of viscosity, while the sound mode is unaffected by screening effects (which are essentially responsible for plasmon excitations). Formally, the two modes coexist but are characterized by different frequencies that are much higher for the plasmon mode. Approximately at wavevectors of the order of $\sim\boldsymbol{\ell_{coll}^{-1}}$, i.e., at the applicability limit of the linear response theory, the sound mode becomes overdamped, which does not happen to the plasmon [10].

1.1.5 References

1. Shavit M., Shytov A. & Falkovich G. Freely Flowing Currents and Electric Field Expulsion in Viscous Electronics. Phys. Rev. Lett. 123, 2, 026801 (2019); https://doi.org/10.48550/arXiv.1812.10981
2. Aharon-Steinberg, A., Völkl, T., Kaplan, A. et al. Direct observation of vortices in an electron fluid. Nature 607, 74–80 (2022); https://doi.org/10.1038/s41586-022-04794-y
3. B. N. Narozhny. Hydrodynamic approach to two-dimensional electron systems. La Rivista del Nuovo Cimento 45, 661 (2022); https://doi.org/10.1007/s40766-022-00036-z
4. B. N. Narozhny, I. V. Gornyi, and M. Titov. Anti-Poiseuille flow in neutral graphene. Phys. Rev. B 104, 075443 (2021); https://doi.org/10.1103/PhysRevB.104.075443
5. V. Gall, B. N. Narozhny, and I. V. Gornyi. Electronic viscosity and energy relaxation in neutral graphene. Phys. Rev. B 107, 045413 (2022); https://doi.org/10.1103/PhysRevB.107.045413
6. V. Gall, B. N. Narozhny, and I. V. Gornyi. Corbino magnetoresistance in neutral graphene. Phys. Rev. B 107, 235401 (2023); https://doi.org/10.1103/PhysRevB.107.235401
7. C. Kumar, J. Birbeck, J.A. Sulpizio, D. Perello *et al.* Imaging hydrodynamic electrons flowing without Landauer–Sharvin resistance. Nature 609, 276 (2022); https://doi.org/10.1038/s41586-022-05002-7
8. S. Danz, M. Titov, and B. N. Narozhny, Giant nonlocality in nearly compensated two-dimensional semimetals. Phys. Rev. B 102, 081114(R) (2020); https://doi.org/10.1103/PhysRevB.102.081114
9. D. A. Abanin, S. V. Morozov, L. A. Ponomarenko *et al*. Giant Nonlocality Near the Dirac Point in Graphene. Science 332, 328 (2011); https://doi.org/10.1126/science.1199595
10. B. N. Narozhny, I. V. Gornyi, and M. Titov. Hydrodynamic collective modes in graphene. Phys. Rev. B 103, 115402 (2021); https://doi.org/10.1103/PhysRevB.103.115402
11. Ady Stern, Thomas Scaffidi, Oren Reuven, Chandan Kumar, John Birkbeck, and Shahal Ilani. How Electron Hydrodynamics Can Eliminate the Landauer-Sharvin Resistance. Phys. Rev. Lett. 129, 157701 (2022); https://doi.org/10.1103/PhysRevLett.129.157701

1.2 The role of boundary conditions in heat transport

The primary objective here was to investigate and elucidate the specific boundary conditions affecting the behaviour of the viscous electronic fluid in graphene and other Dirac materials. A significant challenge in non-Galilean-invariant systems is the lack of equivalence between macroscopic currents and the velocity field. In graphene, for example, the energy current corresponds to momentum, which, because of the linear spectrum, is not equivalent to velocity. Consequently, electron-electron interactions in graphene lead to the conservation of energy current, resulting in unique heat transport characteristics.

1.2.1 Nonlocal hydrodynamic transport and collective excitations in Dirac fluids

In Ref. [1], the response of a Dirac fluid to electric fields and thermal gradients at finite wave numbers and frequencies in the hydrodynamic regime was studied. It was found that nonlocal transport in the hydrodynamic regime is governed by an infinite set of kinetic modes that describe noncollinear scattering events in different angular harmonic channels. The scattering rates of these modes increase as |m|, where m labels the angular harmonics. In earlier works, it was pointed out that this dependence leads to anomalous, Lévy-flight-like phase-space diffusion [2]. In Ref. [1] it was shown how this surprisingly simple, nonanalytic dependence allows one to obtain exact expressions for the nonlocal charge and electronic thermal conductivities. The peculiar dependence of the scattering rates on m also leads to a nontrivial structure of collective excitations. Besides the known plasmon, second-sound, and diffusive modes, one finds nondegenerate damped modes corresponding to excitations of higher angular harmonics. These results were used to investigate the transport of a Dirac fluid through Poiseuille-type geometries under different boundary conditions of different widths and to study the response to surface acoustic waves in graphene-piezoelectric devices.

1.2.2 Heating of inhomogeneous electron flow and the contribution of supercollisions

In Ref. [3], the electron temperature profiles for an inhomogeneous electron flow in the hydrodynamic regime were studied. It was assumed that the inhomogeneity was due to a weakly nonuniform distribution of the momentum relaxation time within a spherically constricted area. The temperature profile was shown to dramatically depend on the drive strength and the viscosity of the electron liquid. In the absence of viscosity, a Landauer-dipole-like temperature distribution, asymmetrically deformed along the current by the inelastic electron-phonon scattering, emerges around the inhomogeneity. Both the Landauer-dipole temperature profile and its asymmetry in the direction of the driving electric field exist in all dimensionalities and are, therefore, universal features of inhomogeneous hydrodynamic electron flow. It was further demonstrated that the electron viscosity suppresses the thermal Landauer dipole and leads to the appearance of a “hot spot” exactly at the centre of the constriction. The phonon temperature distribution was calculated: this can be directly measured in experiments on thermal nanoimaging.

In nearly compensated graphene, disorder-assisted electron-phonon scattering or "supercollisions" are responsible for both quasiparticle recombination and energy relaxation. Within the hydrodynamic approach, these processes contribute weak decay terms to the continuity equations at local equilibrium, i.e., at the level of “ideal" hydrodynamics. In Ref. [1], the derivation of the decay term due to a weak violation of energy conservation was reported. Such terms have to be considered on equal footing with the well-known recombination terms due to the non-conservation of the number of particles in each band. At high enough temperatures in the “hydrodynamic regime", supercollisions dominate both types of interaction. The contribution of supercollisions to the heat transfer equation was also discussed, generalizing the continuity equation for the energy density to viscous hydrodynamics.

1.2.3 Boundary conditions in Corbino geometry

The role of boundary conditions between the graphene sample and metallic leads in the Corbino geometry was also studied, see Refs. [4] and [5] In the circular geometry, the electric current displayed nonuniform behaviour. At charge neutrality, the non-hydrodynamic electric current was accompanied by a hydrodynamic energy flow, resulting in viscous dissipation and energy relaxation. Both currents stemmed from the same source, contributing to the total dissipated energy and determining the voltage drop in the sample. The introduction of an external magnetic field causes the charge and energy currents to flow in orthogonal directions. Unlike the Hall-bar geometry, the lateral energy current in the Corbino disk circulates freely without accumulating quasiparticles at any point. Consequently, the device resistance as a function of the magnetic field remains positive and parabolic. A crossover in the Corbino magnetoresistance from "hydrodynamic" (viscosity-dominated) to "bulk" (disorder-limited) behaviour was predicted with increasing system size compared to the Gurzhi length. In the clean limit, the magnetoresistance provides a means to measure the viscosity coefficient in neutral graphene. While the "contact magnetoresistance" resulting from the dissipation jump at the interfaces between the samples and leads was present, it typically exhibited weaker contributions than the bulk effect.

1.2.4 Giant thermal diffusivity of Dirac fluids

Conducting materials typically exhibit either diffusive or ballistic charge transport. However, when electron-electron interactions dominate, a hydrodynamic regime with viscous charge flow emerges. More stringent conditions eventually yield a quantum-critical Dirac-fluid regime, where electronic heat can flow more efficiently than charge. Heat transport in graphene in the diffusive and hydrodynamic regimes was studied experimentally in Ref. [6], reporting a controllable transition to the Dirac-fluid regime at room temperature, using carrier temperature and carrier density as control knobs. The technique of spatiotemporal thermoelectric microscopy with femtosecond temporal and nanometre spatial resolution, which allows for tracking electronic heat spreading, was introduced. In the diffusive regime, a thermal diffusivity of ∼2,000 cm^2/s, consistent with charge transport, was found. Remarkably, during the hydrodynamic time window before momentum relaxation, an ultrafast heat spreading corresponding to a giant diffusivity up to 70,000 cm^2/Vs was observed, indicative of a Dirac fluid. These results are promising for applications such as nanoscale thermal management.

1.2.5 Thermal transport in compensated semimetals

It is well known that the electronic thermal conductivity of clean compensated semimetals can be greatly enhanced over the electric conductivity by the availability of an ambipolar mechanism of conduction, whereby electrons and holes flow in the same direction experiencing negligible Coulomb scattering as well as negligible impurity scattering. This enhancement—resulting in a breakdown of the Wiedemann-Franz law with an anomalously large Lorenz ratio—has been recently observed in two-dimensional monolayer and bilayer graphene near the charge neutrality point. In contrast to this, three-dimensional compensated semimetals such as WP_2 and Sb are typically found to show a reduced Lorenz ratio. In Ref. [7], the reasons for this difference were investigated, focusing on the low-temperature regime where electron-electron scattering is expected to dominate over other scattering mechanisms. The different regimes of Fermi statistics (nondegenerate electron-hole liquid in graphene versus degenerate electron-hole liquid in compensated semimetals) are not sufficient to explain the reduction of the Lorenz ratio in the latter. The proposed solution to the puzzle lies in the large separation of electron and hole pockets in momentum space, which allows compensated semimetals to sustain sizable regions of electron-hole accumulation near the contacts. These accumulations suppress the ambipolar conduction mechanism and effectively split the system into two independent electron and hole conductors. A quantitative theory of the crossover from ambipolar to unipolar conduction as a function of the size of the electron-hole accumulation regions was presented and shown to naturally lead to a sample-size-dependent thermal conductivity.

1.2.6 Interplay between two mechanisms of resistivity

In Ref. [8] the interplay between two mechanisms of resistivity was studied. Mechanisms of resistivity can be divided into two basic classes: one is dissipative (like scattering on phonons) and another is quasi-elastic (like scattering on static impurities). They are often treated by the empirical Matthiessen rule, which says that total resistivity is just the sum of these two contributions, which are computed separately. This is quite misleading for two reasons. First, the two mechanisms are generally correlated. Second, computing the elastic resistivity alone masks the fundamental fact that the linear-response approximation has a vanishing validity interval at vanishing dissipation. Limits of zero electric field and zero dissipation do not commute for the simple reason that one needs to absorb the Joule heat quadratic in the applied field. A simple model that illustrates these two points was presented in Ref. [9]. The model also illuminates the role of variational principles for non-equilibrium steady states.

1.2.7 References

1. Egor I. Kiselev and Jörg Schmalian, Nonlocal hydrodynamic transport and collective excitations in Dirac fluids, Phys. Rev. B 102, 245434 (2020); https://doi.org/10.1103/PhysRevB.102.245434
2. Egor I. Kiselev and Jörg Schmalian, Lévy Flights and Hydrodynamic Superdiffusion on the Dirac Cone of Graphene, Phys. Rev. Lett. 123, 195302 (2019); https://doi.org/10.1103/PhysRevLett.123.195302
3. Gu Zhang, Valentin Kachorovskii, Konstantin Tikhonov, and Igor Gornyi, Heating of inhomogeneous electron flow in the hydrodynamic regime, Phys. Rev. B 104, 075417 (2021); https://doi.org/10.1103/PhysRevB.104.075417
4. B. N. Narozhny, I. V. Gornyi, Hydrodynamic Approach to Electronic Transport in Graphene: Energy Relaxation, Front. Phys. 9, 1 (2021); https://doi.org/10.3389/fphy.2021.640649
5. Vanessa Gall, Boris N. Narozhny, and Igor V. Gornyi, Electronic viscosity and energy relaxation in neutral graphene, Phys. Rev. B 107, 045413 (2023); https://doi.org/10.1103/PhysRevB.107.045413

6. Vanessa Gall, Boris N. Narozhny, and Igor V. Gornyi, Corbino magnetoresistance in neutral graphene, Phys. Rev. B 107, 235401 (2023); https://doi.org/10.1103/PhysRevB.107.235401
7. Alexander Block, Alessandro Principi, Niels C. H. Hesp, Aron W. Cummings, Matz Liebel, Kenji Watanabe, Takashi Taniguchi, Stephan Roche, Frank H. L. Koppens, Niek F. van Hulst, and Klaas-Jan Tielrooij, Observation of giant and tunable thermal diffusivity of a Dirac fluid at room temperature, Nature Nanotechnology volume 16, 1195 (2021); https://doi.org/10.1038/s41565-021-00957-6
8. Mohammad Zarenia, Alessandro Principi, and Giovanni Vignale, Thermal transport in compensated semimetals: Effect of electron-electron scattering on Lorenz ratio, Phys. Rev. B 102, 214304 (2020); https://doi.org/10.1103/PhysRevB.102.214304
9. A Kapustin and G Falkovich. Interplay between two mechanisms of resistivity arXiv:2407.16284; https://doi.org/10.48550/arXiv.2407.16284

2. Non-locality, non-linearity, and light matter interaction

2.1 Theory of non-local hydrodynamic transport phenomena

The goal of the research presented here is to numerically integrate the hydrodynamic equations and evaluate the non-local conductivity for arbitrary momenta. Additionally, this research aims to study the spatial distribution of the charge flow in various sample geometries, and investigate transport phenomena in quantum critical systems, such as graphene at the charge neutrality point. While local transport coefficients of this system have been thoroughly studied, the non-local conductance properties are much less explored.

2.1.1 Giant nonlocality in graphene-like semimetals

In compensated two-component systems in confined, two-dimensional geometries, nonlocal response may appear due to external magnetic field. In Ref. [1] the evolution of charge flow profiles and the emergence of a giant nonlocal pattern dominating charge transport in magnetic field was demonstrated. Applying our approach to the specific case of intrinsic graphene, a simple physical explanation for the experimental observation of giant nonlocality was suggested. These results provide an intuitive way to predict the outcome of future experiments exploring the rich physics of many-body electron systems in confined geometries as well as to design possible applications.

In particular it was shown that the observed giant nonlocality in neutral graphene at high temperatures in non-quantizing magnetic fields, as reported in Ref. [21], can be attributed to the two-band nature of the quasiparticle spectrum in graphene. This effect is not specific to graphene, and it was anticipated that it should be observable in any compensated two-component system.

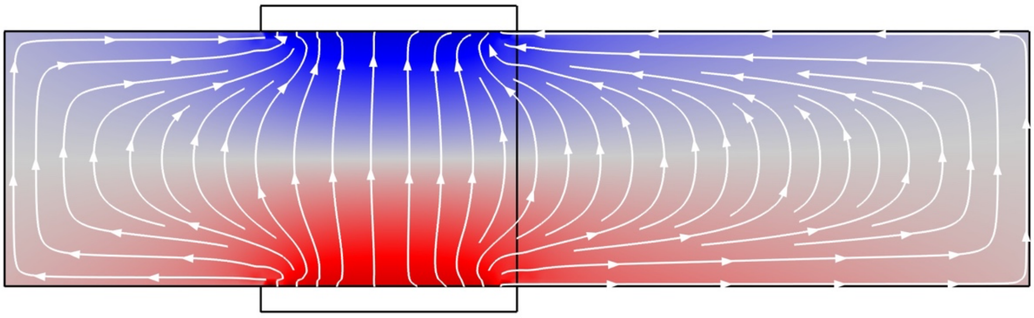


Figure 1: *Giant nonlocality in a compensated semimetal in magnetic field. The arrows indicate the current flow and the color map shows the electrochemical potential (see the main text and Figs. 2 and 3 for specific parameters).*

The theory, with material-specific parameters, provides a quantitative description of the effect observed in experiment. However, for graphene, a more precise calculation involving the solution of the full system of hydrodynamic equations near charge neutrality would be necessary to achieve perfect agreement with data. Nonetheless, this approach demonstrates that the effect is general and does not require additional assumptions of electronic hydrodynamics. The nonlocal flow pattern that emerges in a magnetic field, as shown in Figs. 1, 2, and 3, is distinct from the vortices that arise in viscous hydrodynamic flow, as seen in doped graphene in prior studies. In the latter case, vorticity results from the geometry of the flow and specific boundary conditions, and the solution of the hydrodynamic equations can be obtained by introducing the stream function. This stream function satisfies a biharmonic equation independent of viscosity, which affects the distribution of the electrochemical potential. In contrast, within the present model, the "Ohmic" scattering is the only source of dissipation and cannot be omitted. While one can still introduce a stream function, it is determined not only by the sample geometry but also by the Ohmic scattering and magnetic field. As a result, the flow pattern does not exhibit vortices, unlike those that have been suggested for hydrodynamic flow in intrinsic graphene in the absence of magnetic field in other studies.

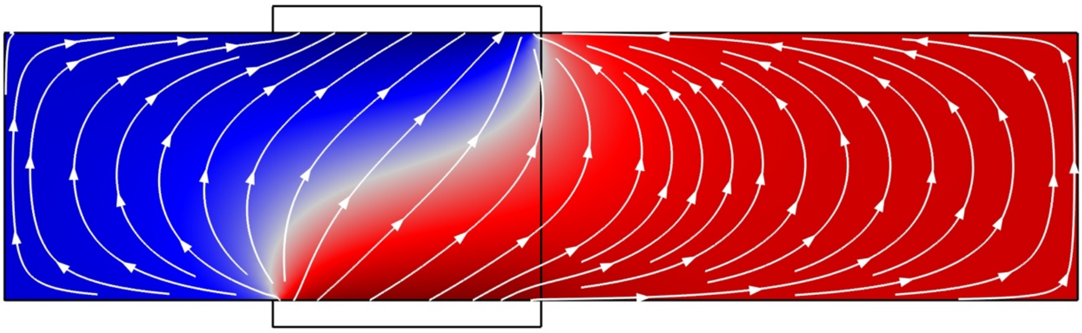


Figure 2: *Classical Hall effect in a one-component electronic system. The current density (shown by the arrows) and the electrochemical potential (shown by the color map) were ob- tained from Eqs. (1) for a sample of the width W = 1 μm and length L = 4 μm with t*

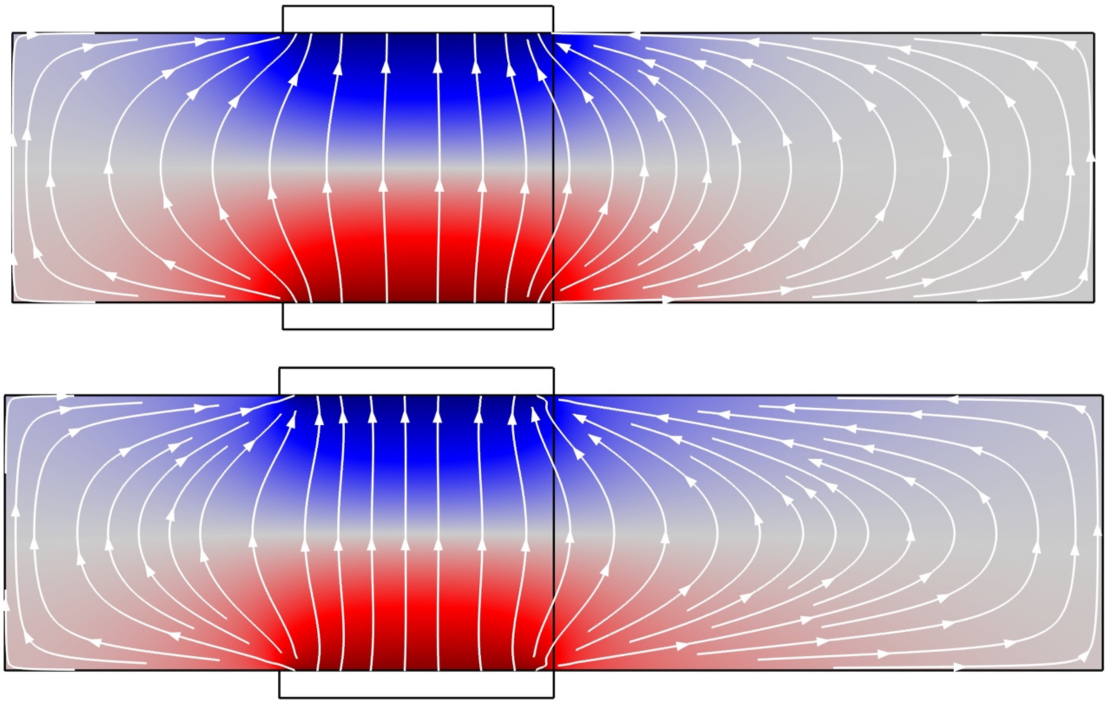


Figure 3: *Charge flow in compensated semimetals. Top: Ohmic flow in the absence of magnetic field. Bottom: emergent non- locality in weak magnetic field B = 0.2T. The associated potential on the sample boundaries grows with the increasing field, see Fig. 1 for the pattern at B = 2T. Stronger fields expel the current from the bulk such that it flows along the boundary.*

2.1.2 Anti-Poiseuille flow in neutral graphene

In a second publication [2] the hydrodynamic flow of charge carriers in graphene was demonstrated to be distinct from the mass flow observed in conventional fluids [1]. In neutral graphene, the energy flow is decoupled from the electric current, making it challenging to observe hydrodynamic effects and measure the viscosity of the electronic fluid via electric current measurements. The hallmark Poiseuille flow in a narrow channel cannot be driven by the electric field, regardless of the boundary conditions at the channel edges. Nevertheless, nonuniform current densities can be observed similarly to the well-known ballistic-diffusive crossover. The standard diffusive behaviour with uniform current density across the channel is achieved under the assumption of specular scattering on the channel boundaries. This flow can also be made nonuniform by applying weak magnetic fields. In this case, the curvature of the current density profile is determined by the quasiparticle recombination processes dominated by disorder-assisted electron-phonon scattering, known as supercollisions.

Despite the challenges in measuring hydrodynamic effects in neutral graphene, this work demonstrates that the Poiseuille flow can be used as a hallmark of viscosity [3]. The Poiseuille flow is a particular solution to the Navier-Stokes equation in the case where a viscous, incompressible fluid is constrained by stationary boundaries. The problem is usually solved under the assumption of the no-slip boundary conditions, meaning that the flow velocity at the boundaries is zero. Then, the Navier-Stokes equation becomes an ordinary second-order differential equation yielding the standard parabolic velocity profile. The solution can be extended to the case of more general Maxwell's boundary conditions with a finite slip length. However, the limit of infinite slip length, i.e., with no-stress boundary conditions, does not admit any solutions for the Poiseuille problem. In other words, a pressure-induced viscous flow in a pipe cannot be homogeneous. On the contrary, an inviscid fluid is described by the Euler equation, which is a nonlinear, first-order differential equation that does not require boundary conditions on the longitudinal component of the velocity and allows for homogeneous solutions. Hence, the Poiseuille flow can be used as a hallmark of viscosity.

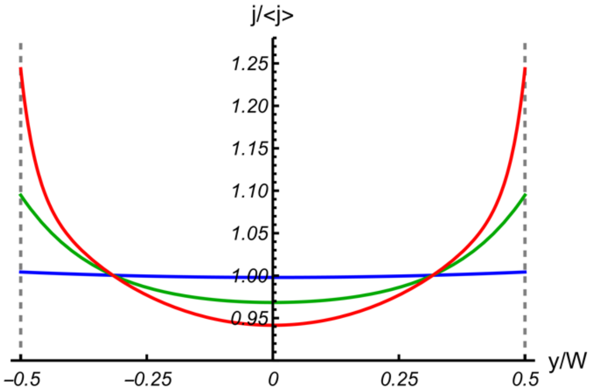


Figure 4: *Catenary curves of the current density in the narrow channel normalized by the averaged current density. The numerical results were obtained for typical parameter values for graphene (τdis ≈ 0.8 THz, αg ≈ 0.2, ν ≈ 0.4m2/s, B = 0.1T, T = 250K) and corre*

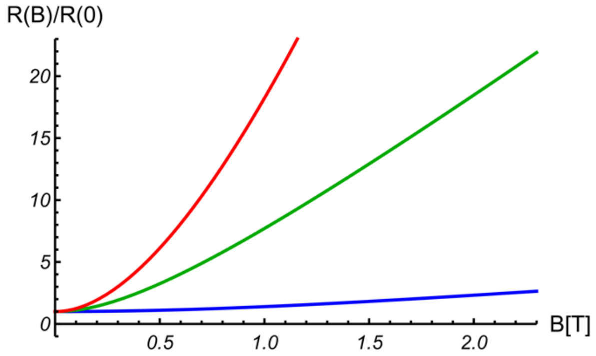


Figure 5: *Magnetoresistance in the narrow graphene channel normalized by the zero-field resistance R0. The numerical results were obtained for typical parameter values (as in Fig 4) and corresponds to three values of the channel width, W = 0.1, 1, 5 µm (blue, gr*

In this study, the behaviour of electronic transport in graphene at charge neutrality is shown to differ significantly from that of any single-component fluid, including strongly doped graphene. When the doping level is weak ($\mu \ll T$), the contribution of hydrodynamics to the electric current is small and only yields a slight correction to our previously presented results. However, for $\mu \sim T$, both the hydrodynamic and dissipative ("kinetic") contributions become equally significant. In this case, there is no small parameter in the theory, and we can represent the full system of linearized hydrodynamic equations by a 6 × 6 matrix. On the other hand, when the doping level is high ($\mu \gg T$), the hydrodynamic contribution dominates, and the boundary scattering becomes diffusive. This dominance results in the Poiseuille profile being observed in the electronic flow in a channel, which is in agreement with the experimental observations in Ref. [4] Based on these findings, the crossover from anti-Poiseuille to Poiseuille flow is anticipated to occur at $\mu \sim T$.

2.1.3 References

1. S. Danz, M. Titov, and B. N. Narozhny, Giant nonlocality in nearly compensated two-dimensional semimetals, Phys. Rev. B 102, 081114(R) (2020); https://doi.org/10.1103/PhysRevB.102.081114
2. B. N. Narozhny, I. V. Gornyi, and M. Titov. Anti-Poiseuille flow in neutral graphene. Phys. Rev. B 104, 075443 (2021); https://doi.org/10.1103/PhysRevB.104.075443
3. M. Polini and A. K. Geim, Viscous electron fluids, Physics Today 73, 28 (2020); https://doi.org/10.1063/PT.3.4497
4. Joseph A. Sulpizio, Lior Ella, Asaf Rozen, John Birkbeck, David J. Perello, Debarghya Dutta, Moshe Ben-Shalom, Takashi Taniguchi, Kenji Watanabe, Tobias Holder, Raquel Queiroz, Alessandro Principi, Ady Stern, Thomas Scaffidi, Andre K. Geim, and Shahal Ilani, Visualizing Poiseuille flow of hydrodynamic electrons, Nature 576, 75 (2019); https://doi.org/10.1038/s41586-019-1788-9

2.2 Non-linear phenomena in electron hydrodynamics

Inspired by the success of electronic hydrodynamics in the linear regime (in particular in graphene), the research presented here encapsulates the work undertaken on non-linear hydrodynamic phenomena by various research groups. The investigations span diverse non-linear phenomena, experimentally relevant geometries, and the exploration of hydrodynamic effects in the presence of weak magnetic fields. The ultimate goal remains to determine if conditions akin to turbulence in real fluids can be achieved in electronic systems. While work in this direction is not yet complete, the path ahead holds exciting potential, not only for fundamental physics but also for practical applications in nanoelectronics and related areas.

2.2.1 Transport properties of strongly coupled electron-phonon liquids

The study [1] focused on the hydrodynamic behaviour of a coupled electron-phonon fluid, specifically under the conditions of strong phonon drag. This regime occurs when the rate of phonon equilibration due to, e.g., Umklapp scattering is much slower than the rate of normal electron-phonon collisions. Then phonons and electrons form a coupled out-of-equilibrium state where the total quasi-momentum of the electron-phonon fluid is conserved. A joint flow velocity emerges as a collective hydrodynamic variable. The study utilized a rigorous approach by deriving the equation of motion for this coupled electron-phonon fluid from the underlying microscopic kinetic theory. This equation elucidated essential parameters such as effective viscosity and thermal conductivity, which are crucial for understanding and characterizing the behaviour of the system. Key findings include:

1. **Decay Times and Super-Diffusive Relaxation:** The study was particularly notable for its derivation of decay times for arbitrary harmonics of the distribution function. It revealed that these harmonics exhibit super-diffusive relaxation on the Fermi surface. This behaviour provides insights into the dynamics of electron-phonon coupling and non-equilibrium states.

2. **Applications to Magneto-Transport Properties:** The analysis was extended to various experimental geometries, including Hall-bar and Corbino-disk setups. These geometries are relevant to magnetotransport experiments. By examining electron-phonon interactions under different conditions, the study contributed to the understanding of magnetotransport properties in these systems.

3. **Consideration of General Boundary Conditions:** A broad spectrum of boundary conditions was considered, ranging from no-slip to no-stress flows. This comprehensive approach allowed them to study the system's behaviour across different regimes, providing valuable insights into the crossover from the Stokes to the Ohmic regime, especially under the conditions of the Gurzhi effect.

4. **Frequency Dependence of Surface Impedance and Non-Equilibrium Noise:** The investigation also encompassed the frequency dependence of the surface impedance and non-equilibrium noise. Of particular interest was the observation that in the diffusive regime, the electron-phonon collision integral exhibited a Burgers-type nonlinearity. This led to the formation of a shock-wave structure in the energy domain of the non-equilibrium distribution function, with implications for the Fano factor of the noise.

In summary, the study made significant contributions to the understanding of non-linear phenomena in electron hydrodynamics, especially in the context of electron-phonon coupling. Their rigorous approach, spanning both theoretical and experimental aspects, provided valuable insights into the behaviour of electron-phonon fluids and their applications in magnetotransport. The findings offer a foundation for future research in this area and have the potential to impact various fields, including materials science and electronics. In particular, the work discussed the connections and limitations of the results in the context of recent electron-phonon drag measurements in Dirac and Weyl semimetals, and layout directions for further extensions and developments.

2.2.2 Corbino field-effect transistors in a magnetic field

Ref. [2] focused on gated field-effect transistors (FETs) with an eccentric Corbino-disk geometry. Their investigation aimed to uncover non-linear electronic phenomena in these systems, particularly concerning their response to external stimuli, such as an AC THz (terahertz) potential difference applied between the source and gate, and a static source-drain voltage rectified by the nonlinearities of the FET. These unique experimental conditions allowed the group to explore and understand the non-linear behaviours exhibited by these electronic devices. When a magnetic field was applied perpendicular to the device, a strong resonance appeared at the cyclotron frequency. The strength of the resonance can be tuned by changing the eccentricity of the disk.

Key findings include:

1. **Resonance at Cyclotron Frequency:** When a magnetic field was applied perpendicular to the Corbino-disk FETs, a strong resonance at the cyclotron frequency was observed. This finding is significant because it reveals that the non-linear response of the system is strongly influenced by the magnetic field's orientation. Moreover, the strength of this resonance could be adjusted by changing the eccentricity of the Corbino-disk, underscoring the importance of geometric parameters in shaping non-linear electronic behaviour.
2. **Optimization of Responsivity:** The study showed that there exists an optimum value of eccentricity for the Corbino-disk that maximizes the responsivity of the FET. This optimization is an exciting discovery, as it suggests that fine-tuning geometric parameters can significantly enhance the device's performance. The ability to control and maximize responsivity holds potential for applications in nanoelectronics, where the efficient detection and manipulation of electronic signals are crucial.
3. **Application of Non-Linearities:** The study illuminated how non-linearities, induced by both the AC THz potential difference and static source-drain voltage, play a pivotal role in shaping the behaviour of the Corbino-disk FETs. This insight offers a more profound understanding of how non-linearities can be harnessed to create functional electronic devices with tailored characteristics.

In summary, the study made substantial contributions to the exploration of non-linear electronic phenomena in the context of Corbino-disk FETs. The work showcased the importance of magnetic fields and geometric parameters in shaping non-linear behaviour. Furthermore, the ability to optimize the responsivity of these devices opens up exciting possibilities for applications in nanoelectronics and related fields, where precise control of electronic responses is of paramount importance. This study bridges the gap between theory and practical applications, shedding light on the potential of non-linear electronic systems in emerging technologies.

2.2.3 Information-based approach to turbulence

A series of studies explored non-linear phenomena in hydrodynamics, aiming to unravel the complex behaviors and interactions that occur in systems with resonantly interacting modes, with a particular focus on noise-driven systems.

When two resonantly interacting modes are in contact with a thermostat, their statistics is exactly Gaussian and the modes are statistically independent despite the strong interaction. Considering noise-driven system, in Ref. [3] it was shown that when one mode is pumped and another dissipates, the statistics of such cascades is never close to Gaussian, no matter the interaction/noise relation. Substantial phase correlation was found in the limit of strong interaction (weak noise). Surprisingly, for both cascades, the mutual information between modes (see also Ref. [4]) increases and entropy further decreases when interaction strength decreases. The model was developed to elucidate the fundamental problems of far-from-equilibrium physics: where the information (entropy deficit) is encoded and how singular measures form. For an instability-driven system (a laser), even a small added noise leads to large fluctuations of the relative phase near the stability threshold, while far from it we show that the conversion into the second harmonic is weakly affected by noise.

Never is the difference between thermal equilibrium and turbulence so dramatic, as when a quadratic invariant makes the equilibrium statistics exactly Gaussian with independently fluctuating modes. That happens in two very different yet deeply connected classes of systems: incompressible hydrodynamics and resonantly interacting waves. In Ref. [5] the first case of a detailed information-theoretic analysis of turbulence in such strongly interacting systems was presented. The analysis elucidated the fundamental roles of space and time in setting the cascade direction and the changes in the statistics along it. A simple yet rich family of discrete models with neighbouring triplet interactions was introduced. It was shown that it had families of quadratic conservation laws defined by the Fibonacci numbers. Depending on the single model parameter, three types of turbulence were found: single direct cascade, double cascade, and the first-ever case of a single inverse cascade. The way deviation from thermal equilibrium all the way to turbulent cascades rendered statistics increasingly non-Gaussian was described quantitatively and it was found the self-similar form of the one-mode probability distribution. In particular, it was revealed where the information (entropy deficit) is encoded and disentangled in the communication channels between modes, as quantified by the mutual information in pairs and the interaction information inside triplets.

Key findings include:

1. **Resonantly Interacting Modes:** The team investigated systems where modes resonantly interact with each other. In such systems, it was found that their statistics were exactly Gaussian, despite the strong interactions between the modes. This result challenges conventional expectations regarding statistical behaviours in interacting systems and highlights the complexity of noise-driven phenomena.
2. **Phase Correlation and Mutual Information:** The studies revealed substantial phase correlation in the limit of strong interaction, even in the presence of weak noise. This

discovery underscores the intricate relationship between mode interactions and their statistical properties. Moreover, the mutual information between modes was observed to increase as the strength of interaction decreased, providing insights into how information is encoded and distributed in these systems.

3. **Insights into Far-From-Equilibrium Physics:** The work addressed the fundamental problem of far-from-equilibrium physics, and investigated how information (entropy deficit) is encoded and how singular measures form in noise-driven systems. This knowledge has broad implications, particularly in understanding the behaviour of systems under various conditions, including instability-driven systems like lasers.
4. **Analysis of Turbulence:** In one of the studies, an information-theoretic analysis of turbulence in strongly interacting systems, such as incompressible hydrodynamics and resonantly interacting waves, was performed. The roles of space and time in setting the cascade direction and the changes in statistical properties along it was explored. This analysis led to the discovery of different types of turbulence, including single direct cascades, double cascades, and a unique single inverse cascade. The research also provided insights into the information encoding and distribution within these turbulent systems.

In summary, the studies have made significant contributions to the understanding of non-linear phenomena in hydrodynamics. The work offers insights into the intricate interplay of modes in noise-driven systems and the statistical behaviours that arise from resonantly interacting modes. These findings have broad implications for understanding far-from-equilibrium physics and the behaviour of various systems, including turbulent flows. The complexity of the behaviours observed underscores the richness of non-linear phenomena in electron hydrodynamics, paving the way for further exploration in this exciting field.

2.2.4 References

1. Alex Levchenko, Jörg Schmalian, Transport properties of strongly coupled electron–phonon liquids, Annals of Physics 419, 168218 (2020); https://doi.org/10.1016/j.aop.2020.168218
2. Bailey Winstanley, Henning Schomerus, and Alessandro Principi, Corbino field-effect transistors in a magnetic field: Highly tunable photodetectors, Phys. Rev. B 104, 165406 (2021); https://doi.org/10.1103/PhysRevB.104.165406
3. N. Vladimirova, M. Shavit, S. Belan, and G. Falkovich, Second-harmonic generation as a minimal model of turbulence, Phys. Rev. E 104, 014129 (2021); https://doi.org/10.1103/PhysRevE.104.014129
4. Michal Shavit and Gregory Falkovich, Singular Measures and Information Capacity of Turbulent Cascades, Phys. Rev. Lett. 125, 104501 (2020); https://doi.org/10.1103/PhysRevLett.125.104501
5. Natalia Vladimirova, Michal Shavit, and Gregory Falkovich, Fibonacci Turbulence, Phys. Rev. X 11, 021063 (2021); https://doi.org/10.1103/PhysRevX.11.021063

2.3 Light-matter interaction

Advances harnessing light-matter interaction are expanding the frontiers of the materials research landscape. Over the last years, a steady series of studies has been changing our understanding of light-matter interaction and its possible reach in the world of materials. Besides serving as a prime characterization tool (with ARPES as a flagship example key in the discovery of topological materials), laser illumination can now be used to transform an otherwise metallic sample onto one endowed with a gap and even topological states (for a recent review see [1]).

Floquet-engineering, the use of light-matter interaction (or more generally, a time-dependent driving) to tailor the properties of materials or systems, has gained maturity over the last decade. The name Floquet is ubiquitous in these systems and signals that we are in the realm of the so-called Floquet theory, the prevalent non-adiabatic and non-perturbative [2] used for this type of driven systems, a temporal analogue of Bloch's theory.

A milestone has been the prediction of laser-induced bandgaps in graphene [3], which was shown to have a favourable spot for experimental observation for laser frequencies in the mid-infrared spectrum [4]. Since 2013, time-resolved ARPES experiments [5, 6] have validated two essential findings: the formation of Floquet-Bloch states, which are hybrid electron-photon states, and the emergence of polarization-dependent bandgaps. The mentioned time-resolved ARPES experiments were carried out at the surface of a three-dimensional topological insulator, in monolayer WS_2 and also more recently, in graphene [7]. Other theoretical studies have also suggested Floquet topological states which have been demonstrated experimentally in laser-illuminated graphene [8].

2.3.1 2D materials under periodic driving

In Ref. [9], the impact of circularly polarized laser illumination on AB-stacked graphite samples' electronic structure was studied. The results suggest potential for implementing a three-dimensional quantum Hall effect in Floquet systems.

Strong light-matter interaction effects on topological phases were also studied. This included graphene in the quantum Hall regime and a two-dimensional material in the quantum-spin-Hall phase.

The first study investigated intense laser illumination's effect on graphene's quantum Hall effect [10]. Using Floquet theory with low energy description and full tight-binding models, selection rules, quasienergy band structure, and their connection to the conductance in experimental device setups, were derived. The findings show that dynamical gaps in the Floquet spectrum switch off quantum Hall edge transport for various edge terminations, except for armchair. Near the Dirac point, laser polarization controls Hall conductance, allowing it to be switched on/off or have its sign flipped. These results provide a new way to manipulate the quantum Hall effect.

In the second study, Floquet scattering theory was combined with atomistic models to examine laser illumination, spin, and topology interplay in a two-dimensional material with spin-orbit coupling [11]. Starting from a topological phase, laser illumination can selectively disrupt topological edge states based on their spin. The study observed pure spin photocurrents and spin-polarized charge photocurrents under linear and circular polarized laser illumination.

These findings offer a method to generate and control spin-polarized photocurrents, with implications for future topological devices.

In all theese studies a key part is the calculation of the Floquet spectrum in such systems. This can be accessed experimentally with time-resolved ARPES techniques and has been the focus of great interest in the last years. Furthermore, this is the basis for the analysis of many transport experiments, which could be simulated with similar techniques. A Python code was made available on Zenodo. The code enables the calculation of the Floquet spectrum for laser irradiated graphene and is crucial for the research mentioned before. The code can be found at https://zenodo.org/records/12809644.

2.3.2 References

1. Mark S. Rudner and Netanel H. Lindner, Band structure engineering and non-equilibrium dynamics in Floquet topological insulators, Nature Reviews Physics 2, 229 (2020); https://doi.org/10.1038/s42254-020-0170-z
2. Sigmund Kohler, Jörg Lehmann, Peter Hänggi, Driven quantum transport on the nanoscale, Physics Reports 406, 379 (2005); https://doi.org/10.1016/j.physrep.2004.11.002
3. Takashi Oka and Hideo Aoki, Photovoltaic Hall effect in graphene, Phys. Rev. B **79**, 081406 (2009); https://doi.org/10.1103/PhysRevB.79.081406
4. Hernán L. Calvo, Horacio M. Pastawski, Stephan Roche, and Luis E. F. Foa Torres, Tuning laser-induced band gaps in graphene, Appl. Phys. Lett. **98**, 232103 (2011); https://doi.org/10.1063/1.3597412
5. Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik, Observation of Floquet-Bloch States on the Surface of a Topological Insulator, Science 342, 453 (2013); https://doi.org/10.1126/science.1239834
6. Edbert J. Sie, James W. McIver, Yi-Hsien Lee, Liang Fu, Jing Kong, and Nuh Gedik, Valley-selective optical Stark effect in monolayer WS2, Nature Materials **14**, 290 (2015); https://doi.org/10.1038/nmat4156
7. Dongsung Choi, Masataka Mogi, Umberto De Giovannini, Doron Azoury, Baiqing Lv, Yifan Su, Hannes Hübener, Angel Rubio, and Nuh Gedik, Direct observation of Floquet-Bloch states in monolayer graphene, arXiv:2404.14392; https://doi.org/10.48550/arXiv.2404.14392
8. J. W. McIver, B. Schulte, F.-U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, Light-induced anomalous Hall effect in graphene, Nature Physics **16**, 38 (2020); https://doi.org/10.1038/s41567-019-0698-y
9. Hernán L. Calvo, Jose E. Barrios Vargas, and Luis E. F. Foa Torres, Floquet boundary states in AB-stacked graphite, Phys. Rev. B **101**, 075424 (2020); https://doi.org/10.1103/PhysRevB.101.075424
10. A. Huamán, L. E. F. Foa Torres, C. A. Balseiro, and Gonzalo Usaj, Quantum Hall edge states under periodic driving: A Floquet induced chirality switch, Phys. Rev. Research **3**, 013201 (2021); https://doi.org/10.1103/PhysRevResearch.3.013201

11. Matías Berdakin*, Esteban A. Rodríguez-Mena, Luis E. F. Foa Torres, Spin-Polarized Tunable Photocurrents, Nano Lett., **21**, 3177 (2021); https://doi.org/10.1021/acs.nanolett.1c00420

3. Emergent phenomena in van der Waals heterostructures

3.1 From hydrodynamics to spin-charge coupling

Van der Waals heterostructures enable precise tuning of electronic properties through modifications of the local environment. In graphene, electron interactions depend strongly on dielectric screening from encapsulants, nearby gates, or adjacent graphene layers. By controlling the twist angle, the system can be tuned between Fermi and Dirac liquid regimes. The hydrodynamic transport regime, arising from electron-electron collisions, is thus highly sensitive to the local environment. For instance, gates placed within ~1 nm of graphene drastically alter viscosity's dependence on carrier density due to screened electron interactions. Furthermore, in two-dimensional conductors, proximity between electron- and hole-doped regions induces Coulomb-mediated electron-hole drag, studied extensively in GaAs and graphene. At reduced interlayer distances, these interactions strengthen, reaching their ultimate regime when electrons and holes coexist in a single layer. Graphene at the charge neutrality point is ideal for such studies, as thermally excited electron-hole pairs can be precisely tuned via gating.

Recent research has also focused on hybrid charge-spin dynamics in van der Waals materials with strong spin-orbit coupling and magnetic order, aiming to replace conventional electronics with low-heat alternatives. While metallic thin films exhibit diffusive transport due to impurity/phonon scattering, van der Waals heterostructures allow exploration of a regime dominated by electron-electron collisions. For example, graphene encapsulated in transition-metal dichalcogenides or magnetic halides combines strong spin-orbit coupling with magnetization coupling, enabling rich phenomena like anomalous and spin-Hall effects. Here, itinerant and localized electrons interact as coupled fluids, generating transverse charge/spin currents while exerting spin torques on the magnetization. The role of electron-electron interactions in these processes remains unexplored.

3.1.1 Electron hydrodynamics in the presence of nearby gates

The interaction amongst electrons plays a fundamental role in defining several important phenomena in solid-state systems, such as magnetism, superconductivity, hydrodynamic transport, etc. By finding ways to control its strength, one can in turn affect the properties of the condensed-matter system under study. One way to do this is to screen the electron-electron interaction by placing a given system in proximity to a metal. In this way, interactions are screened and hence suppressed compared to a system in dielectric matrix.

In Ref. [1], the electron-electron scattering length in graphene in the presence of nearby atomically-thin dielectrics and atomically-flat metallic gates was extracted from measurements of the electron viscosity. Qualitative deviations from the conventional behaviour were observed. Screening effects became important only when the thicknesses of the dielectric separating the graphene sheet from the metal was reduced to a few nm. The condition for the metal screening to dominate is for this distance to become much smaller than the typical separation between electrons, *i.e.* the intrinsic Thomas-Fermi screening wavelength. The scattering rates extracted from measurements of electron viscosity agreed well with the theoretical predictions.

The devices studied in Ref. [1] were graphene monolayers encapsulated between hexagonal boron nitride (hBN) crystals. Graphite monocrystals served as bottom (screening) gates, separated from the graphene layer by ultra-thin hexagonal boron nitride (hBN) dielectrics. The minimum thickness used for the dielectric was ∼1.3nm. Thinner crystals exhibited notable electron tunnelling. An extra metal gate was deposited on top of the heterostructures to tune the carrier density in the graphene sheet. The low-temperature mobility of the devices was generally very high, and the resistivity was found to be independent of dielectric thickness. This ensured that the reported behaviour of screening of electron-electron interactions, measured from the electron viscosity, was due to changes of distance rather than of the transport properties of the system.

For known resistivity and momentum-non-conserving scattering times, it is possible to convert the measured resistances (vicinity, point-contact and in particular viscous-Hall) into the electron-electron scattering length. For all temperatures, screened devices displayed scattering lengths approximately twice longer than those of standard devices with further gates. This agreed well with theory. Furthermore, the screening due to metallic gate in proximity to the graphene sheet qualitatively changed the dependence of the scattering length of density. While in conventional devices the scattering length increases weakly with density away from the neutrality point, a totally opposite behaviour was found. Because of the extra screening, the scattering length becomes a decreasing function of carrier density.

To explain the observed dependences of the scattering length of carrier density and distance from the gate, numerical calculations of the electron self-energy due to dynamically screened electron-electron interactions were carried out. The self-energy can be easily converted into a scattering time, and thus in a length via the Fermi velocity. In these calculations, the gate was modelled as a perfect conductor, thus nearby charges were screened by the presence of image charges. Accounting for the finite density of carriers in the metal only provided small departures from this model. The results were found to be in good agreement with experimental data. No fitting parameters were used.

3.1.2 Electron-hole hydrodynamics

The electron-hole plasma existing in graphene at the neutrality point, commonly referred to as a Dirac fluid, is a quantum-critical system characterized by strong inter-carrier scattering dominated by Planckian dissipation. This makes graphene's Dirac plasma not only a tuneable and conceptually simple system for investigating electron transport but also an analogue for understanding more complex quantum-critical phenomena, such as the transport properties of "strange metals" and high-temperature superconductors in their normal state. The Dirac plasma also provides conceptual parallels to relativistic electron-positron plasmas found in astrophysical contexts, which are otherwise challenging to replicate in laboratory conditions.

Previous experimental studies of the Dirac plasma revealed hallmarks of quantum-critical transport, such as hydrodynamic flow, violations of the Wiedemann-Franz law, and anomalous magnetoresistance. However, until recently, the mutual drag between electron and hole subsystems within the Dirac plasma remained largely unexplored.

In Ref. [2], the first experimental measurements of electron-hole drag within the Planckian plasma formed in monolayer graphene near its Dirac point at temperatures above the liquid-nitrogen range were reported. Frequent electron-hole scattering leads minority carriers to drift against the applied electric field, driven by drag from majority carriers. This phenomenon results in an effective negative mobility for minority carriers. Notably, the strength of the electron-hole drag peaked near room temperature, despite significant contributions from phonon scattering.

The transport properties of graphene near the NP cannot be fully understood without accounting for the strong interactions between electron and hole subsystems within the Dirac plasma. Thorough analysis revealed that minority carriers are dragged along by majority carriers, a behaviour accurately captured by our Boltzmann framework, which enabled quantitative determination of scattering rates. Phonon and impurity scattering play a critical role in limiting the hydrodynamic behaviour of graphene, reducing the ratio of momentum-conserving to momentum-relaxing collision times. Consequently, minority carriers consistently lag behind majority carriers in the Dirac plasma.

For encapsulated high-quality graphene, mutual drag is most pronounced at room temperature, where minority carriers drift at approximately half the velocity of majority carriers. These findings emphasize the influence of impurity and phonon scattering on the transport properties of graphene's Dirac plasma, demonstrating their suppressive effect on its viscous behaviour. These results not only advance the understanding of graphene's transport phenomena but also provide a framework for exploring similar effects in other quantum-critical systems.

3.1.3 Gilbert damping in two-dimensional metallic anti-ferromagnets

The dynamics of magnetization in anti-ferromagnets remain a focal point of interest for the various potential applications. Numerous ideas explore the potential for THz frequency manipulation of anti-ferromagnetic domains to enable ultrafast information storage and beyond-CMOS computation. The emergence of van der Waals magnets has further impacted this field, introducing the prospect of developing adjustable heterostructures involving anti-ferromagnetic and semiconducting layers.

The understanding of relaxation processes of both the Neel vector and non-equilibrium magnetization within anti-ferromagnets is widely acknowledged as pivotal for the optimal functioning of spintronic devices. On one hand, a low level of Gilbert damping is generally expected to enhance the electrical control of magnetic order through mechanisms like domain wall motion or ultrafast domain switching. On the other hand, the effective manipulation of magnetic domains typically necessitates a robust coupling between charge and spin degrees of freedom, attributed to a pronounced spin-orbit interaction, which is commonly regarded as analogous to large Gilbert damping.

The focus has been directed towards a microscopic analysis of Gilbert damping, emphasizing the Dyakonov-Perel and Elliot-Yafet mechanisms. This theoretical framework has been applied to a model representing a two-dimensional Neel anti-ferromagnet characterized by a honeycomb magnetic lattice. In recent investigations, density-functional-theory calculations have been employed to explore single-layer transition metal trichalcogenides. These analyses have foreseen the emergence of numerous metallic anti-ferromagnetic materials featuring a honeycomb lattice and diverse magnetic orderings. Furthermore, predictions indicate the potential induction of anti-ferromagnetism in graphene through its proximity to two-dimensional magnets. This evolving understanding of magnetization dynamics in anti-ferromagnetic systems, combined with advancements in theoretical models and material predictions, propels ongoing explorations for applications in ultrafast information storage and spintronic devices.

Motivated in part by these forecasts and recent strides in producing single-layer anti-ferromagnet crystals, In Ref. [3] an effective model was proposed to examine spin relaxation in a 2D honeycomb anti-ferromagnet with Neel magnetic order. Within this framework, three regimes are predicted, each exhibiting a qualitatively distinct dependence of Gilbert damping on spin-orbit interaction and conduction electron transport time.

In the regime characterized by weak spin-orbit interaction, the dominant factor is the exchange field relaxation of electron spin. The regime featuring moderate spin-orbit strength is governed by Elliot-Yafet spin relaxation. Transitioning into the realm of strong spin-orbit strength, where significant splitting of electron Fermi surfaces occurs, Dyakonov-Perel relaxation of the in-plane spin component takes precedence, coupled with Elliot-Yafet relaxation of the perpendicular-to-the-plane Gilbert damping, resulting in a pronounced damping anisotropy. Isotropic Gilbert damping is reinstated only for finite magnon wave vectors, where the magnon wavelength is smaller than the spin-orbit length. This model contributes to the ongoing exploration of spin dynamics in 2D anti-ferromagnetic systems, adding valuable insights into potential applications in spintronic technologies.

3.1.4 Interactions and topology: the case of massive Dirac fermions

The anomalous Hall effect is a phenomenon observed in condensed matter physics, manifesting as a transverse voltage that arises perpendicular to the applied electric current. Unlike the ordinary Hall effect, which can be explained by the Lorentz force acting on charge carriers due to an applied magnetic field, the anomalous Hall effect is an intriguing quantum mechanical phenomenon that is a result of the combined effects of the intrinsic spin-orbit interaction of the material and its magnetisation. The phenomenon is therefore commonly found in ferromagnetic and strongly correlated electron systems. The presence of a net magnetization in these systems is crucial for the emergence of the anomalous Hall effect. Understanding the anomalous Hall effect is of great importance in both fundamental physics and practical applications. It provides valuable insights into the electronic and magnetic properties of materials, offering a window into the intricate interplay between spin, charge, and orbital degrees of freedom. So far, however, studies have focused on the non-interacting regime. We have been the first to address the role of interactions, to first order in the coupling constant, in Ref. [4].

In a seminal work, Coleman and Hill demonstrated that the anomalous Hall conductivity remains unaltered by two-particle interactions at zero temperature when the Fermi energy is situated within the bulk band gap. This prompted our inquiry into the resilience of the Hall conductivity to interaction effects in systems characterized by finite temperature or chemical potential. To address this question, an examination was conducted on the influence of electron-electron interactions in the prototypical model of the anomalous Hall effect, specifically, a two-dimensional assembly of massive Dirac fermions.

Beyond its foundational implications, this matter also carries practical relevance as real-world materials typically exist at non-zero temperatures and seldom lack doping entirely. It has been established that specific many-body interactions, such as those involving electrons and quenched disorder, can significantly impact Hall responses. In scenarios like the anomalous and spin-Hall effects, the introduction of quenched disorder can yield results notably divergent from non-interacting outcomes. These variations encompass an accelerated decay of Hall responses with increasing chemical potential, alterations in sign, and, in certain instances, the outright elimination of these effects. Although refining the growth process may help mitigate disorder in a material, the omnipresent e-e interactions are challenging to eliminate. Consequently, comprehending their impact on the Hall response becomes crucial for predicting and elucidating experimental findings.

The correction to the anomalous Hall effect of massive Dirac fermions was investigated to first order in the strength of electron-electron interactions. It was observed that the initial response diverged, necessitating the removal of divergences through the renormalization of the bare parameters within the model. Consequently, a finite expression for the first-order correction to the anomalous Hall conductivity was derived. The dependence of the first-order correction on the form of the e-e interaction was explored, addressing both the case of a contact potential and an unscreened Coulomb potential. Despite their distinct characteristics, both cases exhibited similar behaviour, suggesting the robustness and generality of the results to interacting systems.

It was demonstrated that interaction corrections could be substantial enough to counterbalance the non-interacting contribution to the Hall conductivity. Specifically, it was shown that the anomalous Hall conductivity could exhibit a faster decay with chemical potential than predicted in the non-interacting scenario. Moreover, the potential for a change in sign at certain chemical potential values was identified. To exercise caution, this latter outcome was attributed to a potential breakdown in the perturbative result. Subsequent investigations are underway to explore additional corrections and ascertain whether the observed effect dissipates.

3.1.5 Non-conserved density accumulations and generalised Hall effects

Refs. [5,6] explored density accumulations in anomalous transport phenomena, specifically, the valley Hall effect (VHE) and the orbital Hall effect (OHE), focusing on how nonconserved densities (valley, orbital magnetization, etc.) accumulate at the edges of materials under an applied electric field. Ref. [1] examines the VHE in multivalley insulators (both trivial and topological) as well as metals, while Ref. [2] generalizes these findings to a broader class of anomalous effects.

Ref. [5] solved the "valley Hall effect puzzle". the VHE arises in systems with multiple valleys (e.g., graphene/hBN heterostructures), where electrons from different valleys experience opposite Berry curvatures, leading to a transverse valley current but no net charge current. Conventional wisdom suggested that this bulk valley current should produce valley density accumulation at the edges, yet experiments and theory disagreed on whether such accumulations occur in fully gapped insulators. This puzzle was resolved by showing that valley density is not conserved in the presence of an electric field. While the electric field drives a bulk valley Hall current, it also causes an intrinsic decay of valley-density, effectively "short-circuiting" the valley current before it reaches the edges. Consequently, in fully gapped, time-reversal (TR) invariant insulators, no valley accumulation appears at the edges, making the VHE unobservable in such systems. However, in metallic or partially gapped systems, where states at the Fermi level allow dissipation, valley accumulation can occur. A Fermi-surface formula for this accumulation is derived and used to show that in a graphene nanoribbon certain edge terminations can produce a net valley polarization.

Ref. [6] derived a unified theory for anomalous transport, extending these insights to general nonconserved densities (e.g., orbital magnetization, spin, valley) in both insulators and metals. The key distinction lies in whether the system is TR-invariant or not:

- **TR-Invariant systems**: there are no undergap contributions to the density accumulation. In insulators, bulk currents from filled bands (undergap currents) do not lead to edge accumulations unless there are dissipative processes (i.e., states at the Fermi level). Therefore, dissipation is necessary: accumulations arise only in metals or edge-metallic systems where dissipation occurs. A "no-dissipation no-accumulation theorem" is proven in Ref. [6].
- **TR-Broken systems**: nondissipative accumulations becomes possible. Even in insulators, bulk effects like magnetoelectric polarization can induce edge accumulations without requiring dissipation.

Ref. [6] also introduces the concept of net torque, a global nonconservation effect where the integrated density accumulation does not vanish, implying a net imbalance in the system. This contrasts with the distributed torque, which averages to zero and can be absorbed into a redefined current.

These results clarified longstanding controversies in experiments on graphene/hBN and transition metal dichalcogenides, where valley or orbital accumulations were debated. The theory also provided a framework for designing devices that exploit these effects, e.g., by breaking valley degeneracy or using metallic systems where dissipative accumulations are possible.

3.1.6 References

1. M. Kim, M., S.G. Xu, A.I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, R. V. Gorbachev, K. Watanabe, T. Taniguchi, I. V. Grigorieva, V. I. Fal’ko, M. Polini, and A. K. Geim, Control of electron-electron interaction in graphene by proximity screening, Nature Communications **11**, 2339 (2020); https://doi.org/10.1038/s41467-020-15829-1

2. Leonid A. Ponomarenko, Alessandro Principi, Andy D. Niblett, Wendong Wang, Roman V. Gorbachev, Piranavan Kumaravadivel, Alexey I. Berdyugin, Alexey V. Ermakov, Sergey Slizovskiy, Kenji Watanabe, Takashi Taniguchi, Qi Ge, Vladimir I. Fal’ko, Laurence Eaves, Mark T. Greenaway and Andre K. Geim. Extreme electron–hole drag and negative mobility in the Dirac plasma of graphene, Nature Communications **15**, 9869 (2024); https://doi.org/10.1038/s41467-024-54198-x

3. Robert Sokolewicz, Mikhail Baglai, Ivan Ado, Mikhail Katsnelson, Mikhail Titov, Gilbert damping in two-dimensional metallic anti-ferromagnets, Phys. Rev. B 109, 134427 (2024); https://doi.org/10.1103/PhysRevB.109.134427

4. Daria A. Dumitriu-I., Darius A. Deaconu, Alexander E. Kazantsev, Alessandro Principi, First-order effect of electron-electron interactions on the anomalous Hall conductivity of massive Dirac fermions, Phys. Rev. B 109, 165429 (2024); https://doi.org/10.1103/PhysRevB.109.165429

5. Alexander Kazantsev, Amelia Mills, Eoin O’Neill, Hao Sun, Giovanni Vignale, and Alessandro Principi, Nonconservation of the Valley Density and Its Implications for the Observation of the Valley Hall Effect, Phys. Rev. Lett. 132, 106301 (2024); https://doi.org/10.1103/PhysRevLett.132.106301
6. Hao Sun, Alexander Kazantsev, Alessandro Principi, and Giovanni Vignale, Nonconserved density accumulations in orbital Hall transport: Insights from linear response theory, Phys. Rev. B 111, 075432 (2025); https://doi.org/10.1103/PhysRevB.111.075432

3.2 Electronic phenomena in twisted bilayer graphene and other 2D materials

Twisted bilayer graphene (TBG) offers a versatile platform for exploring diverse electronic phenomena due to its unique moiré superlattice. Twist-angle fluctuations, particularly near the magic angle, introduce a soft degree of freedom that strongly couples to electronic states, influencing orbital ferromagnetism. The study presented below reveals that these fluctuations induce damped phasons, enhancing ferromagnetic order at zero temperature while suppressing it at finite temperatures due to interactions with soft phonons.

Plasmonic properties in TBG exhibit a transition from strong coupling at small twist angles to weak coupling at larger angles, where an acoustic plasmon mode emerges. This tuneability suggests applications in mid-infrared nanophotonics and cavity QED, with potential implications for superconductivity and polariton physics. Electronic correlations in magic-angle TBG reveal that Hartree interactions critically influence band structure, compressibility, and paramagnetic state stability near charge neutrality. Beyond the magic angle, TBG can be tuned across distinct regimes—from Fermi liquid to Dirac liquid to compensated semimetal—demonstrating the system’s hydrodynamic versatility. Measurements of hydrodynamic transport in these regimes align closely with microscopic models, offering insights into phenomena like the violation of the Wiedemann-Franz law.

Chirality-induced spin selectivity has also been investigated in atomically thin chiral crystals formed via van der Waals assembly. Structural chirality combined with spin-orbit coupling generates significant spin polarization, highlighting the twisted materials' potential as a tuneable platform for spin-related phenomena relevant to condensed matter physics and chiral chemistry.

3.2.1 Fracton-elasticity duality in twisted moiré superlattices

The dynamic coupling of magnetic collective modes and damped phasons in incommensurate twisted bilayer graphene was studied in Ref. [1]. Twist-angle fluctuations are a natural soft degree of freedom in addition to those found in commensurate crystals, that affect the elasticity of moiré structures, such as twisted bilayer graphene. Fluctuations and inhomogeneities in the twist angle dramatically alter the electronic properties of moiré systems, including the focus of this paper, orbital ferromagnetism. Therefore, for systems near the magic angle, a strong coupling of twist fluctuations to electronic degrees of freedom is expected. The collective modes which are a consequence of twist-fluctuations are not associated with any conservation law due to the adhesion potential between layers and are therefore damped phasons rather than phonons. Such fluctuations induce quantum fluctuations of the Ising orbital ferromagnetism observed in twisted bilayer graphene. In particular, a coupling between the twist angle and orbital magnetic moment reinforces ferromagnetic order at zero temperature, however soft phonons suppress order at finite temperatures, such that coupling decreases the temperature necessary to transition between the ordered and disordered states. The strength of damping of twist modes suppresses order in both regimes. Finally, the connection between such incommensurate fluctuations and low hydrodynamic response of defects of the incommensurate crystal was established.

3.2.2 Weak-coupling theory and plasmons in magic-angle twisted bilayer graphene

Employing a time-dependent Hartree (or GW) approximation Ref. [2] explored strong electronic correlations in magic-angle twisted bilayer graphene. They focused on the situation in which the octet of flat moiré minibands near charge neutrality are partially occupied. The study is motivated by the observation of the fundamental role played by Hartree electrostatic interactions in determining the band structure's filling-factor dependence. It was found that the electronic compressibility is primarily influenced by Hartree interactions, paramagnetic states are stable near charge neutrality, and the energy dependence on flavour polarization is overestimated by mean-field theory.

In Ref. [3] a theoretical study of intrinsic plasmonic properties of twisted bilayer graphene (TBG) as a function of the twist angle θ and other microscopic parameters such as temperature and filling factor. The calculations, based on the random phase approximation, incorporate four critical effects: the layer-pseudospin degree of freedom, spatial non-locality of the density-density response function, crystalline local field effects, and Hartree self-consistency. The findings reveal that the plasmonic spectrum of TBG transitions smoothly from a strongly-coupled regime at small twist angles ($\theta \lesssim 2°$), where the low-energy spectrum is dominated by a weakly dispersive intra-band plasmon, to a weakly-coupled regime at larger twist angles ($\theta \gtrsim 2°$), where an acoustic plasmon mode clearly emerges. This crossover enables the realization of tuneable mid-infrared sub-wavelength cavities, potentially useful for manipulating the ground state of strongly correlated electron systems.

Parallel two-dimensional electron systems (P2DESs), including those based on GaAs/AlGaAs heterostructures and atomically thin 2D materials like graphene, have been explored for phenomena such as Coulomb drag, exciton superfluidity, and broken-symmetry states driven by strong electron-electron interactions. The discovery of correlated insulators and superconductors in TBG has further enriched the many-body physics of P2DESs. TBG, comprising two graphene sheets separated by a vertical distance of approximately 0.3 nm and rotated by a twist angle θ, exhibits significant spectral changes at a small, magic angle (~1.1°). At this angle, the Brillouin zone is characterized by weakly dispersing flat bands, enhancing electron-electron interactions and leading to intriguing many-body phenomena.

The study addressed the placement of TBG within the broader context of P2DESs. Unlike GaAs double quantum wells with a constant tunneling parameter ΔSAS, TBG features spatially modulated inter-layer tunnelling on the moiré superlattice scale. Despite TBG's two-layer structure, theoretical calculations at small twist angles typically reveal only one low-energy plasmon mode, raising questions about the acoustic plasmon mode's presence. To find an intrinsic acoustic plasmon in TBG, it is essential to account for the layer-pseudospin degree of freedom, spatial non-locality, Hartree self-consistency, and crystalline local field effects. The theoretical treatment shows that at small twist angles, TBG behaves as a single entity with a centre-of-mass mode. As the twist angle increases, the layer-pseudospin becomes a quasi-good quantum number, and the plasmonic spectrum exhibits a weakly-damped acoustic plasmon mode, similar to those in other P2DESs.

These results have significant implications for both fundamental and applied research. Plasmons in TBG are potential candidates for explaining superconductivity and enriching the polariton landscape with ultra-slow acoustic plasmons. These plasmons, confined between the two graphene layers, could have important applications in quantum nanophotonics and cavity QED of strongly correlated electron systems. Future research would involve integrating these findings into Eliashberg theory for plasmon-mediated superconductivity in TBG and investigating the spatial distribution of chirality associated with this mode.

3.2.3 Twisted bilayer graphene in the hydrodynamic regime

The ambipolar hydrodynamics in the degenerate regime had long remained inaccessible, due to the fact that an electron-hole system in mono- and bilayer graphene can only be realized through the smearing of the charge neutrality point. In Ref. [4], biased small-angle twisted bilayer graphene (SA-TBG) was shown to be a system in which it is possible to explore the crossover between the Dirac fluid, Fermi liquid and compensated semimetal regimes. In the latter, the system is a degenerate two-component Fermi liquid. *I.e.* electrons and holes form degenerate Fermi liquids while the system keeps an overall charge neutrality. In this case, frequent momentum-conserving (yet velocity-relaxing) electron-hole collisions become the limiting factor for the SA-TBG conductivity.

The single-particle band structure of SA-TBG, folded within a reduced Brillouin zone because of superlattice periodicity, resembles that of monolayer graphene. The band structure exhibits two valleys located at the corners of the hexagonal mini-Brillouin zone, which coincide with the K-points of the two decoupled graphene sheets. However, the Fermi velocity is decreased compared to monolayer graphene. A feature of SA-TBG fundamental for this study is the capability of selectively populate its valleys with charge carriers of opposite types using a perpendicular displacement field. At relatively large, but still experimentally attainable fields, each valley can exhibit an electron or hole Fermi surface, the Fermi temperature of which exceeds the room temperature. At zero displacement field, instead, SA-TBG exhibits a high-temperature Dirac fluid with point-like Fermi surfaces, as in conventional monolayer graphene. This tuneability enables the exploration of various different phases, and the mapping of the entire phase diagram of this system.

To probe the phase diagram, dual-gated multi-terminal Hall bar devices were fabricated, made out of bilayer graphene twisted at about 1.65 degrees. The twisted graphene was produced with tear-and stack techniques and encapsulated between two relatively thin slabs of hexagonal boron nitride. The longitudinal resistance was measured as a function of top and bottom gates. At zero displacement field the longitudinal resistivity exhibited a peak at charge neutrality. Its behaviour changed dramatically at finite displacement field. The longitudinal resistivity at the neutrality point dropped by more than an order of magnitude and became comparable to that of doped SA-TBG. Furthermore, the longitudinal resistivity became a growing, approximately quadratic function of the temperature.

The quadratic growth of the resistivity in compensated SA-TBG at finite displacement was not observed previously. This effect is shown to be due to the friction between electron and hole fluids by solving the steady-state Boltzmann equation for the electron-hole mixture. Comparing theoretical predictions with the experimental results a very good agreement was found.

3.2.4 Chirality-induced spin polarization in twisted transition metal dichalcogenides

In Ref. [5], electron transport through atomically-thin chiral crystals formed via van der Waals assembly was studied. This study is relevant to the chirality-induced spin selectivity effect, where electrons acquire significant spin polarization. Although this effect is of interest in chiral chemistry, its mechanisms remain elusive. This effect can be remarkably large in systems with just two monolayers, provided they have spin-orbit coupling. The large polarization arises from structural chirality combined with spin-flipping spin-orbit coupling. Detailed calculations for twisted homobilayer transition metal dichalcogenides show significant polarization, exceeding

50% for MoTe2. These findings suggest that twisted quantum materials offer a tuneable platform for studying and controlling chirality-induced spin effects in both condensed matter physics and chiral chemistry.

3.2.5 References

1. Jonas Gaa, Grgur Palle, Rafael M. Fernandes, and Jörg Schmalian, Fracton-elasticity duality in twisted moiré superlattices, Phys. Rev. B 104, 064109 (2021); https://doi.org/10.1103/PhysRevB.104.064109
2. Jihang Zhu, Iacopo Torre, Marco Polini, and Allan H. MacDonald, Weak-Coupling Theory of Magic-Angle Twisted Bilayer Graphene, Phys. Rev. B 110, L121117 (2024); https://doi.org/10.1103/PhysRevB.110.L121117
3. Lorenzo Cavicchi, Iacopo Torre, Pablo Jarillo-Herrero, Frank H. L. Koppens, and Marco Polini, Theory of intrinsic acoustic plasmons in twisted bilayer graphene, Phys. Rev. B 110, 045431 (2024); https://doi.org/10.1103/PhysRevB.110.045431
4. D. A. Bandurin, A. Principi, I. Y. Phinney, T. Taniguchi, K. Watanabe, and P. Jarillo-Herrero, Interlayer electron-hole friction in tunable twisted bilayer graphene semimetal, Phys. Rev. Lett. **129**, 206802 (2022); https://doi.org/10.1103/PhysRevLett.129.206802
5. Guido Menichetti, Lorenzo Cavicchi, Leonardo Lucchesi, Fabio Taddei, Giuseppe Iannaccone, Pablo Jarillo-Herrero, Claudia Felser, Frank H. L. Koppens, and Marco Polini, Giant chirality-induced spin polarization in twisted transition metal dichalcogenides, Newton 1, 100013, (2025); https://doi.org/10.1016/j.newton.2025.100013

4. Conclusion

The quest to understand and control the motion of electrons in solid-state materials is a cornerstone of condensed matter physics and a driving force behind technological advances in electronics. Traditionally, the behavior of electrons in conductors has often been described through the lens of single-particle theory, where interactions between electrons are either neglected or treated perturbatively. This framework has been remarkably successful in explaining a wide range of electronic phenomena. In recent years, however, a paradigm shift has emerged, bringing to the fore the collective behavior of strongly interacting electrons, particularly in ultrapure materials. This has led to the burgeoning field of hydrodynamic electronics, where the flow of electrons is not simply a drift of independent particles but rather resembles the viscous flow of a fluid.

Hydrodynamics, as the "oldest and most successful theoretical formalism"[2] has long been the framework of choice for describing the collective motion of strongly interacting particles in diverse physical systems, from the macroscopic scale of water flow to the vastness of interstellar matter. The fundamental foundation of hydrodynamics lies in the conservation laws. Now, driven by significant advances in the nanostructuring of ultrapure materials in recent years, this powerful theoretical tool is finding new applications in the field of nanoelectronics. The observation of the hydrodynamic behavior of electrons in solids has opened up an exciting avenue for both fundamental research and potential technological breakthroughs. This emerging field, as highlighted in this textbook, focuses on the motion of electrons in high-purity materials. The core idea is that in such materials, electron-electron scattering can become the dominant interaction mechanism, leading to a state where momentum is conserved on longer time scales than momentum-relaxing scattering events (such as electron-impurity or electron-phonon scattering). In this regime, the electron system can be effectively described as an electronic fluid with properties analogous to classical fluids, such as viscosity and the ability to flow. This represents a profound intellectual challenge and is of great importance for the future of nanoelectronics.

The study of hydrodynamic electronics is not limited to fundamental physics; it also promises significant technological impact, particularly in nanoelectronics. Understanding and harnessing the viscous flow of electrons could lead to novel electronic devices with enhanced functionality and potentially lower energy dissipation. This introduction sets the stage for a deeper exploration of the fascinating world of hydrodynamic electronics. Drawing on the fundamental principles of fluid dynamics and the unique properties of advanced materials such as graphene, this field is pushing the boundaries of our understanding of electron transport and paving the way for future innovations in electronics technology. The interplay between theoretical predictions and experimental observations continues to drive this exciting area of research, revealing new and unexpected behaviors of electrons in the hydrodynamic regime.

---

[2] See a recent textbook authored by the HYDROTRONICS consortium member (WEIZMANN team) Gregory Falkovich, *Fluid mechanics*, Cambridge University Press (2018).

While the experimental and theoretical results presented in this textbook are very encouraging, we are only at the beginning of the development of understanding and controlling the fluid mechanics of electrons. Further improvements in sample quality and structuring are key aspects to expand the parameter regime to include hydrodynamic behaviour. From a theoretical perspective, there are fundamental challenges in the description of quantum kinetic processes in strongly interacting systems. Examples include the microscopic determination of transport coefficients, the solution of nonlinear kinetic theories, or the general problem of the timescales of local thermalization in closed quantum systems. The phenomena discussed here are only the first examples that make novel applications of hydrodynamic, viscous electronics seem possible.

5. Abstracts of References

Aharon-Steinberg, A., Völkl, T., Kaplan, A. et al. Direct observation of vortices in an electron fluid. Nature 607, 74–80 (2022); https://doi.org/10.1038/s41586-022-04794-y

Vortices are the hallmarks of hydrodynamic flow. Strongly interacting electrons in ultrapure conductors can display signatures of hydrodynamic behaviour, including negative non-local resistance, higher-than-ballistic conduction, Poiseuille flow in narrow channels and violation of the Wiedemann–Franz law. Here we provide a visualization of whirlpools in an electron fluid. By using a nanoscale scanning superconducting quantum interference device on a tip, we image the current distribution in a circular chamber connected through a small aperture to a current-carrying strip in the high-purity type II Weyl semimetal WTe_2. In this geometry, the Gurzhi momentum diffusion length and the size of the aperture determine the vortex stability phase diagram. We find that vortices are present for only small apertures, whereas the flow is laminar (non-vortical) for larger apertures. Near the vortical-to-laminar transition, we observe the single vortex in the chamber splitting into two vortices; this behaviour is expected only in the hydrodynamic regime and is not anticipated for ballistic transport. These findings suggest a new mechanism of hydrodynamic flow in thin pure crystals such that the spatial diffusion of electron momenta is enabled by small-angle scattering at the surfaces instead of the routinely invoked electron–electron scattering, which becomes extremely weak at low temperatures. This surface-induced para-hydrodynamics, which mimics many aspects of conventional hydrodynamics including vortices, opens new possibilities for exploring and using electron fluidics in high-mobility electron systems.

B. N. Narozhny. Hydrodynamic approach to two-dimensional electron systems. La Rivista del Nuovo Cimento 45, 661 (2022); https://doi.org/10.1007/s40766-022-00036-z

The last few years have seen an explosion of interest in hydrodynamic effects in interacting electron systems in ultra-pure materials. One such material, graphene, is not only an excellent platform for the experimental realization of the hydrodynamic flow of electrons, but also allows for a controlled derivation of the hydrodynamic equations on the basis of kinetic theory. The resulting hydrodynamic theory of electronic transport in graphene yields quantitative predictions for experimentally relevant quantities, e.g., viscosity, electrical conductivity, etc. Here I review recent theoretical advances in the field, compare the hydrodynamic theory of charge carriers in graphene with relativistic hydrodynamics and recent experiments, and discuss applications of hydrodynamic approach to novel materials beyond graphene.

B. N. Narozhny, I. V. Gornyi, and M. Titov. Anti-Poiseuille flow in neutral graphene. Phys. Rev. B 104, 075443 (2021); https://doi.org/10.1103/PhysRevB.104.075443

Hydrodynamic flow of charge carriers in graphene is an energy flow unlike the usual mass flow in conventional fluids. In neutral graphene, the energy flow is decoupled from the electric current, making it difficult to observe the hydrodynamic effects and measure the viscosity of the electronic fluid by means of electric current measurements. In particular, we show that the hallmark Poiseuille flow in a narrow channel cannot be driven by the electric field irrespective of boundary conditions at the channel edges. Nevertheless, one can observe nonuniform current densities similarly to the case of the well-known ballistic-diffusive crossover. The standard diffusive behavior with the uniform current density across the channel is achieved under the assumptions of specular scattering on the channel boundaries. This flow can also be made nonuniform by applying weak magnetic fields. In this case, the curvature of the current density profile is determined by the quasiparticle recombination processes dominated by the disorder-assisted electron-phonon scattering—the so-called supercollisions.

V. Gall, B. N. Narozhny, and I. V. Gornyi. Electronic viscosity and energy relaxation in neutral graphene. Phys. Rev. B 107, 045413 (2022); https://doi.org/10.1103/PhysRevB.107.045413

We explore hydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. In the absence of a magnetic field, the bulk Ohmic charge flow and the hydrodynamic energy flow are decoupled. However, the energy flow does affect the overall resistance of the system through viscous dissipation and energy relaxation that has to be compensated by the work done by the current source. Solving the hydrodynamic equations, we find that local temperature and electric potential are discontinuous at the interfaces with the leads as well as the device resistance and argue that this makes Corbino geometry a feasible choice for an experimental observation of the Dirac fluid.

V. Gall, B. N. Narozhny, and I. V. Gornyi. Corbino magnetoresistance in neutral graphene. Phys. Rev. B 107, 235401 (2023); https://doi.org/10.1103/PhysRevB.107.235401

We explore the magnetohydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. Based on the fully consistent hydrodynamic description derived from a microscopic framework and taking into account all peculiarities of graphene-specific hydrodynamics, we report the results of a comprehensive study of the interplay of viscosity, disorder-induced scattering, recombination, energy relaxation, and interface-induced dissipation. In the clean limit, magnetoresistance of a Corbino sample is determined by viscosity. Hence the Corbino geometry could be used to measure the viscosity coefficient in neutral graphene.

C. Kumar, J. Birbeck, J.A. Sulpizio, D. Perello *et al.* Imaging hydrodynamic electrons flowing without Landauer–Sharvin resistance. Nature 609, 276 (2022); https://doi.org/10.1038/s41586-022-05002-7

Electrical resistance usually originates from lattice imperfections. However, even a perfect lattice has a fundamental resistance limit, given by the Landauer conductance caused by a finite number of propagating electron modes. This resistance, shown by Sharvin to appear at the contacts of electronic devices, sets the ultimate conduction limit of non-interacting electrons. Recent years have seen growing evidence of hydrodynamic electronic phenomena prompting recent theories to ask whether an electronic fluid can radically break the fundamental Landauer–Sharvin limit. Here, we use single-electron-transistor imaging of electronic flow in high-mobility graphene Corbino disk devices to answer this question. First, by imaging ballistic flows at liquid-helium temperatures, we observe a Landauer–Sharvin resistance that does not appear at the contacts but is instead distributed throughout the bulk. This underpins the phase-space origin of this resistance—as emerging from spatial gradients in the number of conduction modes. At elevated temperatures, by identifying and accounting for electron–phonon scattering, we show the details of the purely hydrodynamic flow. Strikingly, we find that electron hydrodynamics eliminates the bulk Landauer–Sharvin resistance. Finally, by imaging spiralling magneto-hydrodynamic Corbino flows, we show the key emergent length scale predicted by hydrodynamic theories—the Gurzhi length. These observations demonstrate that electronic fluids can dramatically transcend the fundamental limitations of ballistic electrons, with important implications for fundamental science and future technologies.

S. Danz, M. Titov, and B. N. Narozhny, Giant nonlocality in nearly compensated two-dimensional semimetals. Phys. Rev. B 102, 081114(R) (2020); https://doi.org/10.1103/PhysRevB.102.081114

In compensated two-component systems in confined, two-dimensional geometries, nonlocal response may appear due to an external magnetic field. Within a phenomenological two-fluid framework, we demonstrate the evolution of charge flow profiles and the emergence of a giant nonlocal pattern dominating charge transport in a magnetic field. Applying our approach to the specific case of intrinsic graphene, we suggest a simple physical explanation for the experimental observation of giant nonlocality. Our results provide an intuitive way to predict the outcome of future experiments exploring the rich physics of many-body electron systems in confined geometries as well as to design possible applications.

B. N. Narozhny, I. V. Gornyi, and M. Titov. Hydrodynamic collective modes in graphene. Phys. Rev. B 103, 115402 (2021); https://doi.org/10.1103/PhysRevB.103.115402

Collective behavior is one of the most intriguing aspects of the hydrodynamic approach to electronic transport. Here we provide a consistent, unified calculation of the dispersion relations of the hydrodynamic collective modes in graphene. Taking into account viscous effects, we show that the hydrodynamic sound mode in graphene becomes overdamped at sufficiently large momentum scales. Extending the linearized theory beyond the hydrodynamic regime, we connect the diffusive hydrodynamic charge density fluctuations with plasmons.

Ady Stern, Thomas Scaffidi, Oren Reuven, Chandan Kumar, John Birkbeck, and Shahal Ilani. How Electron Hydrodynamics Can Eliminate the Landauer-Sharvin Resistance. Phys. Rev. Lett. 129, 157701 (2022); https://doi.org/10.1103/PhysRevLett.129.157701

It has long been realized that even a perfectly clean electronic system harbors a Landauer-Sharvin resistance, inversely proportional to the number of its conduction channels. This resistance is usually associated with voltage drops on the system's contacts to an external circuit. Recent theories have shown that hydrodynamic effects can reduce this resistance, raising the question of the lower bound of resistance of hydrodynamic electrons. Here, we show that by a proper choice of device geometry, it is possible to spread the Landauer-Sharvin resistance throughout the bulk of the system, allowing its complete elimination by electron hydrodynamics. We trace the effect to the dynamics of electrons flowing in channels that terminate within the sample. For ballistic systems this termination leads to back-reflection of the electrons and creates resistance. Hydrodynamically, the scattering of these electrons off other electrons allows them to transfer to transmitted channels and avoid the resistance. Counterintuitively, we find that in contrast to the ohmic regime, for hydrodynamic electrons the resistance of a device with a given width can decrease with its length, suggesting that a long enough device may have an arbitrarily small total resistance.

Egor I. Kiselev and Jörg Schmalian, Nonlocal hydrodynamic transport and collective excitations in Dirac fluids, Phys. Rev. B 102, 245434 (2020); https://doi.org/10.1103/PhysRevB.102.245434

We study the response of a Dirac fluid to electric fields and thermal gradients at finite wave numbers and frequencies in the hydrodynamic regime. We find that nonlocal transport in the hydrodynamic regime is governed by an infinite set of kinetic modes that describe noncollinear scattering events in different angular harmonic channels. The scattering rates of these modes $\tau^{-1}m$ increase as $|m|$, where m labels the angular harmonics. In an earlier publication, we pointed out that this dependence leads to anomalous, Lévy-flight-like phase space diffusion [Kiselev and Schmalian, Phys. Rev. Lett. **123**, 195302 (2019)]. Here, we show how this surprisingly simple, nonanalytic dependence allows us to obtain exact expressions for the nonlocal charge and electronic thermal conductivities. The peculiar dependence of the scattering rates on m also leads to a nontrivial structure of collective excitations: Besides the well-known plasmon, second-sound, and diffusive modes, we find nondegenerate damped modes corresponding to excitations of higher angular harmonics. We use these results to investigate the transport of a Dirac fluid through Poiseuille-type geometries of different widths and to study the response to surface acoustic waves in graphene-piezoelectric devices.

Egor I. Kiselev and Jörg Schmalian, Lévy Flights and Hydrodynamic Superdiffusion on the Dirac Cone of Graphene, Phys. Rev. Lett. 123, 195302 (2019); https://doi.org/10.1103/PhysRevLett.123.195302

We show that the hydrodynamic collision processes of graphene electrons at the neutrality point can be described in terms of a Fokker-Planck equation with a fractional derivative, corresponding to a Lévy flight in momentum space. Thus, electron-electron collisions give rise to frequent small-angle scattering processes that are interrupted by rare large-angle events. The latter give rise to superdiffusive dynamics of collective excitations. We argue that such superdiffusive dynamics is of more general importance to the out-of-equilibrium dynamics of quantum-critical systems.

Gu Zhang, Valentin Kachorovskii, Konstantin Tikhonov, and Igor Gornyi, Heating of inhomogeneous electron flow in the hydrodynamic regime, Phys. Rev. B 104, 075417 (2021); https://doi.org/10.1103/PhysRevB.104.075417

We study the electron temperature profiles for an inhomogeneous electron flow in the hydrodynamic regime. We assume that the inhomogeneity is due to a weakly nonuniform distribution of the momentum relaxation time within a spherically constricted area. We show that the temperature profile dramatically depends on the drive strength and the viscosity of the electron liquid. In the absence of viscosity, a Landauer-dipole-like temperature distribution, asymmetrically deformed along the current by the inelastic electron-phonon scattering, emerges around the inhomogeneity. We find that both the Landauer-dipole temperature profile and its asymmetry in the direction of the driving electric field exist in all dimensionalities and are, therefore, universal features of inhomogeneous hydrodynamic electron flow. We further demonstrate that the electron viscosity suppresses the thermal Landauer dipole and leads to the appearance of a "hot spot" exactly at the center of the constriction. We also calculate the phonon temperature distribution, which can be directly measured in experiments on thermal nanoimaging.

Alexander Block, Alessandro Principi, Niels C. H. Hesp, Aron W. Cummings, Matz Liebel, Kenji Watanabe, Takashi Taniguchi, Stephan Roche, Frank H. L. Koppens, Niek F. van Hulst, and Klaas-Jan Tielrooij, Observation of giant and tunable thermal diffusivity of a Dirac fluid at room temperature, Nature Nanotechnology volume 16, 1195 (2021); https://doi.org/10.1038/s41565-021-00957-6

Conducting materials typically exhibit either diffusive or ballistic charge transport. When electron–electron interactions dominate, a hydrodynamic regime with viscous charge flow emerges. More stringent conditions eventually yield a quantum-critical Dirac-fluid regime, where electronic heat can flow more efficiently than charge. However, observing and controlling the flow of electronic heat in the hydrodynamic regime at room temperature has so far remained elusive. Here we observe heat transport in graphene in the diffusive and hydrodynamic regimes, and report a controllable transition to the Dirac-fluid regime at room temperature, using carrier temperature and carrier density as control knobs. We introduce the technique of spatiotemporal thermoelectric microscopy with femtosecond temporal and nanometre spatial resolution, which allows for tracking electronic heat spreading. In the diffusive regime, we find a thermal diffusivity of roughly 2,000 $cm^2 s^{-1}$, consistent with charge transport. Moreover, within the hydrodynamic time window before momentum relaxation, we observe heat spreading corresponding to a giant diffusivity up to 70,000 $cm^2 s^{-1}$, indicative of a Dirac fluid. Our results offer the possibility of further exploration of these interesting physical phenomena and their potential applications in nanoscale thermal management.

Mohammad Zarenia, Alessandro Principi, and Giovanni Vignale, Thermal transport in compensated semimetals: Effect of electron-electron scattering on Lorenz ratio, Phys. Rev. B 102, 214304 (2020); https://doi.org/10.1103/PhysRevB.102.214304

It is well known that the electronic thermal conductivity of clean compensated semimetals can be greatly enhanced over the electric conductivity by the availability of an ambipolar mechanism of conduction, whereby electrons and holes flow in the same direction experiencing negligible Coulomb scattering as well as negligible impurity scattering. This enhancement—resulting in a breakdown of the Wiedemann-Franz law with an anomalously large Lorenz ratio—has been recently observed in two-dimensional monolayer and bilayer graphene near the charge neutrality point. In contrast to this, three-dimensional compensated semimetals such as WP_2 and Sb are typically found to show a reduced Lorenz ratio. We investigate the reasons for this difference, focusing on the low-temperature regime where the electron-electron scattering is expected to dominate over other scattering mechanisms. We show that the different regimes of Fermi statistics (nondegenerate electron-hole liquid in graphene versus degenerate electron-hole liquid in compensated semimetals) are not sufficient to explain the reduction of the Lorenz ratio in the latter. We propose that the solution of the puzzle lies in the large separation of electron and hole pockets in momentum space, which allows compensated semimetals to sustain sizable regions of electron-hole accumulation near the contacts. These accumulations suppress the ambipolar conduction mechanism and effectively split the system into two independent electron and hole conductors. We present a quantitative theory of the crossover from ambipolar to unipolar conduction as a function of the size of the electron-hole accumulation regions, and show that it naturally leads to a sample-size-dependent thermal conductivity.

A Kapustin and G Falkovich. Interplay between two mechanisms of resistivity arXiv:2407.16284; https://doi.org/10.48550/arXiv.2407.16284

Mechanisms of resistivity can be divided into two basic classes: one is dissipative (like scattering on phonons) and another is quasi-elastic (like scattering on static impurities). They are often treated by the empirical Matthiessen rule, which says that total resistivity is just the sum of these two contributions, which are computed separately. This is quite misleading for two reasons. First, the two mechanisms are generally correlated. Second, computing the elastic resistivity alone masks the fundamental fact that the linear-response approximation has a vanishing validity interval at vanishing dissipation. Limits of zero electric field and zero dissipation do not commute for the simple reason that one needs to absorb the Joule heat quadratic in the applied field. Here, we present a simple model that illustrates these two points. The model also illuminates the role of variational principles for non-equilibrium steady states.

Alex Levchenko, Jörg Schmalian, Transport properties of strongly coupled electron–phonon liquids, Annals of Physics 419, 168218 (2020); https://doi.org/10.1016/j.aop.2020.168218

In this work we consider the hydrodynamic behavior of a coupled electron–phonon fluid, focusing on electronic transport under the conditions of strong phonon drag. This regime occurs when the rate of phonon equilibration due to e.g. umklapp scattering is much slower than the rate of normal electron–phonon collisions. Then phonons and electrons form a coupled out-of-equilibrium state where the total quasi-momentum of the electron–phonon fluid is conserved. A joint flow-velocity emerges as a collective hydrodynamic variable. We derive the equation of motion for this fluid from the underlying microscopic kinetic theory and elucidate its effective viscosity and thermal conductivity. In particular, we derive decay times of arbitrary harmonics of the distribution function and reveal its corresponding super-diffusive relaxation on the Fermi surface. We further consider several applications of this theory to magneto-transport properties in the Hall-bar and Corbino-disk geometries, relevant to experiments. In our analysis we allow for general boundary conditions that cover the crossover from no-slip to no-stress flows. Our approach also covers a crossover from the Stokes to the Ohmic regime under the conditions of the Gurzhi effect. In addition, we consider the frequency dependence of the surface impedance and non-equilibrium noise. For the latter, we notice that in the diffusive regime, a Fokker–Planck approximation, applied to the electron–phonon collision integral in the Eliashberg form, reduces it to a differential operator with Burgers type nonlinearity. As a result, the non-equilibrium distribution function has a shock-wave structure in the energy domain. The consequence of this behavior for the Fano factor of the noise is investigated. In conclusion we discuss connections and limitations of our results in the context of recent electron–phonon drag measurements in Dirac and Weyl semimetals, and layout directions for further extensions and developments.

Bailey Winstanley, Henning Schomerus, and Alessandro Principi, Corbino field-effect transistors in a magnetic field: Highly tunable photodetectors, Phys. Rev. B 104, 165406 (2021); https://doi.org/10.1103/PhysRevB.104.165406

We study gated field-effect transistors (FETs) with an eccentric Corbino-disk geometry, such that the drain spans its circumference while the off-center inner ring acts as a source. An alternating current terahertz potential difference is applied between source and gate while a static source-drain voltage, rectified by the nonlinearities of FET electrons, is measured. When a magnetic field is applied perpendicular to the device, a strong resonance appears at the cyclotron frequency. The strength of the resonance can be tuned by changing the eccentricity of the disk. We show that there is an optimum value of the eccentricity that maximizes the responsivity of the FET.

N. Vladimirova, M. Shavit, S. Belan, and G. Falkovich, Second-harmonic generation as a minimal model of turbulence, Phys. Rev. E 104, 014129 (2021); https://doi.org/10.1103/PhysRevE.104.014129

When two resonantly interacting modes are in contact with a thermostat, their statistics is exactly Gaussian and the modes are statistically independent despite strong interaction. Considering a noise-driven system, we show that when one mode is pumped and another dissipates, the statistics of such cascades is never close to Gaussian, no matter what is the relation between interaction and noise. One finds substantial phase correlation in the limit of strong interaction or weak noise. Surprisingly, the mutual information between modes increases and entropy decreases when interaction strength decreases. We use the model to elucidate the fundamental problem of far-from equilibrium physics: where the information, or entropy deficit, is encoded, and how singular measures form. For an instability-driven system, such as laser, even a small added noise leads to large fluctuations of the relative phase near the stability threshold, while far from the equilibrium the conversion into the second harmonic is weakly affected by noise.

Michal Shavit and Gregory Falkovich, Singular Measures and Information Capacity of Turbulent Cascades, Phys. Rev. Lett. 125, 104501 (2020); https://doi.org/10.1103/PhysRevLett.125.104501

How weak is the weak turbulence? Here, we analyze turbulence of weakly interacting waves using the tools of information theory. It offers a unique perspective for comparing thermal equilibrium and turbulence. The mutual information between modes is stationary and small in thermal equilibrium, yet it is shown here to grow with time for weak turbulence in a finite box. We trace this growth to the concentration of probability on the resonance surfaces, which can go all the way to a singular measure. The surprising conclusion is that no matter how small is the nonlinearity and how close to Gaussian is the statistics of any single amplitude, a stationary phase-space measure is far from Gaussian, as manifested by a large relative entropy. This is a rare piece of good news for turbulence modeling: the resolved scales carry significant information about the unresolved scales. The mutual information between large and small scales is the information capacity of turbulent cascade, setting the limit on the representation of subgrid scales in turbulence modeling.

Natalia Vladimirova, Michal Shavit, and Gregory Falkovich, Fibonacci Turbulence, Phys. Rev. X 11, 021063 (2021); https://doi.org/10.1103/PhysRevX.11.021063

Never is the difference between thermal equilibrium and turbulence so dramatic, as when a quadratic invariant makes the equilibrium statistics exactly Gaussian with independently fluctuating modes. That happens in two very different yet deeply connected classes of systems: incompressible hydrodynamics and resonantly interacting waves. This work presents the first detailed information-theoretic analysis of turbulence in such strongly interacting systems. The analysis involves both energy and entropy and elucidates the fundamental roles of space and time in setting the cascade direction and the changes of the statistics along it. We introduce a beautifully simple yet rich family of discrete models with triplet interactions of neighboring modes and show that it has quadratic conservation laws defined by the Fibonacci numbers. Depending on how the interaction time changes with the mode number, three types of turbulence were found: single direct cascade, double cascade, and the first-ever case of a single inverse cascade. We describe quantitatively how deviation from thermal equilibrium all the way to turbulent cascades makes statistics increasingly non-Gaussian and find the self-similar form of the one-mode probability distribution. We reveal where the information (entropy deficit) is encoded and disentangle the communication channels between modes, as quantified by the mutual information in pairs and the interaction information inside triplets.

Hernán L. Calvo, Jose E. Barrios Vargas, and Luis E. F. Foa Torres, Floquet boundary states in AB-stacked graphite, Phys. Rev. B 101, 075424 (2020); https://doi.org/10.1103/PhysRevB.101.075424

We report on the effect of laser illumination with circularly polarized light on the electronic structure of AB-stacked graphite samples. By using Floquet theory in combination with Green's function techniques, we find that the polarized light induces band-gap openings at the Floquet zone edge $\hbar\Omega/2$, bridged by chiral boundary states. These states propagate mainly along the borders of the constituting layers as evidenced by the time-averaged local density of states and the probability current density in several geometries. Semianalytic calculations of the Chern number suggest that these states are of topological nature, similar to those found in illuminated 2D samples like monolayer and bilayer graphene. These states are promising candidates for the realization of a three-dimensional version of the quantum Hall effect for Floquet systems.

A. Huamán, L. E. F. Foa Torres, C. A. Balseiro, and Gonzalo Usaj, Quantum Hall edge states under periodic driving: A Floquet induced chirality switch, Phys. Rev. Research 3, 013201 (2021); https://doi.org/10.1103/PhysRevResearch.3.013201

We report on the fate of the quantum Hall effect in graphene under intense laser illumination. By using Floquet theory combined with both a low energy description and full tight-binding models, we clarify the selection rules, the quasienergy band structure, as well as their connection with the two-terminal and multiterminal conductance in a device setup as relevant for experiments. We show that the well-known dynamical gaps that appear in the Floquet spectrum at $\pm\hbar\Omega/2$ lead to a switch-off of the quantum Hall edge transport for different edge terminations except for the armchair one, where two terms cancel out exactly. More interestingly, we show that near the Dirac point changing the laser polarization (circular right or circular left) controls the Hall conductance, by allowing to switch it on or off, or even by flipping its sign, thereby reversing the chirality of the edge states. This might lead to new avenues to fully control topologically protected transport.

Matías Berdakin*, Esteban A. Rodríguez-Mena, Luis E. F. Foa Torres, Spin-Polarized Tunable Photocurrents, Nano Lett., 21, 3177 (2021); https://doi.org/10.1021/acs.nanolett.1c00420

Harnessing the unique features of topological materials for the development of a new generation of topological based devices is a challenge of paramount importance. Using Floquet scattering theory combined with atomistic models we study the interplay among laser illumination, spin, and topology in a two-dimensional material with spin–orbit coupling. Starting from a topological phase, we show how laser illumination can selectively disrupt the topological edge states depending on their spin. This is manifested by the generation of pure spin photocurrents and spin-polarized charge photocurrents under linearly and circularly polarized laser illumination, respectively. Our results open a path for the generation and control of spin-polarized photocurrents.

M. Kim, M., S.G. Xu, A.I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, R. V. Gorbachev, K. Watanabe, T. Taniguchi, I. V. Grigorieva, V. I. Fal'ko, M. Polini, and A. K. Geim, Control of electron-electron interaction in graphene by proximity screening, Nature Communications 11, 2339 (2020); https://doi.org/10.1038/s41467-020-15829-1

Electron-electron interactions play a critical role in many condensed matter phenomena, and it is tempting to find a way to control them by changing the interactions' strength. One possible approach is to place a studied system in proximity of a metal, which induces additional screening and hence suppresses electron interactions. Here, using devices with atomically-thin gate dielectrics and atomically-flat metallic gates, we measure the electron-electron scattering length in graphene and report qualitative deviations from the standard behavior. The changes induced by screening become important only at gate dielectric thicknesses of a few nm, much smaller than a typical separation between electrons. Our theoretical analysis agrees well with the scattering rates extracted from measurements of electron viscosity in monolayer graphene and of umklapp electron-electron scattering in graphene superlattices. The results provide a guidance for future attempts to achieve proximity screening of many-body phenomena in two-dimensional systems.

Leonid A. Ponomarenko, Alessandro Principi, Andy D. Niblett, Wendong Wang, Roman V. Gorbachev, Piranavan Kumaravadivel, Alexey I. Berdyugin, Alexey V. Ermakov, Sergey Slizovskiy, Kenji Watanabe, Takashi Taniguchi, Qi Ge, Vladimir I. Fal'ko, Laurence Eaves, Mark T. Greenaway and Andre K. Geim. Extreme electron–hole drag and negative mobility in the Dirac plasma of graphene, Nature Communications 15, 9869 (2024); https://doi.org/10.1038/s41467-024-54198-x

Coulomb drag between adjacent electron and hole gases has attracted considerable attention, being studied in various two-dimensional systems, including semiconductor and graphene heterostructures. Here we report measurements of electron–hole drag in the Planckian plasma that develops in monolayer graphene in the vicinity of its Dirac point above liquid-nitrogen temperatures. The frequent electron–hole scattering forces minority carriers to move against the applied electric field due to the drag induced by majority carriers. This unidirectional transport of electrons and holes results in nominally negative mobility for the minority carriers. The electron–hole drag is found to be strongest near room temperature, despite being notably affected by phonon scattering. Our findings provide better understanding of the transport properties of charge-neutral graphene, reveal limits on its hydrodynamic description, and also offer insight into quantum-critical systems in general.

Robert Sokolewicz, Mikhail Baglai, Ivan Ado, Mikhail Katsnelson, Mikhail Titov, Gilbert damping in two-dimensional metallic anti-ferromagnets, Phys. Rev. B 109, 134427 (2024); https://doi.org/10.1103/PhysRevB.109.134427

A finite spin life-time of conduction electrons may dominate Gilbert damping of two-dimensional metallic antiferromagnets or antiferromagnet/metal heterostructures. We investigate the Gilbert damping tensor for a typical low-energy model of a metallic antiferromagnet system with honeycomb magnetic lattice and Rashba spin-orbit coupling for conduction electrons. We distinguish three regimes of spin relaxation: exchange-dominated relaxation for weak spin-orbit coupling strength, Elliot-Yafet relaxation for moderate spin-orbit coupling, and Dyakonov-Perel relaxation for strong spin-orbit coupling. We show, however, that the latter regime takes place only for the in-plane Gilbert damping component. We also show that anisotropy of Gilbert damping persists for any finite spin-orbit interaction strength provided we consider no spatial variation of the Néel vector. Isotropic Gilbert damping is restored only if the electron spin-orbit length is larger than the magnon wavelength. Our theory applies to $MnPS_3$ monolayer on Pt or to similar systems.

Daria A. Dumitriu-I., Darius A. Deaconu, Alexander E. Kazantsev, Alessandro Principi, First-order effect of electron-electron interactions on the anomalous Hall conductivity of massive Dirac fermions, Phys. Rev. B 109, 165429 (2024); **https://doi.org/10.1103/PhysRevB.109.165429**

We investigate the first-order correction to the anomalous Hall conductivity of 2D massive Dirac fermions arising from electron-electron interactions. In a fully gapped system in the limit of zero temperature, we find that this correction vanishes, confirming the absence of perturbative corrections to the topological Hall conductivity. At finite temperature or chemical potential, we find that the total Hall response decays faster than in the noninteracting case, depending on the strength of electron-electron interactions. These features, which could potentially be observed experimentally, show the importance of two-body interactions for anomalous Hall transport.

Alexander Kazantsev, Amelia Mills, Eoin O'Neill, Hao Sun, Giovanni Vignale, and Alessandro Principi, Nonconservation of the Valley Density and Its Implications for the Observation of the Valley Hall Effect, Phys. Rev. Lett. 132, 106301 (2024); https://doi.org/10.1103/PhysRevLett.132.106301

We show that the conservation of the valley density in multivalley insulators is broken in an unexpected way by the electric field that drives the valley Hall effect. This implies that time-reversal-invariant fully gapped insulators, in which no bulk or edge state crosses the Fermi level, can support a valley Hall current in the bulk and yet show no valley density accumulation at the edges. Thus, the valley Hall effect cannot be observed in such systems. If the system is not fully gapped then valley density accumulation at the edges is possible. The accumulation has no contribution from undergap states and can be expressed as a Fermi surface average, for which we derive an explicit formula. We demonstrate the theory by calculating the valley density accumulations in an archetypical valley-Hall insulator: a gapped graphene nanoribbon. Surprisingly, we discover that a net valley density polarization is dynamically generated for certain edge terminations.

Jihang Zhu, Iacopo Torre, Marco Polini, and Allan H. MacDonald, Weak-Coupling Theory of Magic-Angle Twisted Bilayer Graphene, Phys. Rev. B 110, L121117 (2024); https://doi.org/10.1103/PhysRevB.110.L121117

Strong correlations occur in magic-angle twisted bilayer graphene (MATBG) when the octet of flat moiré minibands centered on charge neutrality (CN) is partially occupied. The octet consists of a single valence band and a single conduction band for each of four degenerate spin-valley flavors. Motivated by the importance of Hartree electrostatic interactions in determining the filling-factor-dependent band structure, we use a time-dependent Hartree approximation to gain insight into electronic correlations. We find that the electronic compressibility is dominated by Hartree interactions, that paramagnetic states are stable over a range of density near CN, and that the dependence of energy on flavor polarization is strongly overestimated by mean-field theory.

Lorenzo Cavicchi, Iacopo Torre, Pablo Jarillo-Herrero, Frank H. L. Koppens, and Marco Polini, Theory of intrinsic acoustic plasmons in twisted bilayer graphene, Phys. Rev. B 110, 045431 (2024); https://doi.org/10.1103/PhysRevB.110.045431

We present a theoretical study of the intrinsic plasmonic properties of twisted bilayer graphene (TBG) as a function of the twist angle θ (and other microscopic parameters such as temperature and filling factor). Our calculations, which rely on the random phase approximation, take into account four crucially important effects, which are treated on equal footing: (i) the layer-pseudospin degree of freedom, (ii) spatial nonlocality of the density-density response function, (iii) crystalline local field effects, and (iv) Hartree self-consistency. We show that the plasmonic spectrum of TBG displays a smooth transition from a strongly coupled regime (at twist angles $\theta\lesssim 2\circ$), where the low-energy spectrum is dominated by a weakly dispersive intraband plasmon, to a weakly coupled regime (for twist angles $\theta\gtrsim 2\circ$) where an acoustic plasmon clearly emerges. This crossover offers the possibility of realizing tunable mid-infrared subwavelength cavities, whose vacuum fluctuations may be used to manipulate the ground state of strongly correlated electron systems.

D. A. Bandurin, A. Principi, I. Y. Phinney, T. Taniguchi, K. Watanabe, and P. Jarillo-Herrero, Interlayer electron-hole friction in tunable twisted bilayer graphene semimetal, Phys. Rev. Lett. 129, 206802 (2022); https://doi.org/10.1103/PhysRevLett.129.206802

Charge-neutral conducting systems represent a class of materials with unusual properties governed by electron-hole ($e{-}h$) interactions. Depending on the quasiparticle statistics, band structure, and device geometry these semimetallic phases of matter can feature unconventional responses to external fields that often defy simple interpretations in terms of single-particle physics. Here we show that small-angle twisted bilayer graphene (SA TBG) offers a highly tunable system in which to explore interactions-limited electron conduction. By employing a dual-gated device architecture we tune our devices from a nondegenerate charge-neutral Dirac fluid to a compensated two-component $e{-}h$ Fermi liquid where spatially separated electrons and holes experience strong mutual friction. This friction is revealed through the T^2 resistivity that accurately follows the $e{-}h$ drag theory we develop. Our results provide a textbook illustration of a smooth transition between different interaction-limited transport regimes and clarify the conduction mechanisms in charge-neutral SA TBG.

Guido Menichetti, Lorenzo Cavicchi, Leonardo Lucchesi, Fabio Taddei, Giuseppe Iannaccone, Pablo Jarillo-Herrero, Claudia Felser, Frank H. L. Koppens, and Marco Polini, Giant chirality-induced spin polarization in twisted transition metal dichalcogenides, Newton 1, 100013, (2025);
https://doi.org/10.1016/j.newton.2025.100013

Electrons traveling through chiral materials—that is, materials with structures that differ from their mirror image—can acquire spin polarization through an effect called chirality-induced spin selectivity (CISS). This effect was first observed in chiral organic molecules, and while it can be used to engineer spin polarizers, it remains poorly understood. This study explores CISS in atomically thin layers of transition metal dichalcogenides (TMDs), showing that it is significantly amplified due to the interplay of structural chirality and strong spin-orbit coupling in these materials. Numerical simulations show that in the case of twisted bilayers, TMDs can achieve substantial spin polarization, making them tunable platforms for studying and manipulating spin polarization via the twist angle. These findings highlight the potential of twisted TMDs for spintronic devices.

Direct observation of vortices in an electron fluid

A. Aharon-Steinberg[1,†], T. Völkl[1,†], A. Kaplan[1,†], A. K. Pariari[1], I. Roy[1], T. Holder[1], Y. Wolf[1], A. Y. Meltzer[1], Y. Myasoedov[1], M. E. Huber[2], B. Yan[1], G. Falkovich[3], L. S. Levitov[4], M. Hücker[1], and E. Zeldov[1,]*

Vortices are the hallmarks of hydrodynamic flow. Recent studies indicate that strongly-interacting electrons in ultrapure conductors can display signatures of hydrodynamic behavior including negative nonlocal resistance [1–6], Poiseuille flow in narrow channels [7–10], and a violation of the Wiedemann-Franz law [11,12]. Here we provide the first visualization of whirlpools in an electron fluid. By utilizing a nanoscale scanning superconducting quantum interference device on a tip (SQUID-on-tip) [13] we image the current distribution in a circular chamber connected through a small aperture to an adjacent narrow current-carrying strip in high-purity type-II Weyl semimetal WTe_2. In this geometry, the Gurzhi momentum diffusion length and the size of the aperture determine the vortex stability phase diagram. We find that the vortices are present only for small apertures, whereas the flow is laminar (non-vortical) for larger apertures, consistent with the theoretical analysis of the hydrodynamic regime and in contrast to the expectations of ballistic transport in WTe_2 at low temperatures [10]. Moreover, near the vortical-to-laminar transition, we observe a single vortex in the chamber splitting into two vortices, a behavior that can occur only in the hydrodynamic regime and cannot be sustained by ballistic transport. These findings suggest a novel mechanism of hydrodynamic flow: instead of the commonly considered electron-electron scattering in the bulk, which becomes extremely weak at low temperatures, the spatial diffusion of the charge carriers' momenta is enabled by small-angle scattering at the planar surfaces of thin pure crystals. This surface-induced para-hydrodynamics opens new avenues for exploring and utilizing electron fluidics in high-mobility electron systems.

---

[1] Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 7610001, Israel

[2] Departments of Physics and Electrical Engineering, University of Colorado Denver, Denver, Colorado 80217, USA

[3] Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 7610001, Israel

[4] Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[†] These authors contributed equally to this work

[*] eli.zeldov@weizmann.ac.il

Recent years have seen a quest for systems and regimes in which strong electron-electron interactions may lead to electron flows governed by hydrodynamics [14], as in viscous fluids, rather than by Ohmic transport. Fluids display two distinct hydrodynamic regimes [15]: laminar flows in which neighboring sheets move at gradually varying velocities, and turbulent flows characterized by eddies and vortices with counterflow that develop into strongly fluctuating and chaotic behavior at large scales. The transition from laminar to turbulent flow is usually associated with nonlinear fluid dynamics, described by the Navier–Stokes equations. Yet, already in linear Stokes flow, hydrodynamic vortices in Newtonian fluids readily occur [16].

In contrast to common fluids, which display hydrodynamic phenomena abundantly, evidence for hydrodynamics in electron fluids has remained scarce [17,18] until recently. The advent of high purity single crystals, clean van der Waals heterostructures, and high mobility 2D systems has accelerated the observation of fluid-like behavior in semiconductors and semimetals [1–12,19–29] and triggered a flurry of theoretical works [19,30–61]. Recently, the laminar Poiseuille flow in narrow strips has been demonstrated in graphene [7–9] and in WTe$_2$ [10] with the help of Hall potential imaging and diamond nitrogen-vacancy magnetometry, lending support to the hydrodynamic nature of electron fluids in these systems. Yet, the most striking and ubiquitous feature in the flow of regular fluids, the formation of vortices and turbulence, has not yet been observed in electron fluids despite numerous theoretical predictions based on linear [62–64] and nonlinear [33,39,65] hydrodynamics. Transport measurements showing negative nonlocal resistance in the vicinity of the current injection point are suggestive of electron backflow in graphene and GaAs heterostructures [1–6]. Recent studies, however, propose that the observed negative potentials may arise in ballistic and hydrodynamic regimes even without an actual electron backflow [3,36,40]. Hence, direct observation of vortices and the study of their properties remains an outstanding challenge in electron fluids.

The conventional picture of electron transport [14] involves two distinct length scales -- the momentum-relaxation length l_{mr} describing momentum transfer from electrons to the lattice and the length l_{ee} describing momentum transfer between the carriers due to electron-electron collisions. Hydrodynamic behavior can harbor vortices and other unique effects associated with electron fluidity when $l_{ee} \ll l_{mr}$, at the length scales $l_{ee} < W < l_{mr}$, where W is the characteristic size of the system. In common metals, to the contrary, the shortest length scale is l_{mr}. In this case, the electron transport is described by Ohm's law $\boldsymbol{J} = -\sigma\nabla\phi = \sigma\boldsymbol{E}$ and the continuity equation $\nabla \cdot \boldsymbol{J} = 0$, where ϕ is the electrostatic potential and $\boldsymbol{E}$ is the electric field. In this regime, therefore, no vortices can exist since $\nabla \times \boldsymbol{J} = \sigma\nabla \times \boldsymbol{E} = 0$ in the steady state. Note that since we discuss quasi-2D geometry, $\boldsymbol{J}$ and σ denote here the 2D current density and conductivity, respectively. In ultraclean systems at low temperatures, however, very large values of l_{mr}, exceeding W, can be achieved. If both l_{ee} and l_{mr} exceed W, the transport is ballistic, in which case electrons propagate essentially unimpeded, with scattering occurring mainly at the device edges. Ballistic transport, unlike the ohmic transport, can also lead to vortices with interesting properties.

In this work, we provide a direct visualization of vortices in an electron fluid. By utilizing magnetic imaging with a scanning superconducting quantum interference device fabricated on the apex of a sharp pipette (SQUID-on-tip, SOT) [13], we observe the current flow patterns in the Weyl semimetal WTe$_2$ as shown schematically in Figs. 1a,b. Ultraclean single crystals with residual resistance ratio (RRR) of over 3,000 were grown as described in Methods (Extended Data Fig. 1) and exfoliated into thin flakes of thickness d of 23 to 48 nm. Various sample geometries were patterned using e-beam lithography and plasma etching (Methods). The primary geometry consists of a central strip of width $W = 550$ nm with two truncated circular chambers of radius $R = 900$ nm connected to its sides through apertures defined by the opening angle $\theta \leq 180^\circ$(see AFM images in Fig. 1c and Extended Data Fig. 6). Analogous geometries were patterned in Au films of similar thicknesses for comparison. An *ac* current with rms amplitude I_0 of 1 to 400 μA was applied to the samples at $T = 4.5$ K. The corresponding out-of-plane component of the Oersted field $B_z(x, y)$ was measured by the SOT, scanning at a height $h = 50$ nm above the sample surface (see Figs. 1a,b, Methods, and Extended Data Fig. 2), where the x and y directions are defined in Fig. 1e.

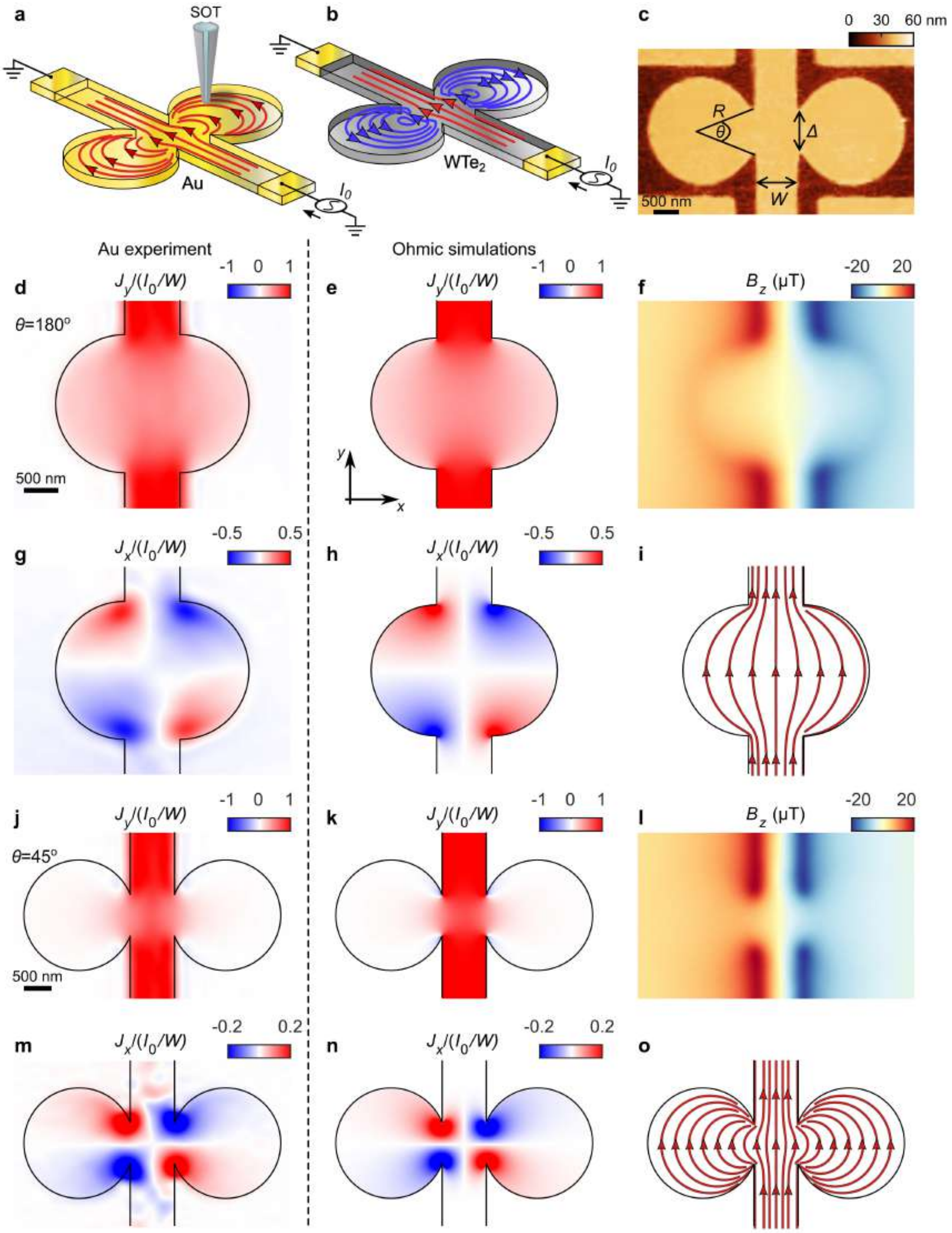


Fig. 1. **Ohmic electron flow in Au film. a-b**, Schematic experimental layout showing the scanning SOT and the Au (**a**) and WTe$_2$ (**b**) samples with double-chamber geometry. The red curves indicate laminar (open-loop) current streamlines while the blue curves represent closed-loop vortex streamlines. **c**, Atomic force microscope (AFM) topography image of WTe$_2$ sample with $\theta = 45°$. W, R, θ, and Δ are the width of the central strip, radius of the circular chambers, aperture angle, and size of the aperture $\Delta = 2R\sin(\theta/2)$, respectively. **d-i**, Measurements and simulations of Au sample with $\theta = 180°$. **d**, Current density $J_y(x,y)$ normalized by I_0/W reconstructed from $B_z(x,y)$ in (**f**). The black contours mark the sample edges. **e**, Simulated $J_y(x,y)$ in the ohmic regime. **f**, $B_z(x,y)$ measured by the SOT above the Au sample under current $I_0 = 50$ μA at 4.5 K. **g**, $J_x(x,y)$ reconstructed from (**f**). The light blue texture outside the sample is an artifact of current reconstruction (Methods). **h**, Simulated $J_x(x,y)$. **i**, Simulated current streamlines. **j-o**, Measurements and simulations of Au sample with $\theta = 45°$. **j**, Current density $J_y(x,y)$ reconstructed from (**l**). **k**, Simulated $J_y(x,y)$ in the ohmic regime. **l**, $B_z(x,y)$ profiles in the Au sample carrying $I_0 = 50$ μA. **m**, $J_x(x,y)$ reconstructed from (**l**). **n**, Simulated $J_x(x,y)$. **o**, Simulated current streamlines.

Ohmic flow

We start by examining the current flow in the Au films. Figure 1f shows $B_z(x, y)$ measured above the Au sample with $\theta = 180°$, corresponding to a strip with two half-disc chambers. By inversion of the magnetic field $B_z(x, y)$ [66], we reconstruct the 2D current density $\boldsymbol{J}(x, y)$ (Methods), with $J_x(x, y)$ and $J_y(x, y)$ components presented in Figs. 1d,g. The longitudinal J_y component demonstrates that the current flowing upwards in the central strip spreads out into the two chambers. The current flows into the right chamber through its lower half (red J_x in Fig. 1g), circulates counterclockwise, and exits through the upper half (blue J_x in Fig. 1g). The left chamber shows a mirrored flow pattern, as expected. The COMSOL numerical simulations in the ohmic regime in Figs. 1e,h show good agreement with the experimental data, describing a laminar (non-vortical) current flow as illustrated by the calculated streamlines in Fig. 1i. Upon decreasing the aperture size, less current enters the chambers as seen by $J_y(x, y)$ in Fig. 1j for the case of $\theta = 45°$. The $J_x(x, y)$ in Fig. 1m shows qualitatively similar behavior, with current flowing counterclockwise in the right chamber. Numerical simulations of J_y and J_x in Figs. 1k,n agree with the experimental data; the streamlines in Figs. 1o show laminar flow, as expected.

As a quick note on terminology – usually 'laminar' means no turbulence. Here, in a linear response regime, turbulence is not encountered. And yet, vortices may or may not appear depending on the dynamical phase of the electron fluid. Indeed, vortex is a flow in which the streamlines form closed loops. In contrast, streamlines that go from source to drain without forming closed loops will be referred to hereafter as laminar.

Vortex flow

For large opening angle, $\theta = 120°$, the current flow pattern in a WTe$_2$ sample (Fig. 2a) looks similar to that of Au with the $J_y(x, y)$ component spreading substantially into the chambers. The corresponding $J_x(x, y)$ in Fig. 2d shows counter-clockwise flow in the right chamber and clockwise flow in the left chamber, similar to the laminar flow in Au in Figs. 1g,m. To quantify the expected behavior in the hydrodynamic regime, $l_{ee} < W < l_{mr}$, we numerically solve the linearized Navier-Stokes equation for a 2D electron fluid [1,34,62],

$$-D^2\nabla^2\boldsymbol{J} + \boldsymbol{J} = -\sigma\nabla\phi, \qquad (1)$$

where D is the Gurzhi length, usually defined as $D = \sqrt{l_{ee}l_{mr}}/2$. The resulting $J_y(x, y)$ and $J_x(x, y)$ in Figs. 2b,e show good agreement with the experimental data. The corresponding calculated current streamlines in Fig. 2f show a laminar flow resembling the ohmic regime in Figs. 1i,o.

The flow pattern changes drastically as the aperture size becomes smaller, as illustrated in Figs. 2g,j for the case of $\theta = 20°$. The $J_y(x, y)$ remains focused in the central strip, while J_y in the chambers is relatively small (Fig. 2g). An essential difference between the flow in WTe$_2$ and the ohmic flow in Au is revealed, however, upon inspecting the transversal component $J_x(x, y)$. On approaching the aperture from below, J_x is initially directed to the right (red) towards the right chamber. Rather than maintaining its flow into the chamber as in Fig. 2d, J_x switches its direction and flows to the left (blue). Similarly, in the top half of the chamber, J_x flows out of the chamber near the aperture (blue), but into the chamber further away from the aperture (red). In other words, near the aperture the current flows counterclockwise while in the interior of the chamber, the current circulates clockwise. Therefore, a clockwise current vortex is formed in the right chamber. Simultaneously, a mirror-symmetric counterclockwise vortex appears in the left chamber. The hydrodynamic simulations of $J_y(x, y)$ and $J_x(x, y)$ in Figs. 2h,k confirm this picture, with the vortices in the two chambers represented by closed-loop streamlines (blue) in Fig. 2l (see also Extended Data Fig. 4). To the best of our knowledge, this constitutes the first direct observation of current vortices in an electron fluid.

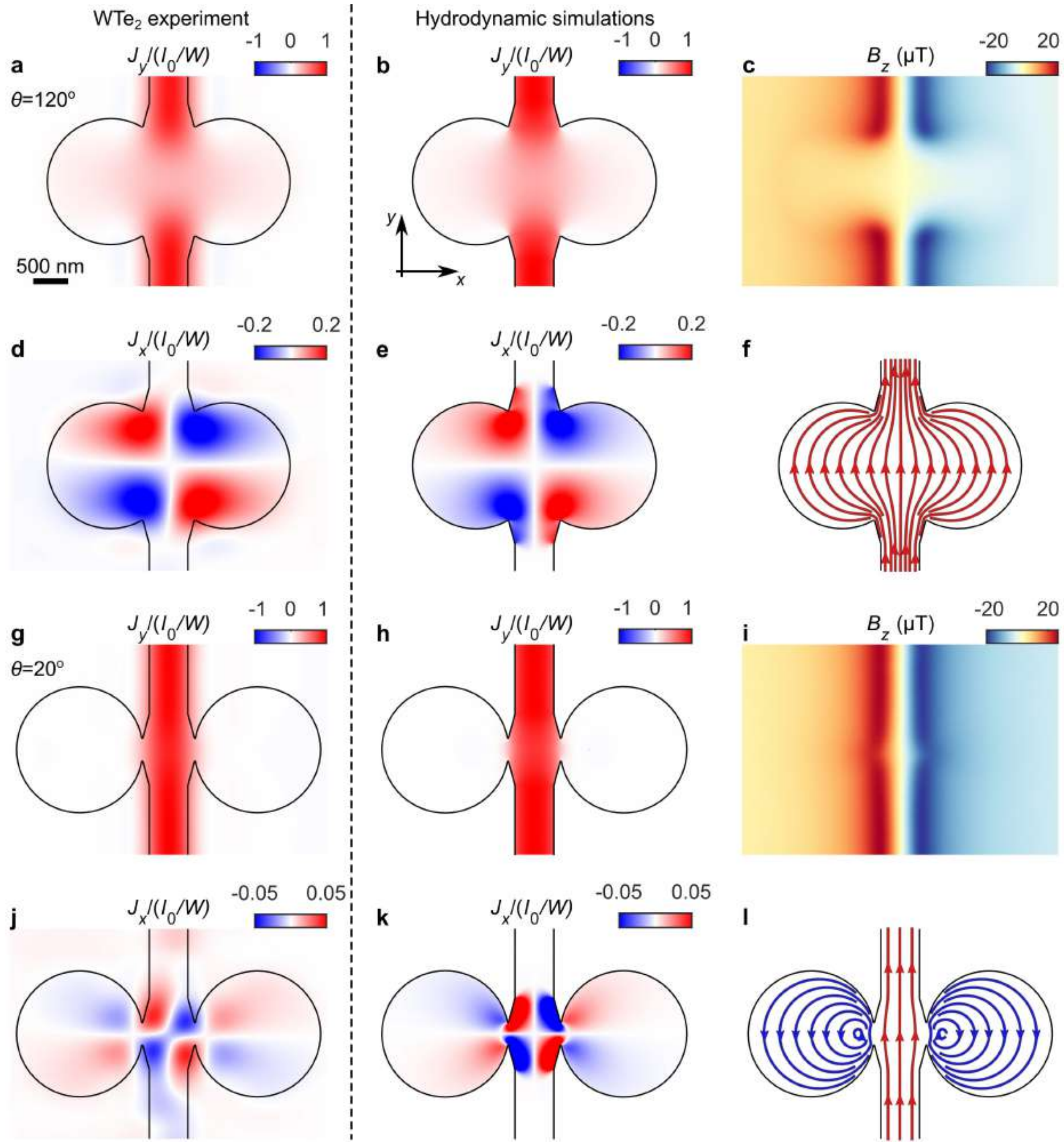


Fig. 2. **Laminar and vortex flow in WTe₂. a-f**, Measurements of WTe₂ sample with $\theta = 120°$ and corresponding simulations in the hydrodynamic regime (Eq. 1) with $D/W = 0.28$ and $\xi = 200$ nm. **a**, Current density $J_y(x,y)$ normalized by I_0/W reconstructed from (**c**). **b**, Simulated $J_y(x,y)$. **c**, $B_z(x,y)$ measured by the SOT above the WTe₂ sample under current $I_0 = 50$ μA at 4.5 K. **d**, $J_x(x,y)$ reconstructed from (**c**). **e**, Simulated $J_x(x,y)$. **f**, Simulated current streamlines showing laminar flow. **g-l**, Measurements of WTe₂ sample with $\theta = 20°$ and corresponding simulations in the hydrodynamic regime. **g**, Current density $J_y(x,y)$ reconstructed from (**i**). See Extended Data Figs. 2k-o in which the color scale is expanded so that the counterflow vortex current $J_y(x,y)$ is resolved. **h**, Simulated $J_y(x,y)$. **i**, $B_z(x,y)$ measured in the WTe₂ sample carrying $I_0 = 50$ μA. **j**, $J_x(x,y)$ reconstructed from (**l**). **k**, Simulated $J_x(x,y)$. **l**, Simulated current flow showing laminar (open-loop, red) and vortex (closed-loop, blue) streamlines.

Hydrodynamic and ballistic vortex stability phase diagram

Though the above hydrodynamic simulations show a good agreement with the experimental data, a question arises whether ballistic trajectories may create similar vortex patterns. Indeed, several studies have pointed out the difficulty in distinguishing the hydrodynamic and ballistic regimes using transport data [3,23,40,46,59]. Moreover, recent studies of current profiles in a WTe₂ whisker suggest the transport should be ballistic at low temperatures [10]. A key aspect that enables formation of vortices in both the hydrodynamic and ballistic regimes is system geometry [36]. Another key aspect is the boundary conditions, which has been an outstanding question in electron hydrodynamics from its early days [1,7–10,14,17,18,34,36,45,63,64,67]. The boundary conditions can be parametrized by the slip length, ξ, which can vary from $\xi = 0$ for no-slip ($\boldsymbol{J}|_{\text{boundary}} = 0$) to $\xi = \infty$ for no-stress boundaries (free-surface, $\hat{n} \cdot \nabla \boldsymbol{J}|_{\text{boundary}} = 0$). A full treatment of the electron transport requires the use of the Boltzmann kinetic equation, solution of which in an arbitrary 2D geometry is quite challenging [6,58]. We note, however, that for realistic parameters, Eq. 1 provides a good approximation in the regime of interest, $D/W \lesssim 1$, as well as in the quasi-ballistic regime, $D/W \gtrsim 1$, provided the no-stress boundary conditions are used [46] (see Methods).

To model vortex formation, we solve Eq. 1 in the two-chamber geometry and compute the total counterflow current I_v carried by the vortex, $I_v = \int_{-w/2}^{w/2} \left(|J_y(x,0)| - J_y(x,0) \right) \mathrm{d}x/4$, where $w = W + 2R[1 + \cos(\theta/2)]$ is the width of the structure in its widest section. Figures 3a,b show the resulting vortex stability phase diagram as a function of the aperture angle θ and the ratio of the Gurzhi length to the strip width, D/W, for no-stress and no-slip boundary conditions (see Methods). The resulting phase diagrams are quite similar, predicting that in the quasi-ballistic regime $D/W \gtrsim 1$ the vortices feature large counterflow I_v, are stable up to large angles θ, and show weak dependence on D/W. In the hydrodynamic regime, to the contrary, the counterflow I_v is lower, the presence of vortices is limited to smaller θ, and the vortex stability is strongly dependent on D/W. At low D/W, the vortex-to-laminar (no-vortex) phase transition line $\theta_t(D/W)$ is linear (dashed green line) and is essentially independent of the boundary conditions. Yet, the total circulating current I_v carried by the vortex is strongly dependent on the boundary conditions, showing about four-fold suppression of the maximum I_v for no-slip boundaries due to enhanced momentum relaxation at the edges.

These insights imply that by fabricating a series of samples with varying θ, one can pinpoint the transition angle θ_t. This would allow one to *i*) determine whether the flow is ballistic (large θ_t) or hydrodynamic (small θ_t), *ii*) evaluate the maximum I_v, and *iii*) extract the value of D in the hydrodynamic case. Moreover, the derived I_v can provide an estimate of the electron slip length ξ. Accordingly, we have fabricated six samples from a single WTe$_2$ flake with a sequence of aperture angles $\theta = 20°$, 35°, 54°, 72°, 90°, and 120° (Extended Data Fig. 6a). The observed current flow patterns are presented in Figs. 2, 3, and 4 (see Methods and Extended Data Figs. 7, 8, and 9 for additional samples and geometries). The six data points are overlaid on the two phase diagrams in Figs. 3a,b. A single vortex in each chamber is observed in two samples with the smallest θ (marked by $\odot$), a double-vortex is found in the $\theta = 54°$ sample (marked by $\ominus$), while laminar (no-vortex) flow is found for the three largest θ (marked by $\times$). As described below, the $\theta = 54°$ sample resides very close to the phase transition line, which allows us to identify the transition angle $\theta_t \cong 54°$. The small value of θ_t clearly establishes the hydrodynamic nature of the observed current vortices.

Gurzhi length and boundary conditions

The obtained θ_t value translates into $D/W \cong 0.28$, which for $W = 550$ nm in our devices results in the Gurzhi length $D \cong 155$ nm, a value nearly independent of the boundary conditions type in this range of parameters. The boundary conditions, however, strongly impact the vortex current I_v as seen in Figs. 3a,b and demonstrated in Figs. 3c-j. The experimentally derived $J_x(x,y)$ in $\theta = 20°$ and 35° samples are shown in Figs. 3c,g alongside simulated $J_x(x,y)$ for three values of the slip length, $\xi = 0$, 200 nm, and ∞. Figures 3f,j demonstrate that for no-slip boundary conditions, the intensity of the circulating current in the vortex is much weaker than the one measured experimentally. Circulating currents comparable to the experimental values in Figs. 3c,g can be achieved only for large $\xi \gtrsim 200$ nm, as shown in Figs. 3e,i. In this limit, the resulting current distribution is nearly identical to the one found for the no-stress boundary conditions, $\xi = \infty$, in Figs. 3d,h. The relatively low values of the measured $J_x/(I_0/W) \cong 0.02$, as compared to the significantly higher maximum values calculated for the ballistic regime, provide additional evidence for the hydrodynamic nature of the observed vortices.

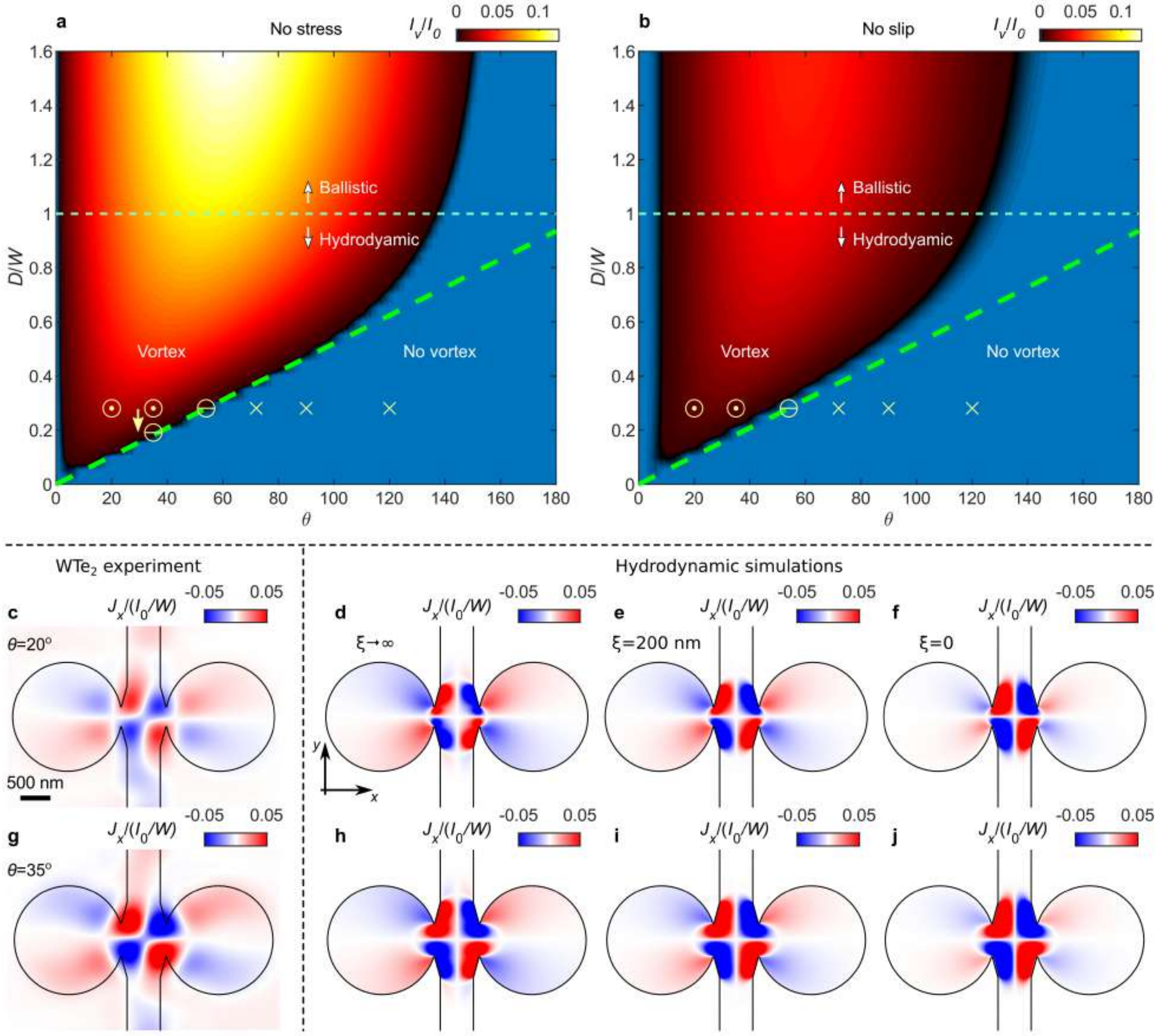


Fig. 3. **Vortex stability phase diagram. a-b**, Vortex stability phase diagram showing the magnitude of the circulating vortex current, I_v, in the chambers vs. the aperture angle θ and the Gurzhi length scaled by the strip width, D/W, for no-stress, $\xi = \infty$ (**a**), and no-slip, $\xi = 0$ (**b**), boundary conditions. The dashed green line indicates the vortical-to-laminar phase transition line $\theta_t(\frac{D}{W})$ in the $D/W \ll 1$ limit. The symbols ⊙, ⊖, and × mark the parameters of the chambers that feature single vortex, double vortex, and no vortices, respectively. The double-vortex state at $\theta = 35°$ is described in Fig. 5. **c-f**, Measured $J_x(x,y)$ in WTe₂ sample with $\theta = 20°$ (**c**) and the corresponding simulated $J_y(x,y)$ for $D/W = 0.28$ and electron slip length at the edges of $\xi = \infty$ (**d**), 200 nm (**e**), and 0 (**f**). **g-j**, Measured $J_y(x,y)$ in WTe₂ sample with $\theta = 35°$ (**g**) and the corresponding simulated $J_x(x,y)$ for $\xi = \infty$ (**h**), 200 nm (**i**), and 0 (**j**).

Our finding of a large slip length $\xi \gtrsim 200$ nm in the hydrodynamic flow is consistent with several transport studies of graphene [1,34,36], but are in an apparent disagreement with recent spatially resolved studies of graphene [7–9] and WTe₂ [10], which have suggested diffuse or no-slip boundary conditions in the hydrodynamic regime. Note, however, that these studies are based on the analysis of current profiles in a strip geometry in which ballistic transport and hydrodynamic flow with large slip length result in essentially indistinguishable current profiles (see further discussion in Methods and Extended Data Fig. 3).

Transition from laminar to vortex flow

We now examine more closely the transition between laminar and vortex flows. Figure 4 shows the experimental current distributions in samples with apertures $\theta = 90°$, $72°$, $54°$, and $35°$. The first two geometries, $\theta = 90°$ and $72°$, show laminar (no vortex) flow, in good agreement with the numerical results (Figs. 4a-j). In the sample with a smaller aperture $\theta = 35°$, a well-resolved vortex is observed in each chamber—also in agreement with simulations (Figs. 4p-t). The intermediate $\theta = 54°$ aperture shows, however, a markedly different $J_x(x,y)$ flow pattern (Fig. 4l). The corresponding numerical simulations (Fig. 4m-o) reveal that at the vortex-to-no-vortex phase transition ($\theta = \theta_t$), rather than vanishing continuously, the vortex elongates into an arc (see Methods and Supplementary Video 1) and eventually splits into two sub-vortices in the top and bottom parts of the chamber as shown by the streamlines in Fig. 4o. As a result, in each chamber, $J_x(x,y)$ shows two pairs of red-blue streaks, one for each sub-vortex (Fig. 4l), instead of a single pair of streaks as in Fig. 4q. The numerical simulations show that the double-vortex flow occurs only in a narrow interval of parameters just below the $\theta_t(D/W)$ phase transition line (where $I_v = 0$ according to our definition, because all the streamlines at $y = 0$ become laminar), which allows us to determine $\theta_t \cong 54°$ and $D/W \cong 0.28$ in our devices. Importantly, the double-vortex state can occur only in the hydrodynamic regime and is precluded in the ballistic transport (see Methods and Supplementary Video 2), thus providing additional strong evidence for the hydrodynamic origin of the observed vortex flow.

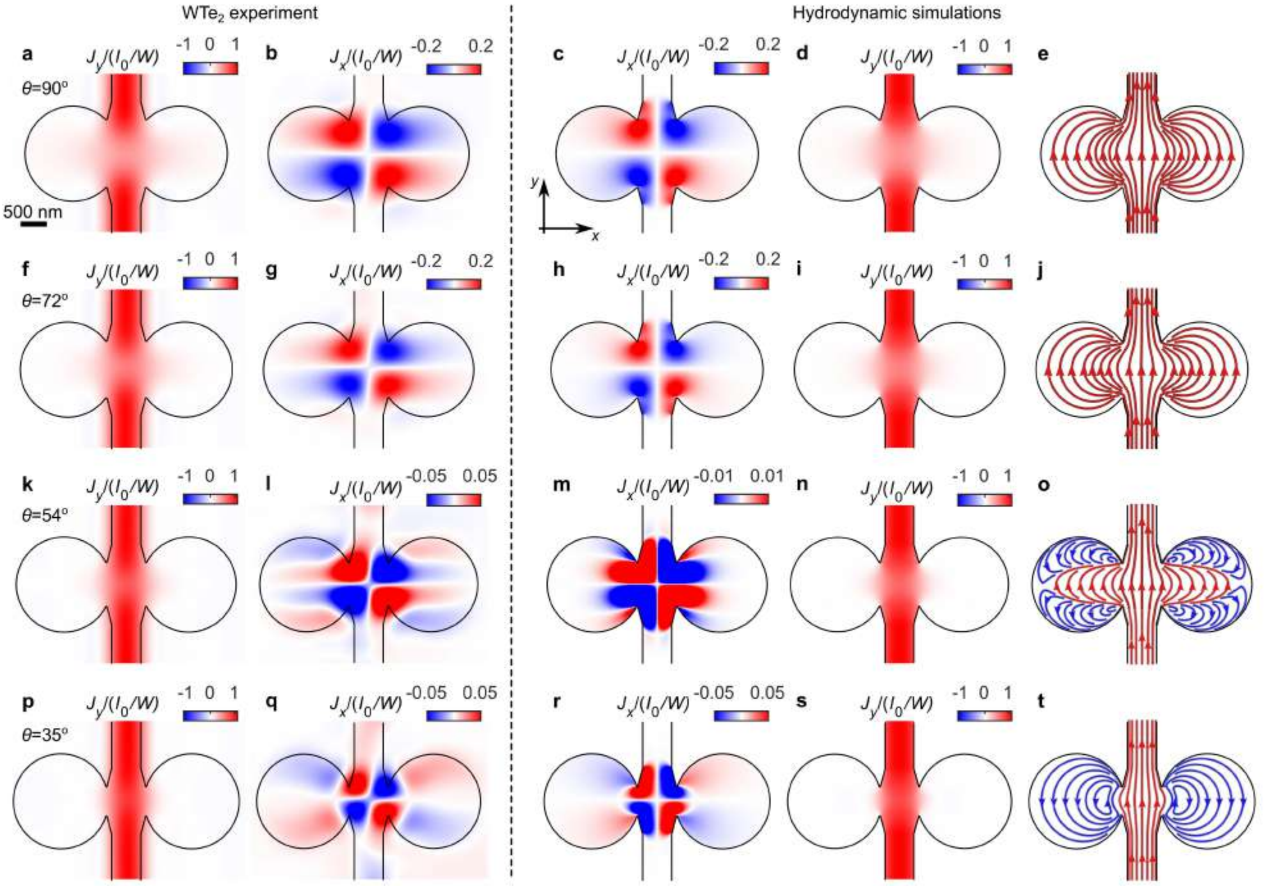


Fig. 4. **Laminar to double-vortex to single-vortex transition. a-e**, Measurements of WTe$_2$ sample with $\theta = 90°$ and corresponding simulations in the hydrodynamic regime with $D = 155$ nm and $\xi = 200$ nm. **a**, Measured current density $J_y(x,y)$ normalized by I_0/W at $I_0 = 50$ μA. **b**, Measured current density $J_x(x,y)$. **c**, Simulated $J_x(x,y)$. **d**, Simulated $J_y(x,y)$. **e**, Simulated laminar (red) current streamlines. **f-j**, Same as (a-e) for $\theta = 72°$ sample. **k-o**, Same as (a-e) for $\theta = 54°$ sample showing double vortex with laminar (red) and vortex (blue) current streamlines in (o). **p-t**, Same as (a-e) for $\theta = 35°$ sample showing single vortex with laminar (red) and vortex (blue) current streamlines in (t).

Current dependence

Our SOT microscope setup is limited to operation at temperatures of about 4 K, but the electron temperature can be raised substantially by increasing the applied current, as was originally employed in the hydrodynamic studies in GaAs 2DEG [17,18]. Since both l_{ee} and l_{mr} decrease with temperature, increasing the current I_0 is expected to reduce the Gurzhi length D. Figure 5 shows the evolution of $J_x(x,y)$ in the $\theta = 35°$ sample upon increasing I_0 up to 400 µA. The normalized $J_x/(I_0/W)$ circulating in the vortex gradually decreases as I_0 grows, until the double-vortex state is formed at our highest applied current. This behavior is qualitatively consistent with the expected decrease of D as indicated by the arrow in Fig. 3a. The degree of the reduction in D is, however, surprising. By measuring the sample resistance as a function of I_0 and comparing it to the temperature dependence of the resistance, we infer that the electron temperature reaches about 18 K at our highest applied current. Based on the theoretical estimate of the temperature dependence of l_{ee} and l_{mr} [10], such temperature increase should have reduced D by more than an order of magnitude relative to the value at 4.5 K. In contrast, Fig. 3a indicates that D has decreased by less than a factor of two. Note that in Ref. [10], the measured temperature evolution was also found to be weaker than predicted. This finding of weak temperature dependence of the Gurzhi length provides an important insight into the mechanism underlying hydrodynamics in our system.

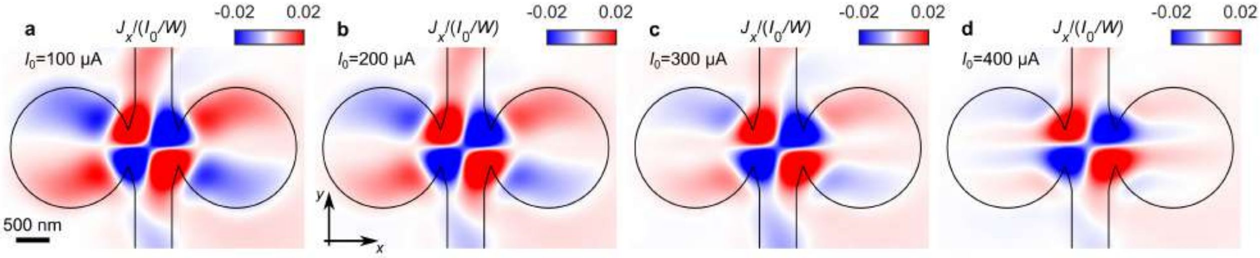


Fig. 5. **Current dependence of the vortex state. a-d**, $J_x(x,y)$ measured in WTe$_2$ sample with $\theta = 35°$ normalized by I_0/W for applied currents of $I_0 = 100$ µA (**a**), 200 µA (**b**), 300 µA (**c**), and 400 µA (**d**), showing a transition from single-vortex to double-vortex state in the chambers.

Discussion

Perhaps our most unexpected finding is that we observe fluid-like spatial momentum transport characterized by $D \cong 155$ nm, a value that is much smaller than $D = \sqrt{l_{mr} l_{ee}}/2$ estimated based on the bulk microscopic parameters. Indeed, transport measurements indicate that l_{mr} values in our bulk samples are in excess of 10 µm (Methods and Extended Data Fig. 1). Recent band structure calculations [10] suggest that bare l_{ee} in WTe$_2$ is of the order of a few mm at 4 K. Accounting for phonon-mediated electron-electron interactions can reduce l_{ee} to about 100 µm [10]. By taking into account the compensated semimetal band structure of WTe$_2$ and the proximity to band edges, we show that the bare l_{ee} can be of the order of 10 µm (Methods and Extended Data Fig. 5). Yet, even with this lower bound, l_{ee} remains much larger than W, which should have resulted in transport deep in the ballistic regime, with $D = \sqrt{l_{mr} l_{ee}}/2 \cong 5$ µm.

To gain insight into the origin of this surprising behavior, it is instructive to recall the general derivation of the Ohm-Stokes law, Eq. 1. Kinetic theory links momentum relaxation to the decay rate γ_1 of the first angular harmonic of the nonequilibrium electron momentum distribution, giving $l_{mr} = v_F/\gamma_1$, whereas the kinematic viscosity $\eta = \frac{1}{4}\frac{v_F^2}{\gamma_2}$ is expressed through the decay rate γ_2 of the second harmonic of the momentum distribution, where v_F is the Fermi velocity (Methods). This results in $D = \sqrt{\eta/\gamma_1} = v_F/\sqrt{4\gamma_1\gamma_2}$ (see also [35]). Importantly, this expression is completely general and is valid for any microscopic momentum-scattering mechanism. Usually γ_1 and γ_2 originate, respectively, from impurity or phonon scattering and electron-electron collisions, giving $\gamma_1 = v_F/l_{mr}$, $\gamma_2 = v_F/l_{ee}$, and resulting in $D =$

$\sqrt{l_{mr} l_{ee}}/2$. As discussed above, because of the large values of bulk l_{mr} and l_{ee}, these relations are inconsistent with D inferred from the observed behavior.

However, momentum-conserving electron-electron scattering is not required in order to generate spatial diffusion of the momentum. There is another scattering mechanism that could result in both the diffusion and relaxation of electron momenta, related to the finite thickness d of the sample. The resistivity of WTe_2 flakes has been found to be strongly thickness dependent [68,69], a behavior attributed to surface oxidation. Our transport measurements show that the conductivity of the flakes is one to two orders of magnitude lower than that of the bulk crystals (see Methods), indicating a large reduction in l_{mr} and an enhancement of γ_1 induced by the momentum relaxing scattering off the surfaces. However, an enhancement of γ_1 alone, such that $\gamma_1 > \gamma_2$, would of course lead to an ohmic transport with no vorticity, in contrast to the observed hydrodynamic flow, which requires $\gamma_2 > \gamma_1$. Yet, it has recently been pointed out [35] that enhancement in γ_1 implies an enhancement in γ_2. Indeed, using the method described in [59], we show in Methods that small-angle scattering results in $\gamma_2 \cong 4\gamma_1$, giving rise to hydrodynamic-like transport with $D = v_F/\sqrt{4\gamma_1\gamma_2} = v_F/4\gamma_1 = l_{mr}/4$. Based on $D \cong 155$ nm derived from the vortex stability diagram, we arrive at the effective surface-induced $l_{mr} = 4D \cong 620$ nm in our samples. This value compares well with the effective $l_{mr} \cong 530$ nm derived independently from transport measurements of conductivity in thin flakes (Methods).

The emerging picture is therefore as follows. For a fully specular surface scattering, the transport is ballistic. Small-angle scattering at the surfaces results in two effects: enhancement in momentum relaxation and a concurrent enhancement of the lateral momentum diffusion. In this para-hydrodynamic mechanism, momentum diffusion does not occur through the usual momentum-conserving electron-electron scattering, but rather through multiple, close-to-specular scattering events between the top and bottom surfaces without inter-electron momentum transfer. The resulting surface-induced para-hydrodynamics opens a unique possibility of observing and utilizing hydrodynamic phenomena in a wide range of high mobility materials without the necessity of the hard-to-achieve strong bulk electron-electron interactions.

References

1. D. A. Bandurin, I. Torre, R. K. Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V Grigorieva, L. A. Ponomarenko, A. K. Geim, and M. Polini, ''Negative local resistance caused by viscous electron backflow in graphene'', *Science* **351**, 1055–1058 (2016).
2. A. D. Levin, G. M. Gusev, E. V. Levinson, Z. D. Kvon, and A. K. Bakarov, ''Vorticity-induced negative nonlocal resistance in a viscous two-dimensional electron system'', *Phys. Rev. B* **97**, 245308 (2018).
3. D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, ''Fluidity onset in graphene'', *Nat. Commun.* **9**, 4533 (2018).
4. A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin, ''Measuring Hall viscosity of graphene's electron fluid'', *Science* **364**, 162–165 (2019).
5. M. Kim, S. G. Xu, A. I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, R. V. Gorbachev, K. Watanabe, T. Taniguchi, I. V. Grigorieva, V. I. Fal'ko, M. Polini, and A. K. Geim, ''Control of electron-electron interaction in graphene by proximity screening'', *Nat. Commun.* **11**, 2339 (2020).

6. A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, ''Hydrodynamic and Ballistic Transport over Large Length Scales in GaAs/AlGaAs'', *Phys. Rev. Lett.* **126**, 076803 (2021).

7. J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A. K. Geim, and S. Ilani, ''Visualizing Poiseuille flow of hydrodynamic electrons'', *Nature* **576**, 75–79 (2019).

8. M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M. M. Fogler, P. Kim, A. Yacoby, and R. L. Walsworth, ''Imaging viscous flow of the Dirac fluid in graphene'', *Nature* **583**, 537–541 (2020).

9. A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, K. Watanabe, T. Taniguchi, A. Lucas, A. F. Young, and A. C. B. Jayich, ''Imaging the breakdown of ohmic transport in graphene'', *arXiv:2002.05065* (2020).

10. U. Vool, A. Hamo, G. Varnavides, Y. Wang, T. X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C. A. C. Garcia, A. T. Pierce, J. Gooth, P. Anikeeva, C. Felser, P. Narang, and A. Yacoby, ''Imaging phonon-mediated hydrodynamic flow in WTe2'', *Nat. Phys.* **17**, 1216–1220 (2021).

11. J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, ''Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene'', *Science* **351**, 1058–1061 (2016).

12. J. Gooth, F. Menges, N. Kumar, V. Süβ, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, ''Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide'', *Nat. Commun.* **9**, 4093 (2018).

13. D. Vasyukov, Y. Anahory, L. Embon, D. Halbertal, J. Cuppens, L. Neeman, A. Finkler, Y. Segev, Y. Myasoedov, M. L. Rappaport, M. E. Huber, and E. Zeldov, ''A scanning superconducting quantum interference device with single electron spin sensitivity'', *Nat. Nanotechnol.* **8**, 639–644 (2013).

14. R. N. Gurzhi, ''Hydrodynamic effects in solids at low temperature'', *Sov. Phys. Uspekhi* **11**, 255–270 (1968).

15. L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*. Elsevier, 1987.

16. J. Mayzel, V. Steinberg, and A. Varshney, ''Stokes flow analogous to viscous electron current in graphene'', *Nat. Commun.* **10**, 937 (2019).

17. L. W. Molenkamp and M. J. M. de Jong, ''Observation of Knudsen and Gurzhi transport regimes in a two-dimensional wire'', *Solid. State. Electron.* **37**, 551–553 (1994).

18. M. J. M. de Jong and L. W. Molenkamp, ''Hydrodynamic electron flow in high-mobility wires'', *Phys. Rev. B* **51**, 13389–13402 (1995).

19. O. E. Raichev, G. M. Gusev, A. D. Levin, and A. K. Bakarov, ''Manifestations of classical size effect and electronic viscosity in the magnetoresistance of narrow two-dimensional conductors: Theory and experiment'', *Phys. Rev. B* **101**, 235314 (2020).

20. D. Taubert, G. J. Schinner, C. Tomaras, H. P. Tranitz, W. Wegscheider, and S. Ludwig, ''An electron jet pump: The venturi effect of a fermi liquid'', *J. Appl. Phys.* **109**, (2011).

21. P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, ''Evidence for hydrodynamic electron flow in PdCoO2'', *Science* **351**, 1061–1064 (2016).

22. R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, and A. K. Geim, ''Superballistic flow of viscous electron fluid through graphene constrictions'', *Nat. Phys.* **13**, 1182–1185 (2017).

23. B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, M. Polini, T. Ihn, and K. Ensslin, ''Scanning gate microscopy in a viscous electron fluid'', *Phys. Rev. B* **98**, 241304 (2018).

24. G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, ''Stokes flow around an obstacle in viscous two-dimensional electron liquid'', *Sci. Rep.* **10**, 7860 (2020).

25. G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, ''Viscous magnetotransport and Gurzhi effect in bilayer electron system'', *Phys. Rev. B* **103**, 075303 (2021).

26. Z. J. Krebs, W. A. Behn, S. Li, K. J. Smith, K. Watanabe, T. Taniguchi, A. Levchenko, and V. W. Brar, ''Imaging the breaking of electrostatic dams in graphene for ballistic and viscous fluids'', *arXiv:2106.07212* (2021).

27. J. Geurs, Y. Kim, K. Watanabe, T. Taniguchi, P. Moon, and J. H. Smet, ''Rectification by hydrodynamic flow in an encapsulated graphene Tesla valve'', *arXiv:2008.04862* (2020).

28. S. Samaddar, J. Strasdas, K. Janßen, S. Just, T. Johnsen, Z. Wang, B. Uzlu, S. Li, D. Neumaier, M. Liebmann, and M. Morgenstern, ''Evidence for Local Spots of Viscous Electron Flow in Graphene at Moderate Mobility'', *Nano Lett.* **21**, 9365–9373 (2021).

29. C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, and S. Ilani, ''Imaging Hydrodynamic Electrons Flowing Without Landauer-Sharvin Resistance'', *arXiv:2111.06412* (2021).

30. A. O. Govorov and J. J. Heremans, ''Hydrodynamic Effects in Interacting Fermi Electron Jets'', *Phys. Rev. Lett.* **92**, 026803 (2004).

31. M. Müller, J. Schmalian, and L. Fritz, ''Graphene: A Nearly Perfect Fluid'', *Phys. Rev. Lett.* **103**, 2–5 (2009).

32. A. V. Andreev, S. A. Kivelson, and B. Spivak, ''Hydrodynamic Description of Transport in Strongly Correlated Electron Systems'', *Phys. Rev. Lett.* **106**, 256804 (2011).

33. M. Mendoza, H. J. Herrmann, and S. Succi, ''Preturbulent Regimes in Graphene Flow'', *Phys. Rev. Lett.* **106**, 156601 (2011).

34. I. Torre, A. Tomadin, A. K. Geim, and M. Polini, ''Nonlocal transport and the hydrodynamic shear viscosity in graphene'', *Phys. Rev. B* **92**, 165433 (2015).

35. P. S. Alekseev, ''Negative Magnetoresistance in Viscous Flow of Two-Dimensional Electrons'', *Phys. Rev. Lett.* **117**, 166601 (2016).

36. F. M. D. Pellegrino, I. Torre, A. K. Geim, and M. Polini, ''Electron hydrodynamics dilemma: Whirlpools or no whirlpools'', *Phys. Rev. B* **94**, 155414 (2016).

37. H. Guo, E. Ilseven, G. Falkovich, and L. S. Levitov, ''Higher-than-ballistic conduction of viscous electron flows'', *PNAS* **114**, 3068–3073 (2017).

38. T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, ''Hydrodynamic Electron Flow and Hall Viscosity'', *Phys. Rev. Lett.* **118**, 226601 (2017).

39. V. Galitski, M. Kargarian, and S. Syzranov, ''Dynamo Effect and Turbulence in Hydrodynamic Weyl Metals'', *Phys. Rev. Lett.* **121**, 176603 (2018).

40. A. Shytov, J. F. Kong, G. Falkovich, and L. Levitov, ''Particle Collisions and Negative Nonlocal Response of Ballistic Electrons'', *Phys. Rev. Lett.* **121**, 176805 (2018).

41. D. Svintsov, ''Hydrodynamic-to-ballistic crossover in Dirac materials'', *Phys. Rev. B* **97**, 121405 (2018).

42. P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, ''Nonmonotonic magnetoresistance of a two-dimensional viscous electron-hole fluid in a confined geometry'', *Phys. Rev. B* **97**, 085109 (2018).

43. P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, ''Counterflows in viscous electron-hole fluid'', *Phys. Rev. B* **98**, 125111 (2018).

44. I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, ''Dissipative and Hall Viscosity of a Disordered 2D Electron Gas'', *Phys. Rev. Lett.* **123**, 26804 (2019).

45. K. A. Guerrero-Becerra, F. M. D. Pellegrino, and M. Polini, ''Magnetic hallmarks of viscous electron

flow in graphene'', *Phys. Rev. B* **99**, 041407 (2019).

46. T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, ''Ballistic and hydrodynamic magnetotransport in narrow channels'', *Phys. Rev. B* **100**, 245305 (2019).
47. P. Ledwith, H. Guo, A. Shytov, and L. Levitov, ''Tomographic Dynamics and Scale-Dependent Viscosity in 2D Electron Systems'', *Phys. Rev. Lett.* **123**, 116601 (2019).
48. B. N. Narozhny and M. Schütt, ''Magnetohydrodynamics in graphene: Shear and Hall viscosities'', *Phys. Rev. B* **100**, 035125 (2019).
49. D. Di Sante, J. Erdmenger, M. Greiter, I. Matthaiakakis, R. Meyer, D. R. Fernández, R. Thomale, E. van Loon, and T. Wehling, ''Turbulent hydrodynamics in strongly correlated Kagome metals'', *Nat. Commun.* **11**, 3997 (2020).
50. R. Toshio, K. Takasan, and N. Kawakami, ''Anomalous hydrodynamic transport in interacting noncentrosymmetric metals'', *Phys. Rev. Res.* **2**, 032021 (2020).
51. E. H. Hasdeo, J. Ekström, E. G. Idrisov, and T. L. Schmidt, ''Electron hydrodynamics of two-dimensional anomalous Hall materials'', *Phys. Rev. B* **103**, 125106 (2021).
52. Y. Huang and M. Wang, ''Nonnegative magnetoresistance in hydrodynamic regime of electron fluid transport in two-dimensional materials'', *Phys. Rev. B* **104**, 155408 (2021).
53. A. Hui, V. Oganesyan, and E. Kim, ''Beyond Ohm's law: Bernoulli effect and streaming in electron hydrodynamics'', *Phys. Rev. B* **103**, 235152 (2021).
54. B. N. Narozhny, I. V. Gornyi, and M. Titov, ''Anti-Poiseuille flow in neutral graphene'', *Phys. Rev. B* **104**, 075443 (2021).
55. B. N. Narozhny, I. V. Gornyi, and M. Titov, ''Hydrodynamic collective modes in graphene'', *Phys. Rev. B* **103**, 115402 (2021).
56. O. Tavakol and Y. B. Kim, ''Artificial electric field and electron hydrodynamics'', *Phys. Rev. Res.* **3**, 013290 (2021).
57. G. Zhang, V. Kachorovskii, K. Tikhonov, and I. Gornyi, ''Heating of inhomogeneous electron flow in the hydrodynamic regime'', *Phys. Rev. B* **104**, 075417 (2021).
58. S. Li, M. Khodas, and A. Levchenko, ''Conformal maps of viscous electron flow in the Gurzhi crossover'', *Phys. Rev. B* **104**, 155305 (2021).
59. K. G. Nazaryan and L. Levitov, ''Robustness of vorticity in electron fluids'', *arXiv:2111.09878* (2021).
60. A. Principi, G. Vignale, M. Carrega, and M. Polini, ''Bulk and shear viscosities of the two-dimensional electron liquid in a doped graphene sheet'', *Phys. Rev. B* **93**, 125410 (2016).
61. A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, and S. Ilani, ''Spread and erase -- How electron hydrodynamics can eliminate the Landauer-Sharvin resistance'', *arXiv:2110.15369* (2021).
62. L. Levitov and G. Falkovich, ''Electron viscosity, current vortices and negative nonlocal resistance in graphene'', *Nat. Phys.* **12**, 672–676 (2016).
63. G. Falkovich and L. Levitov, ''Linking Spatial Distributions of Potential and Current in Viscous Electronics'', *Phys. Rev. Lett.* **119**, 066601 (2017).
64. S. Danz and B. N. Narozhny, ''Vorticity of viscous electronic flow in graphene'', *2D Mater.* **7**, 035001 (2020).
65. A. Gabbana, M. Polini, S. Succi, R. Tripiccione, and F. M. D. Pellegrino, ''Prospects for the Detection of Electronic Preturbulence in Graphene'', *Phys. Rev. Lett.* **121**, 236602 (2018).
66. A. Y. Meltzer, E. Levin, and E. Zeldov, ''Direct Reconstruction of Two-Dimensional Currents in Thin Films from Magnetic-Field Measurements'', *Phys. Rev. Appl.* **8**, 064030 (2017).
67. E. I. Kiselev and J. Schmalian, ''Boundary conditions of viscous electron flow'', *Phys. Rev. B* **99**,

035430 (2019).

68. L. Wang, I. Gutiérrez-Lezama, C. Barreteau, N. Ubrig, E. Giannini, and A. F. Morpurgo, ''Tuning magnetotransport in a compensated semimetal at the atomic scale'', *Nat. Commun.* **6**, 8892 (2015).

69. J. M. Woods, J. Shen, P. Kumaravadivel, Y. Pang, Y. Xie, G. A. Pan, M. Li, E. I. Altman, L. Lu, and J. J. Cha, ''Suppression of Magnetoresistance in Thin WTe2 Flakes by Surface Oxidation'', *ACS Appl. Mater. Interfaces* **9**, 23175–23180 (2017).

70. M. N. Ali, J. Xiong, S. Flynn, J. Tao, Q. D. Gibson, L. M. Schoop, T. Liang, N. Haldolaarachchige, M. Hirschberger, N. P. Ong, and R. J. Cava, ''Large, non-saturating magnetoresistance in WTe2'', *Nature* **514**, 205–208 (2014).

71. Y.-Y. Lv, X. Li, B.-B. Zhang, W. Y. Deng, S.-H. Yao, Y. B. Chen, J. Zhou, S.-T. Zhang, M.-H. Lu, L. Zhang, M. Tian, L. Sheng, and Y.-F. Chen, ''Experimental Observation of Anisotropic Adler-Bell-Jackiw Anomaly in Type-II Weyl Semimetal WTe1.98 Crystals at the Quasiclassical Regime'', *Phys. Rev. Lett.* **118**, 096603 (2017).

72. Y. Wu, N. H. Jo, M. Ochi, L. Huang, D. Mou, S. L. Bud'ko, P. C. Canfield, N. Trivedi, R. Arita, and A. Kaminski, ''Temperature-Induced Lifshitz Transition in WTe2'', *Phys. Rev. Lett.* **115**, 166602 (2015).

73. M. N. Ali, L. Schoop, J. Xiong, S. Flynn, Q. Gibson, M. Hirschberger, N. P. Ong, and R. J. Cava, ''Correlation of crystal quality and extreme magnetoresistance of WTe2'', *EPL Europhys. Lett.* **110**, 67002 (2015).

74. Z. Zhu, X. Lin, J. Liu, B. Fauqué, Q. Tao, C. Yang, Y. Shi, and K. Behnia, ''Quantum Oscillations, Thermoelectric Coefficients, and the Fermi Surface of Semimetallic WTe2'', *Phys. Rev. Lett.* **114**, 176601 (2015).

75. F.-X. Xiang, M. Veldhorst, S.-X. Dou, and X.-L. Wang, ''Multiple Fermi pockets revealed by Shubnikov-de Haas oscillations in WTe2'', *EPL* **112**, 37009 (2015).

76. Q. Zhang, Z. Liu, Y. Sun, H. Yang, J. Jiang, S. Mo, Z. Hussain, X. Qian, L. Fu, S. Yao, M. Lu, C. Felser, B. Yan, Y. Chen, and L. Yang, ''Lifshitz Transitions Induced by Temperature and Surface Doping in Type-II Weyl Semimetal Candidate Td -WTe2'', *Phys. Status Solidi RRL* **11**, 1700209 (2017).

77. Y. Luo, H. Li, Y. M. Dai, H. Miao, Y. G. Shi, H. Ding, A. J. Taylor, D. A. Yarotski, R. P. Prasankumar, and J. D. Thompson, ''Hall effect in the extremely large magnetoresistance semimetal WTe2'', *Appl. Phys. Lett.* **107**, 182411 (2015).

78. P. Wang, G. Yu, Y. Jia, M. Onyszczak, F. A. Cevallos, S. Lei, S. Klemenz, K. Watanabe, T. Taniguchi, R. J. Cava, L. M. Schoop, and S. Wu, ''Landau quantization and highly mobile fermions in an insulator'', *Nature* **589**, 225–229 (2021).

79. N. Kumar, Y. Sun, N. Xu, K. Manna, M. Yao, V. Süss, I. Leermakers, O. Young, T. Förster, M. Schmidt, H. Borrmann, B. Yan, U. Zeitler, M. Shi, C. Felser, and C. Shekhar, ''Extremely high magnetoresistance and conductivity in the type-II Weyl semimetals WP2 and MoP2'', *Nat. Commun.* **8**, 1642 (2017).

80. Y. Anahory, H. R. Naren, E. O. Lachman, S. Buhbut Sinai, A. Uri, L. Embon, E. Yaakobi, Y. Myasoedov, M. E. Huber, R. Klajn, and E. Zeldov, ''SQUID-on-tip with single-electron spin sensitivity for high-field and ultra-low temperature nanomagnetic imaging'', *Nanoscale* **12**, 3174–3182 (2020).

81. M. E. Huber, P. A. Neil, R. G. Benson, D. A. Burns, A. M. Corey, C. S. Flynn, Y. Kitaygorodskaya, O. Massihzadeh, J. M. Martinis, and G. C. Hilton, ''DC SQUID series array amplifiers with 120 MHz bandwidth'', *IEEE Trans. Appiled Supercond.* **11**, 1251–1256 (2001).

82. A. Finkler, Y. Segev, Y. Myasoedov, M. L. Rappaport, L. Ne'eman, D. Vasyukov, E. Zeldov, M. E. Huber, J. Martin, and A. Yacoby, ''Self-Aligned Nanoscale SQUID on a Tip'', *Nano Lett.* **10**, 1046–1049 (2010).

83. A. Finkler, D. Vasyukov, Y. Segev, L. Ne'eman, E. O. Lachman, M. L. Rappaport, Y. Myasoedov, E. Zeldov, and M. E. Huber, ''Scanning superconducting quantum interference device on a tip for

magnetic imaging of nanoscale phenomena'', *Rev. Sci. Instrum.* **83**, 073702 (2012).

84. D. Halbertal, J. Cuppens, M. Ben Shalom, L. Embon, N. Shadmi, Y. Anahory, H. R. Naren, J. Sarkar, A. Uri, Y. Ronen, Y. Myasoedov, L. S. Levitov, E. Joselevich, A. K. Geim, and E. Zeldov, ''Nanoscale thermal imaging of dissipation in quantum systems'', *Nature* **539**, 407–410 (2016).
85. D. A. Broadway, S. E. Lillie, S. C. Scholten, D. Rohner, N. Dontschuk, P. Maletinsky, J.-P. Tetienne, and L. C. L. Hollenberg, ''Improved Current Density and Magnetization Reconstruction Through Vector Magnetic Field Measurements'', *Phys. Rev. Appl.* **14**, 024076 (2020).
86. L. Dell'Anna and W. Metzner, ''Fermi surface fluctuations and single electron excitations near Pomeranchuk instability in two dimensions'', *Phys. Rev. B* **73**, 045127 (2006).

Acknowledgments The authors thank Michal Shavit and Victor Steinberg for useful discussions. This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant No. 785971), by the German-Israeli Foundation for Scientific Research and Development (GIF) grant No. I-1505-303.10/2019, and by the Israel Science Foundation (ISF) grant No. 994/19. G.F. was supported by the Scientific Excellence Center at WIS, the Simons Foundation grant 662962, the EU Horizon 2020 programme grant 873028, the U.S.-Israel Binational Science Foundation (BSF) grant 2018033 and NSF-BSF grant 2020765. B.Y. acknowledges the financial support by the European Research Council (ERC Consolidator Grant "NonlinearTopo", No. 815869) and the ISF grant No. 2932/21. E.Z. acknowledges the support of the Andre Deloro Prize for Scientific Research. L.S.V. and E.Z. acknowledge the support of the Sagol Weizmann-MIT Bridge Program. M.H. and E.Z. acknowledge the support of the Leona M. and Harry B. Helmsley Charitable Trust grants No. 2018PG-ISL006 and 2112-04911. A.K.P. acknowledges the postdoctoral fellowship support from the Council for Higher Education, Israel through 'Study in Israel' program.

Author contributions A.A., T.V., A.K. and E.Z. conceived the experiments. A.K.P. and M.H. grew and characterized the bulk WTe_2 crystals. A.K. fabricated and characterized the devices. T.V. and A.A. carried out the SOT magnetic imaging measurements and data analysis. I.R. and Y.M. fabricated the SOTs and the tuning fork feedback. A.Y.M. developed the current density reconstruction method. M.E.H. designed and built the SOT readout system. Y.W., T.H. and B.Y. carried out the band structure and the electron-electron scattering calculations. T.H. developed the vortex stability model. A.A. performed the numerical simulations. G.F., L.S.L. and E.Z. developed the para-hydrodynamic model. A.A., T.V., T.H., A.K.P, E.Z., G.F. and L.S.L. wrote the manuscript with contributions from the rest of the authors.

Competing interests The authors declare no competing interests.

Data availability The data that support the findings of this study are available from the corresponding authors on reasonable request.

Code availability The current reconstruction codes used in this study are available from the corresponding authors on reasonable request.

Methods

Synthesis of WTe_2 crystals

To obtain high quality WTe_2 single crystals we conducted a series of synthesis experiments using both chemical vapor transport (CVT) and the flux growth technique, as well as starting materials with different purity. These experiments led to progressively better crystals with increased $RRR = \frac{\rho(300\ \mathrm{K})}{\rho(2\ \mathrm{K})}$ and magneto resistance ratio $MR = \frac{\rho(9\ \mathrm{T})-\rho(0\ \mathrm{T})}{\rho(0\ \mathrm{T})}$, as summarized in Extended Data Fig. 1c, and are described here in chronological order. Eventually all devices for the hydrodynamic flow experiments were fabricated from our highest quality single crystals described last. The first crystals were grown by CVT [70,71] using elemental W (99.95%) and Te (99.99%) from Stanford Advanced Materials. Initially, polycrystalline WTe_2 was prepared by solid-state reaction in a vacuum-sealed quartz ampule at 750 °C. The obtained precursor material was then vacuum-sealed (1.33×10^{-5} mbar) in a 16 cm long quartz tube with a minute amount of $TeBr_4$ transport agent, and placed in a temperature gradient of 850 °C – 750 °C for several days. Few millimeters wide sheet-shaped single crystals were collected from the cold end of the ampule after cool down. However, we found that both RRR and MR of the CVT grown crystals were extremely low, despite testing a range of different growth conditions.

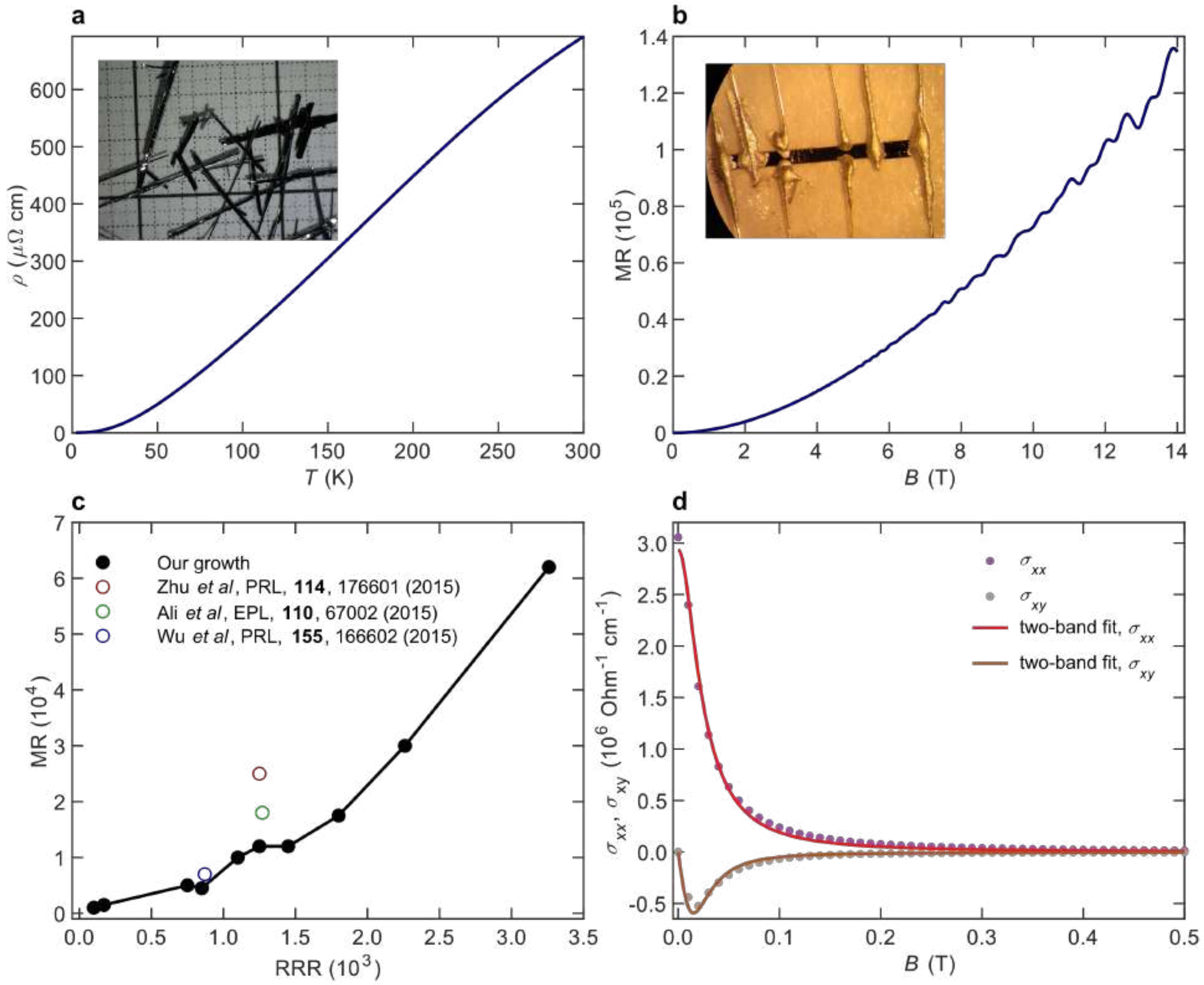
Much higher quality crystals were obtained using tellurium self-flux growth [72]. Elemental W and Te were mixed at a molar ratio of 1:30, loaded into frit-disc alumina crucibles, and sealed in a quartz ampule under vacuum. All steps of materials handling were performed in an Ar glove box with O_2 and H_2O concentration < 0.1 ppm. Quartz ampule, alumina crucibles, and quartz wool for cushioning were heat treated at 800 °C prior to the growth experiment. The tungsten tellurium mixture was heated in a box furnace to 1100 °C at a rate of 30 °C/h, followed by soaking at 1100 °C for 10 h. The metal solution was then slowly cooled down to 650 °C at a rate of 2 °C/h, followed by centrifuging to separate the Te flux from the crystals. To remove any trace of Te flux from the crystal surfaces, they were again vacuum sealed in a quartz ampule and placed on the hot side of a temperature gradient of 400 °C – 190 °C. Although these needle shaped crystals were of much higher quality than those grown by CVT, they were still inferior to the best crystals reported in literature [72]. Subsequent optimization of the growth parameters, such as a higher sample-to-flux ratio of W:Te = 1:50 and W:Te = 1:120, varying cooling rates, changing crucible arrangements, using pre-reacted WTe_2 and excess tellurium as starting materials, led to further improvements but the RRR and MR values were still unsatisfactory; see two points with lowest RRR in Extended Data Fig. 1c.

The deciding factor that led to our best quality crystals was the use of higher purity W (99.999%) and Te (99.9999%) from Furuuchi Chemical Corporation. In addition the starting materials were loaded directly into the quartz tube, using a quartz wool filter instead of the frit-disc alumina crucibles. This allowed us to use larger amounts of starting materials, which should improve the volume to surface ratio of the melt in favor of a lower impurity density in the crystals from contact contamination with the quartz tube. Using our previously optimized mixing ratio of W:Te = 1:120, and a cooling rate of 2 °C/h from 1000 °C to 600 °C followed by centrifuging, we obtained crystals with excellent RRR and MR values on par with the best crystals reported in literature (see Extended Data Fig. 1c). We note that a cooling rate of 1 °C/h resulted in crystals with slightly lower quality.

Device fabrication

WTe_2 crystals were mechanically exfoliated onto an oxidized silicon wafer (290 nm of SiO_2) and suitable flakes were identified by optical microscopy. Standard nano-fabrication techniques were used for device fabrication: electron beam lithography (EBL), inductively coupled plasma etching (ICP), and electron gun metal deposition (E-gun). Separate EBL steps were used to define the mesa and the contact geometries. The WTe_2 mesa was etched with ICP using SF_6 (15 sccm) and O_2 (5 sccm), RF power of 20 W, resulting in an etch rate of ~0.8 nm/s.

The Au reference samples and contacts to WTe_2 flakes were fabricated by E-gun deposition of Ti (2 nm) and Au (30 to 60 nm) followed by a lift-off procedure in acetone. Ar ion milling was used prior to the contact metal deposition for removal of the WTe_2 oxidation layer. For transport characterization, separate devices were fabricated in Hall bar geometry with width $W = 5$ µm, thickness $d \cong 40$ nm, and distance between the voltage contacts $L \cong 3.3$ µm.



Extended Data Fig. 1. **Transport characterization of bulk WTe_2 single crystals.** **a**, Resistivity, ρ, as a function of temperature of our highest purity crystal. At $T = 2$ K, the resistivity is $\rho = 0.23$ μΩ·cm corresponding to $RRR \cong$ 3,250. Inset: optical image of crystals from the optimized quality growth. **b**, Magnetoresistance, $MR = \frac{\rho(B)-\rho(0)}{\rho(0)}$, as a function of magnetic field at 2 K showing $MR \cong$ 62,000 at 9 T. **c**, MR vs. RRR at $T = 2$ K and $B = 9$ T of our different crystals synthesized by flux growth (black dots) in comparison to reported values (open circles) in the literature [72–74]. The black line is a guide to the eye. **d,** Longitudinal and transverse conductivities σ_{xx} and σ_{xy} vs. magnetic field at 4.2 K and their fit to the two band model with resulting parameters $n_e = 2.4\times10^{19}$ cm^{-3}, $n_h = 2.3\times10^{19}$ cm^{-3}, $\mu_e = 5.1\times10^5$ cm^2/Vs, and $\mu_h = 2.7\times10^5$ cm^2/Vs.

Magnetotransport measurements

For bulk transport measurements, crystals with elongated geometry were selected with typical dimensions of width $W = 250$ to 350 µm, thickness $d = 22$ to 240 µm, and distance between the voltage contacts $L =$ 1 to 3.7 mm. Electrical contacts were made with conductive silver epoxy resin (EPO-TEK H20E) using 50 µm diameter gold wire. The epoxy contacts were cured at 150 °C under continuous N_2 flow. An optical image of a representative crystal with current and voltage contacts is shown in the inset of Extended Data Fig. 1b. The transport measurements (temperature and field dependence of resistivity) were carried out in a physical property measurement system (PPMS, Quantum Design) using a *dc* current of 1 to 5 mA for bulk samples and an ac current of 100 nA at frequency of $f = 11.51$ Hz for crystal flakes. Transverse and longitudinal voltages were symmetrized and anti-symmetrized with respect to the magnetic field.

WTe$_2$ is a nearly compensated semimetal with electron and hole pockets contributing to transport [70,72–77]. We thus use a two-band conductivity model for the analysis of the magnetotransport:

$$\sigma_{xx}(B) = e\left(\frac{n_h\mu_h}{1+\mu_h^2B^2} + \frac{n_e\mu_e}{1+\mu_e^2B^2}\right), \qquad \sigma_{xy}(B) = e\left(\frac{n_h\mu_h^2B}{1+\mu_h^2B^2} - \frac{n_e\mu_e^2B}{1+\mu_e^2B^2}\right),$$

where e and B are the elementary charge and applied magnetic field, and n_e, n_h, μ_e, and μ_h are the electron and hole densities and mobilities, which are the fitting parameters.

Extended Data Fig. 1a shows the resistivity $\rho(T)$ measurement of a crystal from our best quality batch. The attained $RRR \cong 3{,}250$ slightly exceeds the values of $RRR = 900$ to 2500 reported in previous landmark studies [72–74,78]. Also, the magnetoresistance (MR) at 2 K (Extended Data Fig. 1b) shows an exceptionally high value of ~ 62,000 at 9 T, even exceeding the values of $MR \cong 42{,}000$ in WP$_2$ [79] and $MR \cong 17{,}500$ in WTe$_2$ [73], measured at the same field. At 14 T, our WTe$_2$ crystal attains $MR \cong 140{,}000$. In Extended Data Fig. 1c, we compare the RRR and MR values in some of our crystals from different batches with previously reported values. By fitting the conductivity at low fields to the two-band model, we obtained the electron and hole concentrations and their mobilities shown in Extended Data Fig. 1d. We took the average electron mobility in our bulk samples to be $\mu_e \cong 2.5 \times 10^5$ cm^2/Vs at 4.2 K, from which we derived the electron momentum-relaxing mean free path $l_{mr} = \hbar k_F \mu_e / e \cong 20$ µm ($k_F = 1.22$ nm^{-1} is the Fermi wavelength [74]).

Transport measurements of our WTe$_2$ flakes in Hall bar geometry with thickness $d \cong 40$ nm show typical conductivities of $\sigma_{flake} \cong 8\times10^4$ Ohm^{-1}cm^{-1} at 4.2 K, which are significantly lower than the bulk conductivities, $\sigma_{bulk} \cong 3\times10^6$ Ohm^{-1}cm^{-1}, leading to estimated effective $l_{mr} \cong 530$ nm in our flakes.

SQUID-on-tip and magnetic imaging

For the magnetic imaging measurements, Pb SOTs were fabricated with diameters ranging from 120 nm to 140 nm following the methods described in Ref. [13]. The SOTs were protected from oxidation by deposition of 3 to 5 nm thick Ti films below and on top of the Pb film. The SOTs included integrated shunt resistors on the tip [80] and had magnetic sensitivity of approximately 50 nT/Hz$^{1/2}$ in applied magnetic field of 60 mT. The SOT readout was carried out using a cryogenic SQUID series array amplifier (SSAA) [81–83]. For height control, the SOT was attached to a quartz tuning fork as described in Ref. [84].

Magnetic imaging was carried out at 4.5 K in 25 µbar residual He pressure in the chamber. For the measurements in Figs. 1 to 4, an ac current of $I_0 = 50$ µA at frequency $f = 186.4$ Hz was applied to the WTe$_2$ or Au samples and the corresponding out-of-plane component of the Oersted field $B_z(x,y)$ was measured by a lock-in amplifier at a constant height of 50 nm above the sample surface. For the scans in Fig. 5, the ac current was varied between $I_0 = 100$ µA and $I_0 = 400$ µA. The images were acquired with a pixel size of 13 nm, acquisition time of 40 ms/pixel, and image size of 430 × 305 pixels.

Current density reconstruction

For the reconstruction of the 2D current density $\boldsymbol{J}(x,y)$ from the measured $B_z(x,y)$, we have used the inversion method described in detail in Ref. [66]. The procedure allows for correction of a small possible tilt of the SOT from the vertical axis and for its finite size, and takes into account the finite thickness d of the sample.

The inversion, however, is an ill posed problem and as such is prone to various artifacts, including high sensitivity to noise and fluctuations, boundaries of the imaging window, fields arising from sources outside the imaging window, ringing at sharp edges due to scanning height related low-pass filtering, and high sensitivity to the assumed height of the sensor. To stabilize the solution with respect to fluctuations, filtering and regularization methods are required [66,85]. As a result, the qualitative features of the resulting $\boldsymbol{J}(x, y)$ are well reproduced; however, the precise quantitative details and the fine structure are less reliable. In the following, we detail the artifacts arising due to ringing and sensor height.

For controlling the scanning height, the SOT is attached to a quartz tuning fork (TF) [84]. The TF is exited electrically at its resonance frequency ~33 kHz and the shift in its phase is monitored as a function of height h upon approaching the sample surface. A threshold of 1° phase shift is defined as the “poking” height, $h = 0$. We then retract the SOT and scan at a nominal height of $h = 50$ nm. We note that the actual effective height of the magnetic imaging should be larger due to the low phase shift threshold, possible surface residues, and the finite thickness of the Ti/Pb/Ti film of the SOT. In addition, the accuracy of the calibration of the vertical displacement of the piezoelectric scanner is limited.

Extended Data Fig. 2 demonstrates the effect of the assumed effective sensor height h on the current distribution reconstructed from the measured $B_z(x, y)$ in the $\theta = 35^\circ$ sample. It shows that the qualitative features of the current flow patterns, including the vortices in the chambers, are robust with respect to the assumed height in the range of $h = 20$ to 150 nm. Inspection of J_y shows, as expected, that low h of 20 and 50 nm results in some broadening of the derived current profiles (Extended Data Figs. 2g,h), whereas higher h of 100 and 150 nm gives rise to enhanced ringing at the edges with negative current visible outside the sample edges (light blue in Extended Data Figs. 2i,j). This ringing is much less pronounced in the J_x distribution in Extended Data Figs. 2b-d and is noticeable in Extended Data Fig. 2e predominantly near the chamber apertures. The ringing in J_x is less significant because the value of J_x in the strip and in the chambers is comparable and its absolute value is a much lower than J_y in the strip. To minimize the ringing in J_y, we therefore use the nominal height $h = 50$ nm for the current reconstruction in all of the figures in the main text.

The ringing at the sharp edges, however, is an unavoidable feature of the low-pass filtering of the inversion procedure and is one of the limiting factors in determining quantitatively the accurate current profiles. This is exemplified by taking a uniform current density J_y in an infinite strip of width $W = 550$ nm and thickness $d = 48$ nm (as expected in a Au strip). We then calculate numerically $B_z(x)$ at $h = 150$ nm, and perform numerical inversion back to current. The resulting reconstructed J_y (green dots in Extended Data Fig. 3a) deviates substantially from the original uniform current density (light green line). It shows ringing both inside and outside of the strip and finite slope at the edges. The ringing artifacts depend on the various parameters of the inversion procedure, including the pixel size, but they cannot be eliminated. Thus deriving precise current profiles from the measured magnetic field has always a limited accuracy. The black solid line in Extended Data Fig. 3a shows the current profile reconstructed from the experimentally measured $B_z(x)$ across the Au strip assuming effective height of 150 nm. It shows a qualitative agreement with the green dotted curve, consistent with a uniform current distribution in the ohmic regime in Au. Assuming a lower effective height results in visible broadening of $J_y(x)$, while higher effective h causes large oscillations.

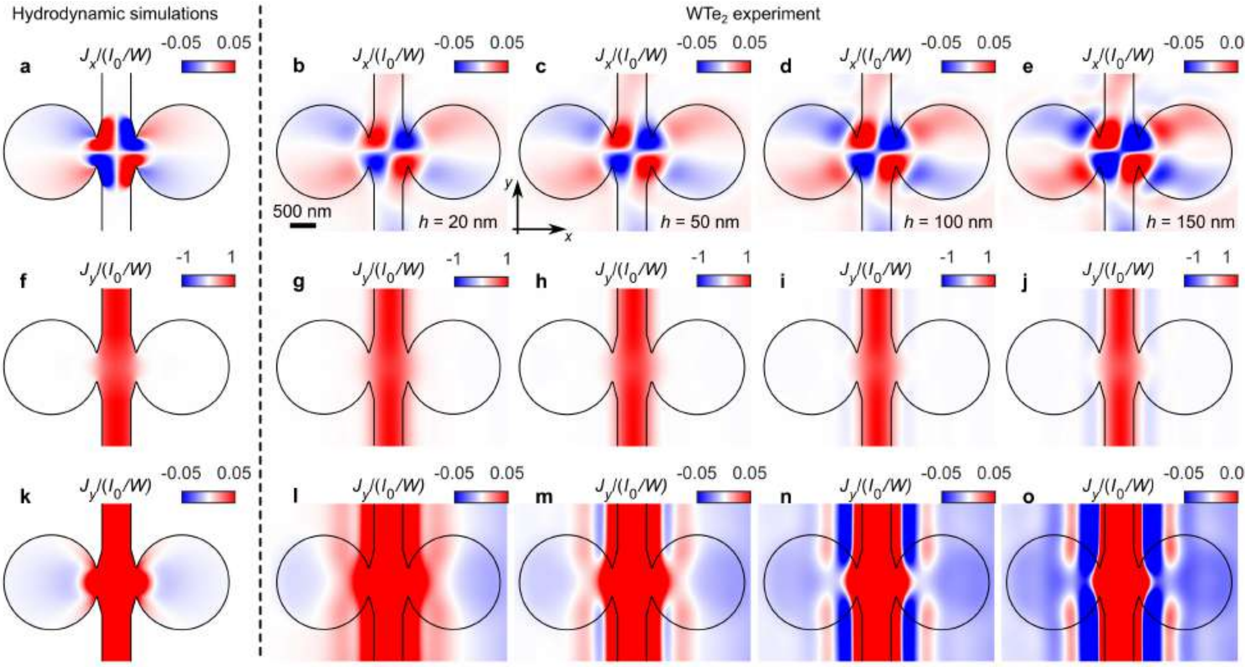
In the strip geometry the current profile in the hydrodynamic regime is given by

$$J_y(x) = J_0 \left[\frac{1 + (\xi/D)\tanh(W/2D) - \cosh(x/D)/\cosh(W/2D)}{1 + (\xi/D - 2D/W)\tanh(W/2D)}\right], \qquad (2)$$

where ξ is the slip length at the boundaries. For no slip conditions ($\xi = 0$), one obtains the familiar Poiseuille profile as shown by the light blue curve in Extended Data Fig. 3b. Even though the no-slip boundary conditions have been considered for the analysis of current profiles [7–10], it has been argued that they are not physical in electron fluids, and that ξ of the order of l_{ee} should be expected even in the case of fully diffuse boundaries [67]. From the analysis of vortex intensity in Fig. 3, we conclude $\xi = 200$ nm or larger in our samples. In this regime, the hydrodynamic current profiles in the strip geometry are almost indistinguishable from the ballistic profiles given by

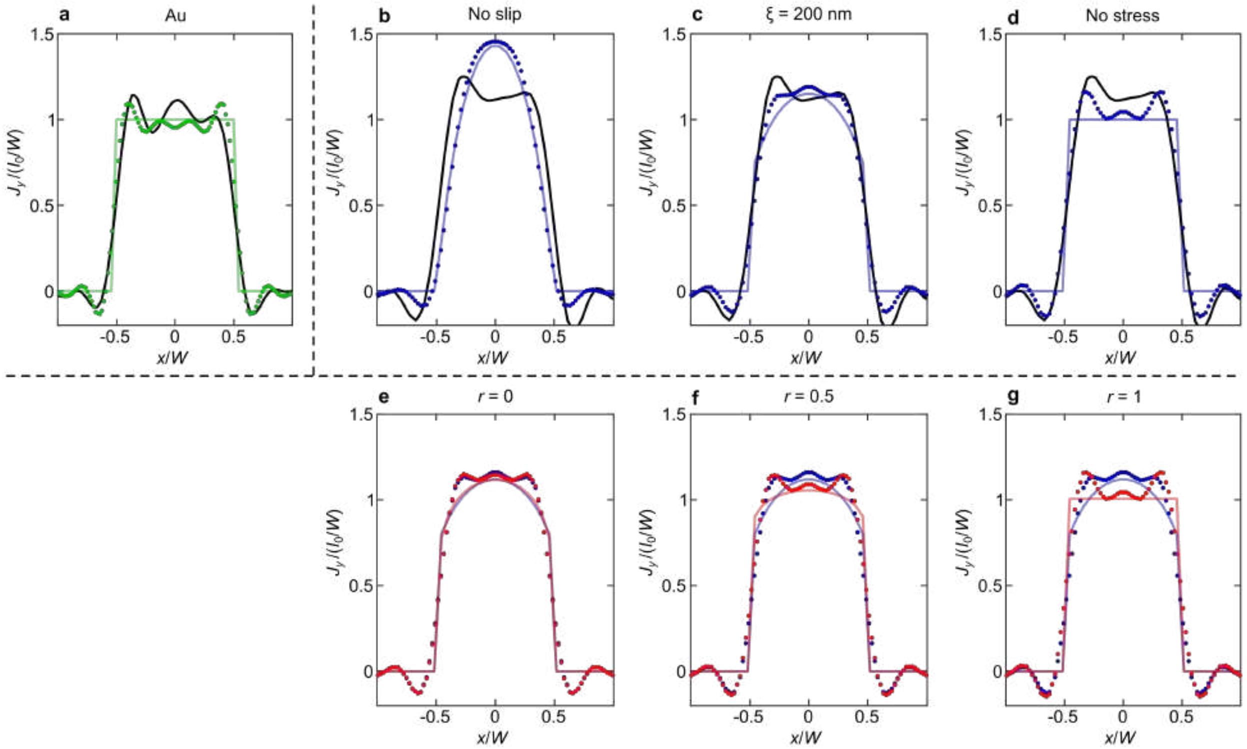
$$J_y(x) = \frac{J_0}{\pi}\int_{-\pi}^{\pi} d\theta \cos^2\theta \left[1 - (1 - r)\frac{\cosh(\csc\theta[W - 2x\,\mathrm{sgn}(\theta)]/2l_{mr}) + \sinh(\csc\theta[W - 2x\,\mathrm{sgn}(\theta)]/2l_{mr})}{-r + \cosh(W\cdot\csc|\theta|/l_{mr}) + \sinh(W\cdot\csc|\theta|/l_{mr})}\right], \qquad (3)$$

where $0 < r < 1$ is the reflectivity coefficient, with $r = 0$ corresponding to fully diffuse boundaries and $r = 1$ describing specular boundaries. Extended Data Figs. 3e-g compare the ballistic profiles for $r = 0$, 0.5, and 1 (red) with the hydrodynamic case with $\xi = 200$ nm (blue). These results show that in a strip geometry, the difference between these cases based on the reconstructed $J_y(x)$ (blue and red dots) is small and experimentally insignificant. In contrast, in the chamber geometry, the vortex stability differs greatly between ballistic and hydrodynamic cases and is hardly affected by the boundary conditions, as shown in Fig. 3.



Extended Data Fig. 2. Dependence of the reconstructed current densities on the assumed SOT scanning height. a, Numerical simulation of $J_x(x, y)$ normalized by the average current density I_0/W in the strip in $\theta = 35^\circ$ sample for $D/W = 0.28$ and $\xi = 200$ nm. The span of the color scale is ± 0.05. **b-e**, Current densities $J_x(x, y)$ reconstructed from the inversion of the measured $B_z(x, y)$ in WTe₂ sample A with $\theta = 35^\circ$ assuming effective SOT scanning heights of $h = 20$ nm (**b**), 50 nm (**c**), 100 nm (**d**) and 150 nm (**e**). The nominal scanning height was 50 nm. The span of the color scale is ± 0.05. **f-j**, Same as **a-e**, but for $J_y(x, y)$ on color scale of ± 1. **k-o**, Same as **f-j**, but on expanded color scale of ± 0.05. The J_y vortex counterflow current (light blue) is resolved in the chambers on a large artificial ringing background outside the strip edges.

Taking the above limitations into account, we now inspect $J_y(x, y)$ more carefully within the chambers. Extended Data Fig. 2k shows the calculated J_y with color scale expanded 20 times, such that the color of the laminar current in the strip (red) is strongly saturated. On this expanded scale, the J_y counterflow becomes visible (light blue). Note that the density of the vortex current counterflow both in J_x and J_y is only about 1% of the laminar current density in the strip as seen in Extended Data Figs. 2a,k. By expanding the color scale of the experimental $J_y(x, y)$ in Extended Data Figs. 2l-o by the same factor of 20, the counterflow in J_y (light blue) becomes visible. On this expanded scale the ringing in J_y is very pronounced even for $h = 20$ nm and grows significantly with h, but the enhanced light blue signal of J_y in the far side of the chambers is resolved at all values of h and shows little dependence on h. Note that the experimental J_x in Extended Data Figs. 2b-e is presented on the same color scale as J_y in Extended Data Figs. 2l-o, demonstrating that the magnitude of the backflow in J_y matches the scale of J_x in the chambers, providing an independent confirmation for the observation of a vortex. Since the ringing problem in J_x is much less pronounced, the counterflow of the vortex current is readily resolved in J_x despite being only about 1% of the driving current density.



Extended Data Fig. 3. Current profiles in narrow Au and WTe₂ strips. a, A uniform current density $J_y(x)$ in $W = 550$ nm strip (light green line) from which $B_z(x)$ is computed at a height $h = 150$ nm. The $J_y(x)$ (green dotted symbols) is then reconstructed by inversion of the calculated $B_z(x)$, showing the unavoidable distortions and ringing. The $J_y(x)$ reconstructed from the experimental $B_z(x)$ in the Au strip (black line) shows consistency with a uniform current distribution in the ohmic regime. **b**, Same as (**a**) for a Poiseuille current profile (light blue) with $D/W = 0.28$ and no-slip boundary conditions. The reconstructed $J_y(x)$ from the experimentally measured $B_z(x)$ in WTe₂ strip (black) is inconsistent with the theoretically reconstructed $J_y(x)$ (dotted blue) from $B_z(x)$ corresponding to the Poiseuille profile. **c**, Same as (**b**) for hydrodynamic flow with $D/W = 0.28$ and slip length $\xi = 200$ nm (light blue) showing good correspondence between the theoretically reconstructed $J_y(x)$ (dotted blue) and the experimentally derived $J_y(x)$ (black) in accord with the conclusions in the main text. **d**, Same as (**b**) for hydrodynamic flow with $D/W = 0.28$ and no-stress boundary conditions (light blue). The reconstructed theoretical $J_y(x)$ (dotted blue) underestimates the experimentally derived $J_y(x)$ (black) supporting the conclusion of a finite slip length. **e**-**g**, Comparison between theoretically calculated current profiles in the hydrodynamic regime with $\xi = 200$ nm (light blue line) and in the ballistic flow (light red line) with boundary reflectivity coefficients of $r = 0$ (fully diffuse) (**e**), $r = 0.5$ (**f**), and $r = 1$ (specular) (**g**). The solid lines show $J_y(x)$ calculated from Eqs. 2 and 3 while the dotted lines are the current profiles reconstructed from the calculated corresponding $B_z(x)$. These results demonstrate the difficulty in using reconstructed current profiles in strip geometry for distinguishing between the hydrodynamic flow with finite slip length and the ballistic transport, in contrast to vastly different vortex stability phase diagrams in these two regimes.

COMSOL numerical simulations

The 2D finite-element numerical simulation of an ohmic electron flow and of transport described by Eq. 1 for the ohmic, ballistic and hydrodynamic regimes, as discussed in the main text, were carried out using COMSOL Multiphysics 5.4. We used the Coefficients Form PDE module, which solves the general equation:

$$e_a \frac{\partial^2 \boldsymbol{u}}{\partial t^2} + d_a \frac{\partial \boldsymbol{u}}{\partial t} + \nabla \cdot (-c\nabla \boldsymbol{u} - \alpha \boldsymbol{u} + \gamma) + \beta \cdot \nabla \boldsymbol{u} + a\boldsymbol{u} = f,$$

where the field $\boldsymbol{u}$ is:

$$\boldsymbol{u} = \begin{pmatrix} \phi \\ J_x \\ J_y \end{pmatrix},$$

and the coefficients were chosen to match Eq. 1:

$$e_a = d_a = \alpha = \gamma = f = 0,$$

$$c = \begin{pmatrix} 0 & 0 & 0 \\ 0 & D^2 & 0 \\ 0 & 0 & D^2 \end{pmatrix}, \quad a = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \beta = \begin{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} \sigma \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ \sigma \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{pmatrix},$$

where D is the Gurzhi length, σ is the conductivity, and the source and drain are simulated by Dirichlet boundary conditions on the potential ϕ. The resulting current density field $\boldsymbol{J}(x,y)$ was normalized by the average current density $J_0 = \int_W J_y(x,y)dx/W$ in the strip. A complete description of the problem is given by employing boundary conditions. The boundary conditions for the perpendicular current were $\boldsymbol{J}_\perp = \boldsymbol{J} \cdot \hat{n} = 0$ in all simulations. For the tangential current, $\boldsymbol{J}_\parallel = (\boldsymbol{J} - \boldsymbol{J}_\perp \cdot \hat{n})$, we have used the three types of boundary conditions described in the main text: (*i*) no slip, $\boldsymbol{J}_\parallel = 0$, (*ii*) no stress, $\hat{n} \cdot \nabla \boldsymbol{J}_\parallel = 0$, and (*iii*) a finite slip length, $\boldsymbol{J}_\parallel = \xi \hat{n} \cdot \nabla \boldsymbol{J}_\parallel$. The dimensions of the simulated devices were chosen to be equal to those of our samples, with the width of the central strip $W = 550$ nm in the dual chamber geometry and the disk radius of the chambers of $R = 900$ nm.

For derivation of the vortex stability phase diagram, a geometry depicted in Fig. 3 was used, with variable D and θ. The vortex current I_v for each simulated geometry and D was calculated according to

$$I_v(D,\theta) = \tfrac{1}{4}\int_l \left(|J_y(x,0)| - J_0 \right) \mathrm{d}x,$$

where the integral was carried out along the cross-section through the central horizontal line connecting the centers of the disk chambers. For $D/W > 1$, the maximum of I_v occurs at $\theta = 60°$ when the aperture and the radius of the disk chambers form an equilateral triangle. An equilateral triangle in the context of hydrodynamic flow was also reported in Ref. [23]. Upon D/W decreasing, the maximum point of I_v shifts to lower θ.

Supplementary Videos 1 and 2 show the evolution of the streamlines upon decreasing θ in the hydrodynamic and ballistic regimes, respectively. As θ increases, on approaching the vortex-to-laminar transition, the laminar streamlines (red) penetrate deeper into the chambers causing distortion of the vortex into a banana shape. In the ballistic case the vortex (blue streamlines) is pushed out of the chamber as a whole, while in the hydrodynamic regime it splits into two vortices at the top and bottom of the chambers. This enhanced stability of the two-vortex solution in the hydrodynamic regime is well captured by the analytical estimates presented in the theory section below.

Theory

Angular scattering, diffusion along the Fermi surface and para-hydrodynamics

This section aims to provide a microscopic justification of Eq. 1 which is used as a benchmark model in the main text. We consider a Fermi gas in two dimensions with a cylindrically symmetric dispersion and a circular Fermi surface. Transport in the system weakly perturbed away from equilibrium is described by a steady-state carrier distribution of the form $\bar{f}(p,r) = \bar{f} + \delta f(r,\varphi)$, with $\bar{f}$ being the equilibrium distribution and $\delta f(r,\varphi)$ a perturbation, where φ is the angle on the Fermi surface, p is electron momentum and $r = (x,y)$. The steady-state distribution satisfies the linearized kinetic equation:

$$v_F \cos\varphi \frac{\partial \delta f}{\partial x} + v_F \sin\varphi \frac{\partial \delta f}{\partial y} - \hat{I}\delta f = -\mathbf{F} \cdot \nabla_p \bar{f} \tag{4}$$

Here $\hat{I}$ is the linearized collision operator of the elastic scattering and $\mathbf{F}$ is an external force. Equation 4 is valid for a collision operator of a general form; below, we apply it to describe scattering at the sample upper and lower surfaces, the process discussed in the main text. Following [59], we expand our perturbation over angular harmonics:

$$\delta f(y,\varphi) = \int \frac{d^2k}{(2\pi)^2} \sum_{m=-\infty}^{m=\infty} f_m(k) e^{im\varphi + i\boldsymbol{kr}}.$$

Statistical isotropy of scattering means that the angular harmonics are eigenfunctions of the scattering operator: $\hat{I} f_m = -\gamma_m f_m$. Because of the $\cos\varphi$ and $\sin\varphi$ structure of the streaming term in Eq. 4, the harmonics $f_m(k)$ satisfy a tridiagonal system of linear equations. The contribution generated by the m-th harmonic of the right-hand side of Eq. 4 (denoted B_m) satisfies the equation:

$$\gamma_m f_m + iAf_{m+1} + i\bar{A}f_{m+1} = B_m, \tag{5}$$

where $A = (k_x - ik_y)v_F/2$. Following [59], we solve Eq. 5 for the ratios $\alpha_m = if_{m+1}/f_m$, which for $n \neq m$ satisfy the recursive relation

$$\alpha_{n-1} = \frac{\bar{A}}{\gamma_n + A\alpha_{n-1}}.$$

The solution of this problem is given in a closed form as a continued fraction

$$\alpha_{n-1} = \cfrac{\bar{A}}{\gamma_n + \cfrac{|A|^2}{\gamma_{n+1} + \cfrac{|A|^2}{\gamma_{n+2} + \cdots}}}$$

Similarly for $\beta_m = if_{m-1}/f_m$ we obtain the fraction running down:

$$\beta_{n+1} = \cfrac{\bar{A}}{\gamma_n + \cfrac{|A|^2}{\gamma_{n-1} + \cfrac{|A|^2}{\gamma_{n-2} + \cdots}}}.$$

Substituting into Eq. 5 yields a contribution to the m-th harmonic of perturbed distribution as follows:

$$f_m = \cfrac{B_m}{\gamma_m + \cfrac{(kv_F/2)^2}{\gamma_{m-1} + \cfrac{(kv_F/2)^2}{\gamma_{m-2} + \cdots}} + \cfrac{(kv_F/2)^2}{\gamma_{m+1} + \cfrac{(kv_F/2)^2}{\gamma_{m+2} + \cdots}}}. \tag{6}$$

Considering long-wavelength limit, we retain only the terms up to quadratic in wavenumber, which gives dissipation and diffusion terms:

$$\left(\gamma_m + \frac{(kv_F)^2}{4\gamma_{m-1}} + \frac{(kv_F)^2}{4\gamma_{m+1}}\right) f_m = B_m.$$

For $m = 0$, one usually has $\gamma_0 = 0$ due to charge conservation, so that we have pure diffusion. For a locally homogeneous electric field we have

$$v_F \cos\varphi \frac{\partial \delta f}{\partial y} + v_F \sin\varphi \frac{\partial \delta f}{\partial x} - \hat{I}\delta f = -e\frac{\partial \bar{f}}{\partial \epsilon}\left(v_F \cos\varphi \frac{\partial \phi}{\partial y} + v_F \sin\varphi \frac{\partial \phi}{\partial x}\right)$$

$$\gamma_m f_m + iAf_{m+1} + \overline{\imath A} f_{m+1} = B\delta_{m,1} + \bar{B}\delta_{m,-1}.$$

Here, $2B = e(E_x + iE_y)v_F \partial \bar{f}/\partial \epsilon$, where ϵ is electron energy. In this case, Eq. 6 is the current-field relation with a nonlocal conductivity:

$$\boldsymbol{J_k} = \sigma(k)\boldsymbol{E_k}, \tag{7}$$

$$\sigma(k) = \frac{ne^2/m}{\gamma_1 + \Gamma(k)}, \qquad \Gamma(k) = \frac{(kv_F/2)^2}{\gamma_2 + \frac{(kv_F/2)^2}{\gamma_3 + \cdots}}. \tag{8}$$

In the long-wavelength limit set by the $m = 3$ harmonic decay rate, such that $(kv_F)^2 < 4\gamma_2\gamma_3$, Eqs. 7 and 8 give

$$\frac{(kv_F)^2}{4\gamma_1\gamma_2}\boldsymbol{J_k} + \boldsymbol{J_k} = -\sigma\boldsymbol{E_k}. \tag{9}$$

After applying a Fourier transform, we recover Eq. 1 from the main text with $D = v_F/\sqrt{4\gamma_1\gamma_2}$:

$$-D^2\nabla^2\boldsymbol{J} + \boldsymbol{J} = -\sigma\nabla\phi. \tag{10}$$

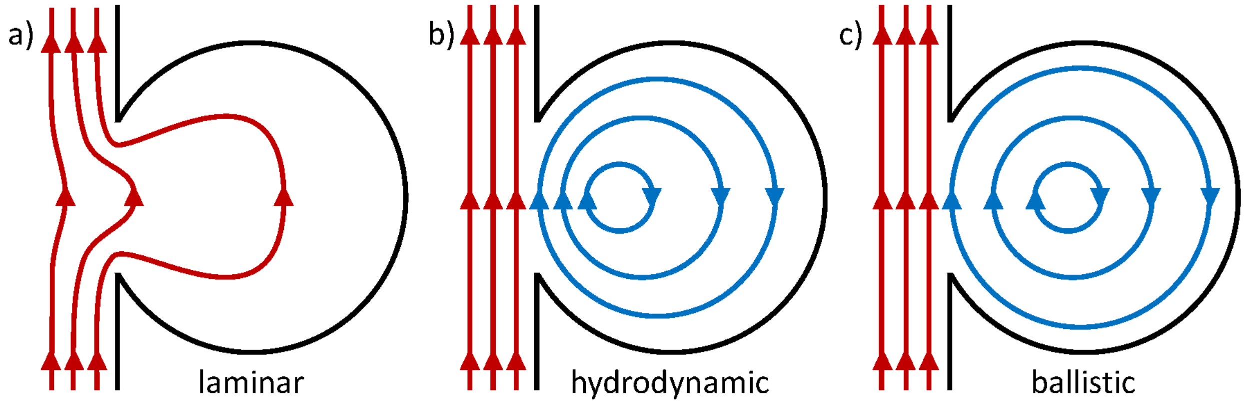
Notably, this equation is applicable not only in the hydrodynamic and ohmic regimes but also in the ballistic regime. This behavior is unique to the situation when $\gamma_3 \gg \gamma_2$, since in this case the condition for the length scales $(kv_F)^2 < 4\gamma_2\gamma_3$ used to derive Eq. 10 is valid in both the hydrodynamic and ballistic regimes. This is in contrast to the conventional electron fluids, where Eq. 10 is valid only in the hydrodynamic and ohmic regimes, but not in the ballistic regime. The extended validity range of Eq. 10 is a salient feature due to small-angle scattering.

The relation between our interpretation of the observed hydrodynamic behavior in terms of Eq. 10 and the assumption that the rates $\gamma_1, \gamma_2, \gamma_3$ … are determined by the small-angle scattering on sample surfaces can be further substantiated as follows. The requirement for Eq. 10 to hold is $\gamma_3 \gg \gamma_2 \gg \gamma_1$. This condition can be approximately fulfilled for small-angle scattering that leads to angular diffusion. Indeed, in this case we have $\hat{I} \approx \gamma \frac{\partial^2}{\partial\varphi^2}$, and therefore $\gamma_m = \gamma_1 m^2$. The condition for the long-wavelength limit is then $kv_F < \sqrt{4\gamma_2\gamma_3} = 12\gamma_1$. In terms of the effective momentum-relaxation length $l_{mr} = v_F/\gamma_1$, the condition takes the form $kl_{mr} < 12$, which is not too restrictive. Indeed, if we put $k \cong 1/W$, the Eq. 10 with $D = l_{mr}/4$ is expected to work reasonably well for $l_{mr} < 12W$, a regime well satisfied for our parameters. In contrast, for large-angle scattering, the decay rates for different harmonics take similar values, $\gamma_3 \cong \gamma_2 \cong \gamma_1$. In this case, depending on the size of $(kv_F)^2$, Eq. 8 can only be truncated at zeroth order or never, thus precluding a regime where Eq. 10 holds.

Hydrodynamic vs. ballistic vortex formation: general considerations and scaling analysis

The discussion in this section provides qualitative arguments in support of the hydrodynamic origin of the observed vortices, as suggested by our simulation results. In a system of size W, the flow pattern depends on two dimensionless parameters, l_{ee}/W and l_{mr}/W. The hydrodynamic regime occurs when $l_{ee}/W \to 0$ and $l_{mr}/W \to \infty$, whereas the ballistic regime corresponds to $l_{ee}/W \to \infty$ and $l_{mr}/W \to \infty$, and the ohmic regime takes place when both $D/W \to 0$ and $l_{mr}/W \to 0$. To gain insight into the character of the flow patterns in all three limits we consider the dissipation in the flow: $\int dxdy\left[\sigma|J|^2 + \sum_{ij}(\sigma_{ij})^2/2\eta\right]$. Here, the first term is the ohmic dissipation, while the second term is the viscous dissipation, determined by the stress $\sigma_{ij} = \eta\left(\partial v_i/\partial x_j + \partial v_j/\partial x_i\right)$. The physical flow can be obtained by minimizing this dissipation functional supplemented with suitable boundary conditions, a procedure known as the principle of minimum entropy production.

The flow in the ohmic limit tends to minimize the current density everywhere, while the ballistic limit tends to minimize stress globally. Finally, the hydrodynamic limit minimizes stress locally, but not globally (Extended Data Fig. 4). As discussed below, the vortex stability phase boundary and the crossovers between different regimes can be obtained directly from the analysis based on the principle of minimum entropy production.



Extended Data Fig. 4. Schematic streamlines for purely ohmic, hydrodynamic and ballistic flow. a, If the sample is purely ohmic, the current leaks into the chamber, forming a current dipole decaying as inverse distance squared. **b**, For a purely hydrodynamic flow, no laminar current (red) leaks from the strip into the chamber. Instead, a vortex forms in the vicinity of the aperture in the chamber (blue) in order to decrease the shear due to the gradient in the velocity profile. **c**, In a purely ballistic flow, only the geometry dictates the streamlines, producing a vortex (blue) whose center is positioned near the chamber center.

Hydrodynamic-to-ohmic and ballistic-to-ohmic crossovers

Since there is no vortex in the ohmic regime, the transition to this regime (upon a decrease in l_{mr}, for instance) from either hydrodynamic or ballistic flow can be understood as a weakening of the vortex in the chamber. As vortex weakens, the laminar current j_l from the strip is expected to penetrate deeper and deeper into the chamber. As a result, the laminar streamlines superimpose onto the vortex streamlines in panels (b) or (c) in Extended Data Fig. 4. In the core of the ballistic or hydrodynamic vortex, the superimposed laminar flow points in the direction of the strip flow, which means that the combined flow pattern in the presence of a small laminar component has its vortex center (where the current vanishes) displaced further away from the aperture compared to the purely ballistic or hydrodynamic case.

In the following analysis, we consider three components that comprise the total current density: laminar current j_l, hydrodynamic current j_{hyd}, and ballistic current j_{bal}. The aperture presents a small dipolar source and sink of the laminar component of the current, which leaks from the strip into in the chamber. Along the chamber boundary and as a function of the distance δ from the aperture, the density of the laminar component thus decays as $j_l \sim 1/\delta^2$ (for $l_{mr} > W$). In contrast, the vortical components (blue in Extended Data Fig. 4) do not flow into or out of the strip, but circulate entirely within the chamber. The ballistic vortex is almost rotationally symmetric and is positioned close to the center of the chamber. As a result, its density along the chamber boundaries is almost constant, $j_{bal} \sim const$. The hydrodynamic vortex, in contrast, forms directly outside the aperture, producing a flow that decays as $j_{hyd} \sim 1/\delta$ far outside the vortex core. Thus, if the laminar flow is strong enough to push the vortex core out to the far side of the chamber, it inevitably overpowers the flow in the upper and lower segments of the chamber, meaning that no vortical flow remains in either the hydrodynamic or the ballistic case. However, for hydrodynamic flow, the streamlines are not uniquely fixed by the geometry, and more than one vortex might form. This is indeed the case, as we demonstrate next: in the hydrodynamic regime, vortical flow can be stabilized by creating two vortex cores even if the single, large vortex is annihilated. In contrast, ballistic flow cannot form two vortex cores next to each other, because ballistic trajectories intersect each other without any effect, which means that the vortex cores do not repel each other, and instead merge into a single large vortex.

Principal components of hydrodynamic flow

To illustrate the basic physics behind the appearance and disappearance of vortices in the chamber, we develop a simple model which allows us to estimate dissipation and choose a state that minimizes it. We want to model the vortex that forms when current in a strip of width W leaks into a large chamber or open space through an aperture in the strip wall at $x = 0$ of size $\Delta = 2R\sin(\theta/2)$, where R is the chamber radius. The relevant current densities are j_c in the strip, j_v of the vortex, and j_l the laminar current in the open space. These current densities are taken to be additive, where the relative amplitudes are determined by the requirement that the resulting flow minimizes the dissipation.

First we consider laminar flow, which can be modelled as a dipole describing current that flows out into the $x > 0$ half-plane in the interval $-\frac{\Delta}{2} < y < 0$ and returns in the interval $0 < y < \Delta/2$. For concreteness, we take the profile of the outflow through the aperture to be parabolic of the form $\sim y(|y| - \Delta/2)$, with the y dependence sign-changing with a zero net current. For weak ohmic dissipation, a laminar ballistic flow in a half plane injected through an aperture with such a profile yields

$$\boldsymbol{j_l}(x,y) = j_{l0}\int_{-\Delta/2}^{\Delta/2} dy' \frac{y'(|y'| - \Delta/2)}{(y-y')^2 + x^2}\begin{pmatrix} x \\ y-y' \end{pmatrix}, \tag{11}$$

with the x component odd in y and the y component even in y. This is a dipole source, and the integral can be done analytically. Expanding in large x for $y = 0$ yields $|\boldsymbol{j_l}(x,0)| = \frac{j_{l0}\Delta^4}{6x^2}$, meaning that the current decays inversely with distance squared at large distances from the aperture, as expected.

Next we consider a vortex positioned inside the chamber in proximity to the aperture. We assume that it has the shape of a Kaufmann vortex, with current density in polar coordinates given by $\boldsymbol{j_v}(r,\varphi) = j_{v0}\, rR_c/(r^2 + R_c^2)\hat{\boldsymbol{e}}_\varphi$. Here, R_c is the size of the vortex core, and j_{l0} and j_{v0} measure the strength of the laminar and vortical flows, respectively. We assume that the vortex is located at $x = d_v$, and account for the boundary conditions close to the walls by adding a counter-rotating image vortex centered around $x = -d_v$. We further impose a condition that the total current I_0 is conserved in any cross-section with a fixed y, which leads to the condition that in the presence of the laminar outflow into the half-space, the current in the strip at $y = 0$ becomes $I_0 - c_{c\rho}j_l\Delta$, where $c_{c\rho}$ is a dimensionless geometrical factor.

Turning to the discussion of dissipation, we note that the decreased current density in the strip also reduces viscous dissipation, which we account for by subtracting the current densities from both the laminar current and the vortical flow. The current profile across the strip is parabolic if $D{\sim}W$. However, in the hydrodynamic limit when $D \ll W$, the current profile becomes very flat and the viscous dissipation is only relevant in a small boundary region. In this latter case, the term for the viscous dissipation acquires an additional factor of D in the denominator. We therefore write an approximate expression for the total dissipation in our system area which is affected by the chamber (i.e. for the chamber itself and for the strip section between $-\Delta/2$ and $\Delta/2$), in which we treat the laminar and vortical flows as additive contributions. This gives

$$P(\boldsymbol{j_l}, \boldsymbol{j_v}, R_c) = P_{\rho,W} + P_{v,W} + P_{\rho,R} + P_{v,R} + P_{\rho,mixed} + P_{v,mixed}. \tag{12}$$

Here, ohmic dissipation $P_\rho = P_{\rho,W} + P_{\rho,R} + P_{\rho,mixed}$ measures the local current density squared, integrated over the system area, while the viscous dissipation $P_v = P_{v,W} + P_{v,R} + P_{v,mixed}$ measures the square of the current density gradients. For the central strip, the effective area is $W\Delta$ for current in the strip, while it is Δ^2 for the laminar flow that penetrates from the strip into the chamber, and R_c^2 for the vortical flow. For the viscous dissipation, we note that the gradients are usually smooth and essentially cancel the area integral. Only in the strip, for small $D \ll W$, this is not true and the current density gradients are restricted to a narrow boundary region of size D. Therefore, $P_{\rho,W} = \frac{\rho}{2}(I_0 - c_{c\rho}j_l\Delta)^2\frac{\Delta}{W}$ is the ohmic dissipation in the strip section, while $P_{v,W} = \frac{\eta}{2}(\frac{I_0}{W} - c_{c\eta}(j_l + j_v))^2\frac{\Delta}{\min[D,W]}$ is the viscous dissipation in the same area (ρ is the sheet resistance). The leading term, proportional to $\frac{\rho}{2}{I_0}^2$, does not affect the competition between the vortex and no-vortex states. Analogously, in the chamber we find $P_{\rho,R_c} =$

$\frac{\rho}{2}(c_{l\rho}j_l^2\Delta^2 + c_{v\rho}j_v^2R_c^2)$ and $P_{v,R_c} = \frac{\eta}{2}(c_{l\eta}j_l^2 + c_{v\eta}j_v^2)$, up to logarithmic corrections in R/W. Here, $c_{c\eta}$, $c_{c\eta}$, $c_{l\rho}$, $c_{v\rho}$, $c_{l\eta}$, $c_{v\eta}$ are dimensionless form factors.

One implication of the construction of the flow in terms of j_l and j_v is that for both ohmic and viscous dissipation, the mixed terms in Eq. 12 between both components are negligible, a property that greatly simplifies the analysis. From our general considerations, we know that the effective size of the vortex goes to zero as the chamber opening is made smaller, i.e. $R_c(\Delta\rightarrow 0) = 0$. As the first main observation, we thus find that $P_{\rho,R_c}(\Delta\rightarrow 0) = 0$. In contrast, the viscous dissipation $P_{v,R}$ does not depend explicitly on the size of Δ (and R_c), and therefore the ratio $P_{\rho,R_c}/P_{v,R_c} \rightarrow 0$, i.e. for small apertures, the ohmic dissipation is negligible compared to the viscous one. Consequently, for small opening sizes Δ, the solution is independent of P_{ρ,R_c}. This suffices to find a unique solution to j_{l0} and j_{v0} (cf. Eq. 11 and below), and subsequently construct an upper bound for the existence of a single vortex depending on whether the total dissipation is smaller with rather than without vortical flow. The vortex is energetically favorable as long as

$$\frac{\Delta_1}{W} < \frac{c_{l\eta}c_{v\eta}}{c_{c\eta}^2(c_{l\eta} + 2c_{v\eta})}\frac{\min[D,W]}{W}. \tag{13}$$

This result provides the general form of the vortex stability phase diagram: for $D/W \ll 1$, the vortex to no-vortex phase transition line is linear with $\Delta_1/W \cong \theta$, while for $D/W > 1$, θ saturates at a finite value, consistent with the numerically derived phase diagram in Figs. 3a,b.

Using the same ansatz, but starting from two vortex cores, the same condition Eq. 13 is recovered for the limiting opening size Δ_2, but where $c_{v\eta}$ is replaced by $4c_{v\eta}$, (the factor 4 is due to the square of the current which enters in the dissipation) to account for the viscous dissipation from two cores. Since it holds that $\Delta_2 > \Delta_1$ this means that there is a narrow range of opening sizes where not one but instead two vortices can form. We note that by the same argument, even more vortex cores might be stabilized. However, at the same time, the positive effect of the vortex formation on the stress at the strip opening diminishes, making these more exotic solutions incompatible with the geometrical constraints imposed by the chamber. Numerically, all form factors are of order 1 and depend little on the particular choices of integration cutoffs. Specifically, for the dipole laminar flow and a Kaufmann vortex, we find that $c_{l\eta} \approx 4.0$ and $c_{v\eta} \approx 2.0 + 0.8\log(W/R_c)$. We further estimate that $W/R_c \sim 3$ and $c_{c\eta} \sim 1/2$, in which case $\Delta_1 < 4.7\min[D,W]$ and $\Delta_2 < 6.8\min[D,W]$. These numbers are reasonably close to both the numerical and experimental findings, even though the ratio of Δ_2/Δ_1 seems to be somewhat smaller in the simulations. In summary, for weakly ohmic flow, a single hydrodynamic vortex can form inside the chamber close to the aperture, with a size of vortex core that increases with the aperture size. For large apertures, this vortex becomes unstable, and we find a narrow range of parameters where it is favorable to form two vortices instead. Vortical flow disappears once the aperture size becomes sufficiently large to allow spreading of the laminal currents over the entire area of the chamber.

Electron-electron scattering length in a compensated semimetal

Here, we estimate the electron-electron scattering rate using the fermionic self-energy calculated in the random phase approximation (RPA). To that end, we consider three contributions that might be relevant in reducing l_{ee} below the values reported in [10]. They are all related to the compensated nature of WTe$_2$, where the band edges of several bands are close to (but not at) the Fermi level.

(i) Firstly, the proximity of the band edges violates the core requirement of semiclassical estimates of the relaxation rate that the ratio of temperature over Fermi energy, $T/E_F \ll 1$, so that a fully quantum-mechanical treatment is needed. In this latter approach [86], the relaxation rate explicitly contains the occupation functions which account for both virtual and thermal fluctuations of the electron fluid. Given a dispersion ϵ_{mk} and eigenfunctions $|u_{mk}\rangle$, where m is the band index and k the momentum, the imaginary part of the self-energy Σ assumes the form,

$$\operatorname{Im}\Sigma(q,\epsilon_{nq}) = \sum_m \int \frac{d^3k}{(2\pi)^3} \operatorname{Im}\left[\frac{C(k-q)\left|\langle u_{nq}|u_{mk}\rangle\right|^2}{1-\Pi(k-q,\epsilon_{mk}-\epsilon_{nq})C(k-q)}\right]\left(b(\epsilon_{mk}-\epsilon_{nq})+n(\epsilon_{mk})\right). \tag{14}$$

Here b and n denote the Bose and Fermi functions, respectively. The dielectric function is determined by the charge susceptibility Π in the RPA-approximation,

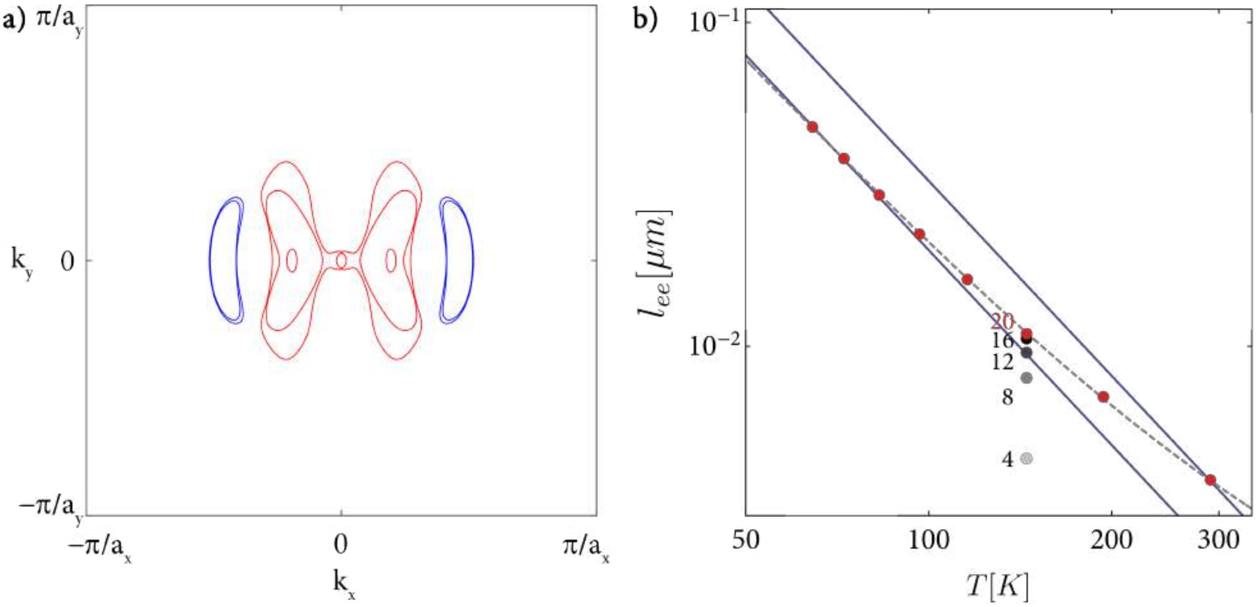
$$\Pi(q,\omega) = \sum_{ij} \int \frac{d^3k}{(2\pi)^3} \left|\langle u_{ik}|u_{j\,k+q}\rangle\right|^2 \frac{n(\epsilon_{ik})-n(\epsilon_{j\,k+q})}{i0^+ + \omega + \epsilon_{ik} - \epsilon_{jk+q}}, \tag{15}$$

while the Coulomb interaction is

$$C(k) = \frac{4\pi e^2}{|k|^2}. \tag{16}$$

(ii) Secondly, the presence of band edges also precludes the extrapolation of relaxation rates obtained at high temperatures under the assumption of a simple Fermi-liquid T^2-scaling, at least *a priori*. We find in particular that the T^2-dependence is violated for temperatures above 100 K.

(iii) Thirdly, the compensated nature of the material with both hole and electron Fermi surfaces requires a high-fidelity calculation with a fine momentum space grid to properly resolve the nesting between electron and hole Fermi surfaces. For example, we observe convergence of the obtained relaxation rates only for grid sizes larger than 100×50×7 in the x-y-z momentum volume.

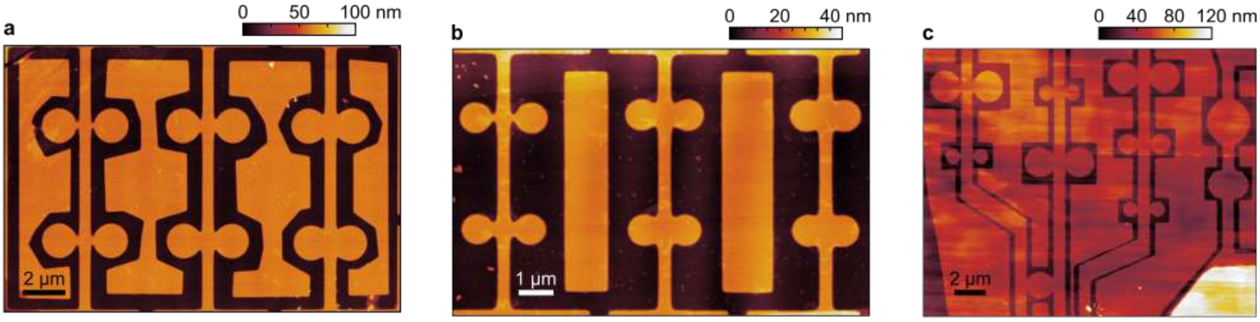


Extended Data Fig. 5. Fermi surface and electron-electron mean free path. a, Fermi surface cut for $k_z = 0$. Typical for a compensated semimetal, small electron and hole pockets appear close to the compensation point. If the hole density is slightly larger than the electron density, the Fermi surface features hole pockets near the Gamma point (red) and electron pockets (blue). **b**, l_{ee} as calculated from Eq. 14 for 20 bands as a function of temperature (red points). For $T = 145$ K, we also show the values for a smaller number of bands. The blue lines denote upper and lower estimates for the T^{-2} dependence of l_{ee}, where the lower one corresponds to the low-temperature asymptotics.

Taking all these issues into account, we calculated l_{ee} for temperatures between 70 K and 300 K, checked for convergence in terms of grid resolution, and extrapolated to lower temperatures based on a power-law fit. To decrease runtime, as an approximation we restricted the effects of the Coulomb interaction to the first Brillouin zone and set the imaginary part of the dielectric function to zero for bands far from the chemical potential. Extended Data Fig. 5 shows the Fermi surface near the compensation point, and the scaling of l_{ee} with temperature, averaged for band 55 which forms one of the hole pockets, taking into account a total of 20 bands closest to the chemical potential. Comparable values are obtained for the other Fermi pockets. As is clearly visible, the low-temperature asymptotic temperature dependence sets in only below 100 K. We also confirmed the convergence of our calculation with respect to contributions from far bands by comparing the obtained relaxation rates for 4, 8, 12, 16 and 20 bands at an intermediate temperature of 145 K. Extrapolating our results to low temperatures under the assumption of a T^{-2} dependence, we obtain $l_{ee} =$ 0.5 µm at $T =$ 20 K and $l_{ee} =$ 4 µm at $T =$ 7 K. We note that these values are lower than the ones reported for the electron-electron mean free path in [10], which was calculated with a much smaller grid resolution. However, they are comparable to the latter's estimate for the phonon-mediated interacting mean free path. While a full analysis is beyond the scope of this work, we point out that the combined momentum-conserving mean free path could thus be even slightly smaller. These estimated lower values of l_{ee} give credence to the mechanism described in the main text whereby a weakly ballistic flow can effectively become hydrodynamic in thin samples.

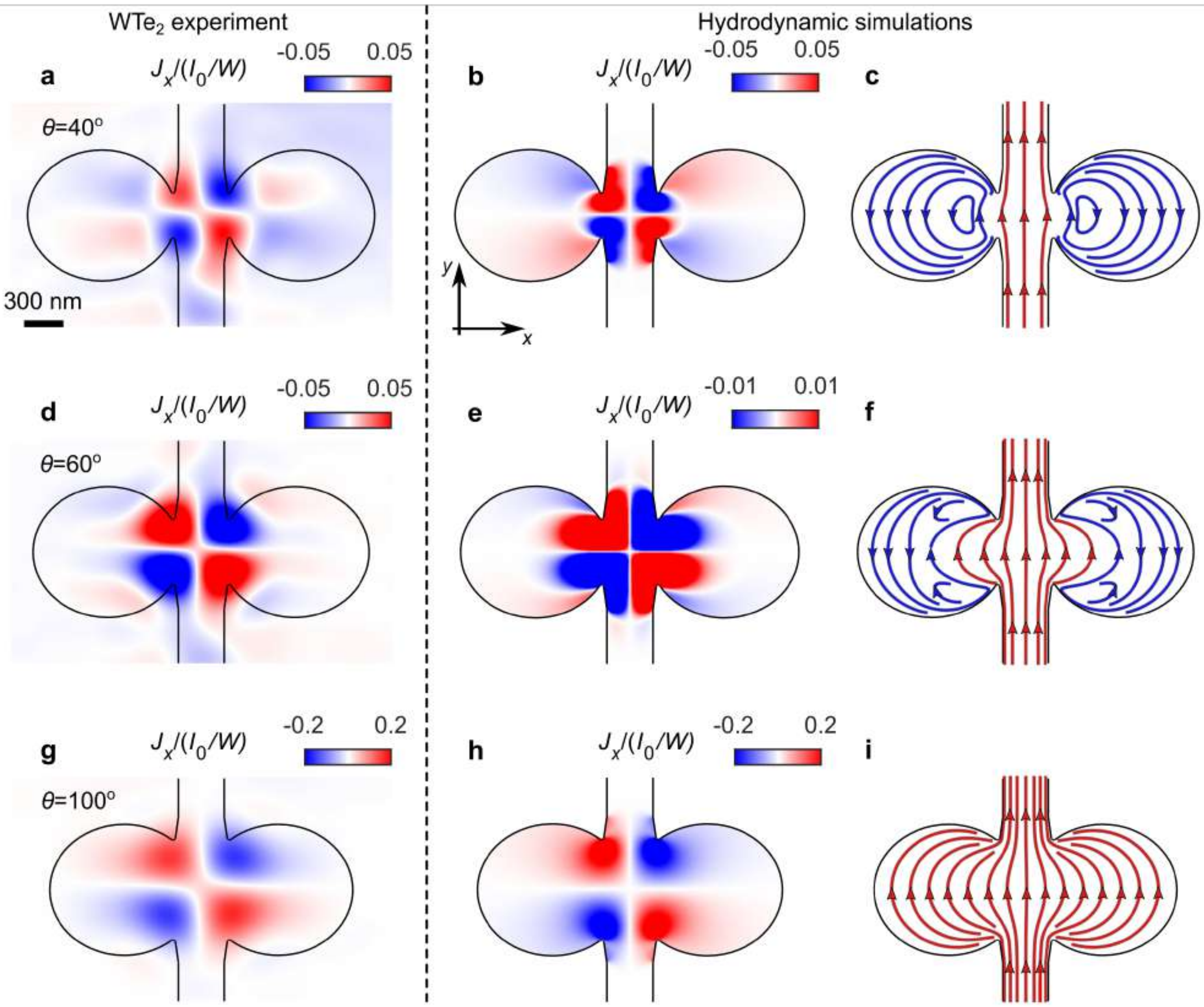
Transition from vortex to laminar flow in additional samples

Analogously to the analysis in the main text (sample A), we show the transition from vortex to laminar flow in two additional samples, B and C. These samples provide additional insight into the dependence of the flow on the geometrical parameters, including the width the central strip, W, the chamber radius, R, and the sample thickness, d. Importantly, samples A and B were fabricated from the same batch of WTe_2 crystals, while sample C was exfoliated from a different batch of lower quality. AFM images of samples A, B, and C are shown in Extended Data Fig. 6.



Extended Data Fig. 6. AFM images of WTe_2 devices. a, AFM image of device A analyzed in the main text with $W =$ 550 nm, $R =$ 900 nm, $d =$ 48 nm, and aperture angles $\theta =$ 20°, 35°, 54°, 72°, 90°, and 120°. **b**, Device B used for Extended Data Fig. 7 with $W =$ 350 nm, $R =$ 450 nm, and $d =$ 23 nm. **c**, Device C with $W =$ 770 nm and $d =$ 30 nm, and $R =$ 950, 725, and 500 nm (Extended Data Figs. 8) and dual-drive geometry at the bottom part (Extended Data Figs. 9).

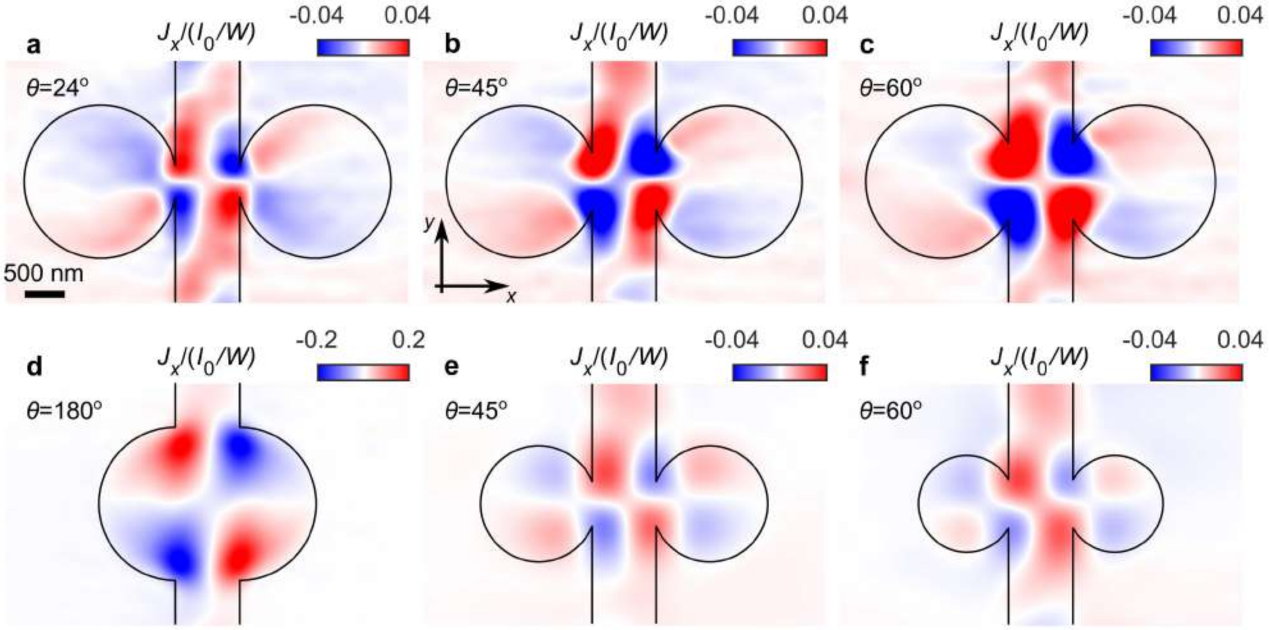
Sample B is characterized by $W =$ 350 nm, $R =$ 450 nm, and $d =$ 23 nm. To avoid current heating due to the narrower central strip, the excitation current was reduced to $I_0 =$ 25 µA. Similar to the behavior in sample A in Fig. 4, the $J_x(x, y)$ images in Extended Data Figs. 9a,d,g show a transition from single-vortex to two-vortex to no-vortex state upon increasing θ. This is in good agreement with simulations of $J_x(x, y)$ (Extended Data Figs. 7b,e,h) as well as simulated current streamlines (Extended Data Figs. 7c,f,i). Here, the two-vortex state occurs at $\theta_t =$ 60° as compared to $\theta_t =$ 54° in sample A. This aperture angle allows us to deduce $D/W =$ 0.35 and resulting $D =$ 123 nm in sample B. Thus, the smaller strip width W of sample B results in an effective vertical upshift of the data points in Fig. 3a.



Extended Data Fig. 7. Transition from single-vortex to two-vortex to laminar flow in sample B. a-c, Measurement of single vortex state in device B with $\theta = 44°$ and corresponding simulations in the hydrodynamic regime with $D = 123$ nm and $\xi = 200$ nm. **a**, Measured current density $J_y(x, y)$ normalized by I_0/W at $I_0 = 25$ µA. **b**, Simulated $J_y(x, y)$. **c**, Simulated current streamlines showing laminar (red) flow in the central strip and vortex flow (blue) in the chambers. **d-f**, Same as (a-c) for $\theta = 60°$ showing banana-shaped vortex at the transition from a single to double-vortex state. **g-i**, Same as (a-c) for $\theta = 100°$, showing laminar flow.

Extended Data Fig. 8 shows $J_x(x, y)$ distributions in sample C, which is characterized by $W = 770$ nm and $d = 30$ nm. Here, the disc chambers of radius $R = 950$ nm show single-vortex state for aperture angles of $\theta = 24°$ (a), $\theta = 45°$ (b), and $\theta = 60°$ (c), and laminar flow for $\theta = 180°$ (d). Although, no chamber with θ between 60° and 180° was available in this sample, the presence of a single vortex at $\theta = 60°$ indicates $D/W > 0.35$. The resulting $D > 270$ nm significantly exceeds the values derived for samples A and B. This may be attributed to the different batches of source material with different microscopic parameters or to variations in fabrication resulting in differences in the top and bottom surface quality.

In addition, single-vortex behavior was found in chambers with $R = 725$ nm and $\theta = 45°$ as well as for $R = 500$ nm and $\theta = 60°$ (Extended Data Figs. 8e,f), which shows the stability of the single-vortex state with respect to chamber radius.



Extended Data Fig. 8. Vortex flow in sample C with different geometrical parameters. Current density $J_x(x,y)$ in sample C with $W = 770$ nm and $d = 30$ nm and various chamber parameters: $\theta = 24°$ and $R = 950$ nm (**a**), $\theta = 45°$ and $R = 950$ nm (**b**), $\theta = 60°$ and $R = 950$ nm (**c**), $\theta = 180°$ and $R = 950$ nm (**d**), $\theta = 45°$ and $R = 725$ nm (**e**) and $\theta = 60°$ and $R = 500$ nm (**f**). Laminar flow is observed in (d), while vortex flow is present in all the rest of the geometries.

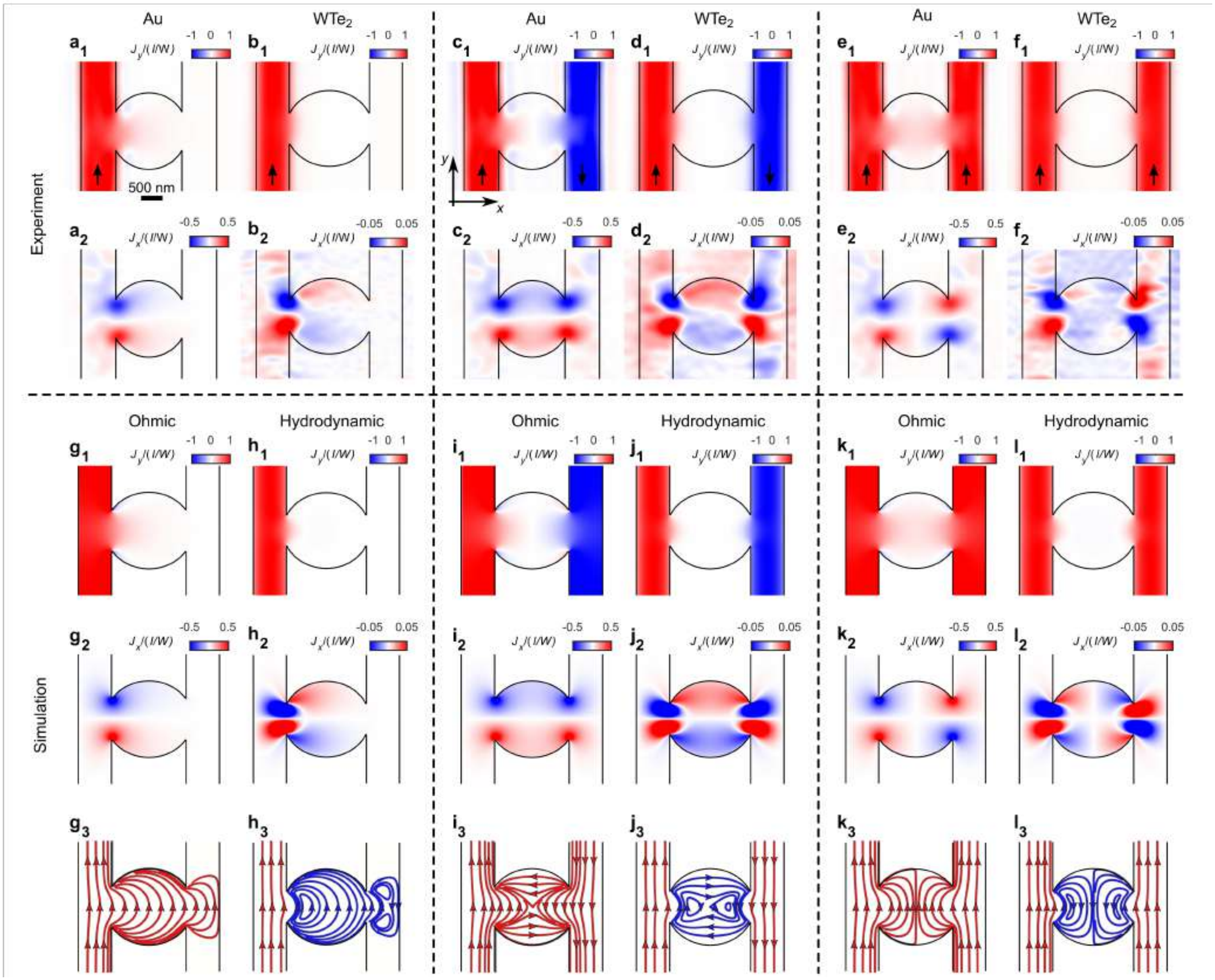
Dual-drive geometry

We have also studied an alternative geometry of a central disk with two apertures with $\theta = 44°$ on opposite sides connected to two current-driven strips patterned in WTe$_2$ sample C with $d = 30$ nm, and in Au film of similar thickness, as shown in Extended Data Fig. 9. An *ac* current of $I_L = 50$ µA was applied to the left strip with source at the bottom and drain at the top. A lower excitation frequency of $f = 17.73$ Hz was used to reduce capacitive currents between the two strips. A separate floating current source was used to apply current to the right strip with three values, $I_R = 0$, -50 µA, and 50 µA.

Panels a$_1$ and a$_2$ in Extended Data Fig. 9 show the $J_y(x,y)$ and $J_x(x,y)$ current distributions in Au sample for $I_R = 0$. In the ohmic regime, the current penetrates substantially into the central chamber similar to Fig. 1j. In the hydrodynamic case of WTe$_2$ (panel b$_1$), in contrast, the J_y is mostly concentrated along the left strip with little protrusion into the chamber, analogous to Fig. 2g and consistent with the numerical simulations in panels g$_1$ and h$_1$. The transverse current, $J_x(x,y)$, reveals a laminar flow in Au (panel a$_2$) and a vortical flow in WTe$_2$ (panel b$_2$) in the chamber, consistent with the numerical simulations in panels g$_2$ and h$_2$, and with the simulated streamlines in panels g$_3$ and h$_3$. This configuration is equivalent to Figs. 1m and 2j of the main text.

Upon applying $I_R = -50$ µA, interesting flow patterns are observed in the central chamber. In the ohmic case, instead of flowing in and out of the chamber as observed for a single drive in panel a$_2$, part of the current traverses the chamber in its lower part (red J_x in panel c$_2$), and then flows down to the bottom drain of the right strip as corroborated by the simulated streamlines in panel i$_3$. In the top half of the chamber, an opposite flow from top source of the right strip to the top drain on the left strip occurs (blue J_x), exchanging part of the currents from the two sources. Remarkably, in the hydrodynamic case, the shear forces of counter propagating currents in the two strips add up constructively and propel a single massive clockwise vortex in the entire chamber as observed in panel d$_2$ and simulated in panel j$_2$. The simulated streamlines in panel j$_3$ show that in this case the currents of the left and right sources do not mix: the laminar streams in the two strips are isolated by the whirlpool in the central chamber.

One would then expect that in the case of copropagating currents in the two strips (panel f_1), the opposing shear forces at the two apertures act destructively, annihilating the massive vortex. Instead, we find that a vortex–antivortex pair is nucleated in the chamber as visualized by $J_x(x, y)$ in panel f_2 and simulated in panels l_2 and l_3.



Extended Data Fig. 9. **Vortex–antivortex formation in dual-drive geometry. a-f**, Experimentally derived current densities $J_y(x, y)$ (top row) and $J_x(x, y)$ (bottom row) in Au and WTe₂ samples. **a-b**, Current $I_L =$ 50 µA is driven in the up direction in the left strip with no current applied to the right strip resulting in a single vortex in the WTe₂ chamber in **b₂**. **c-d**, Counterpropagating currents $I_L = 50$ µA and $I_R = -50$ µA applied to the right and left strips, giving rise to a single massive vortex in **d₂**. **e-f**, Copropagating currents $I_L = 50$ µA and $I_R = 50$ µA applied to both strips which generates a vortex–antivortex pair in **f₂**. **g-l**, Numerical simulations of current densities $J_y(x, y)$ (top row), $J_x(x, y)$ (middle row) and the corresponding streamlines (bottom row) in the ohmic and hydrodynamic regimes for the three current configurations. The laminar streamlines are marked in red and the vortex streamlines in blue. The experimental data were acquired with pixel size of 10 nm, acquisition time of 40 ms/pixel, and image size of 600×350 pixels/image.

Captions of Supplementary Videos

Supplementary Video 1 | Simulations of vortical-to-laminar flow transition in the para-hydrodynamic regime vs. $\boldsymbol{\theta}$. Numerical simulation of the current density $J_x(x, y)$ (top right) and the corresponding streamlines (bottom right) in the double-chamber geometry upon increasing the aperture angle θ for $D/W = 0.28$. The left panel shows the vortex stability phase diagram with no stress boundary conditions as presented in Fig. 3a. The purple dot marks the value of the varying θ along the $D/W = 0.28$ line. For $\theta \leq 54°$, there is a single vortex in each chamber (blue streamlines). Upon increasing θ further, the laminar flow (red streamlines) splits the single vortex in each chamber into two vortices, which are stable up to $\theta \leq 60°$. Finally, for $\theta > 60°$, the laminar streamlines fill the entire area of the chambers.

Supplementary Video 2 | Simulations of vortical-to-laminar flow transition in the quasi-ballistic regime vs. $\boldsymbol{\theta}$. Numerical simulation of the current density $J_x(x, y)$ (top right) and the corresponding streamlines (bottom right) in the double-chamber geometry upon increasing θ for $D/W = 1.5$. The left panel shows the vortex stability phase diagram with no stress boundary conditions as presented in Fig. 3a. The purple dot marks the value of the varying θ along the $D/W = 1.5$ line. With increasing θ, the laminar streamlines (red) gradually penetrate deeper into the chambers, distorting the vortices (blue streamlines) and pushing them towards the outer boundaries. The vortices become extinct at $\theta \cong 150°$ without splitting into double vortices as is the case in the hydrodynamic regime in Supplementary Video 1. For $\theta > 150°$, the laminar streamlines fill the entire area of the chambers.

Hydrodynamic approach to two-dimensional electron systems

Boris N. Narozhny

Abstract The last few years have seen an explosion of interest in hydrodynamic effects in interacting electron systems in ultra-pure materials. One such material, graphene, is not only an excellent platform for the experimental realization of the hydrodynamic flow of electrons, but also allows for a controlled derivation of the hydrodynamic equations on the basis of kinetic theory. The resulting hydrodynamic theory of electronic transport in graphene yields quantitative predictions for experimentally relevant quantities, e.g. viscosity, electrical conductivity, etc. Here I review recent theoretical advances in the field, compare the hydrodynamic theory of charge carriers in graphene with relativistic hydrodynamics and recent experiments, and discuss applications of hydrodynamic approach to novel materials beyond graphene.

Keywords Hydrodynamics · Kinetic theory · Electronic transport · Viscosity · Hall effect · Graphene · Compensated semimetals · Topological materials

1 Hydrodynamics and condensed matter

Collective excitations in solid state physics – phonons, magnons, plasmons, etc. – are often considered in the long-wavelength (small wavevector) limit with the corresponding observables describing long-distance properties of matter. One way to develop a macroscopic theory reflecting such physics [1] is to combine continuity equations (manifesting conservation laws) with thermodynamic arguments to identify how the entropy of the system responds to local density fluctuations of the conserved quantities. Requiring the total entropy

---

Boris N. Narozhny
Institut for Theoretical Condensed Matter Physics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia
E-mail: boris.narozhny@kit.edu

production rate to be non-negative, one may establish the "constitutive relations" between the macroscopic currents and the external bias. Closing the equations with the help of the thermodynamic relations one can complete the description of the long-wavelength dynamics of the system. The resulting theories are macroscopic since their variables are densities of physical quantities and the corresponding currents. They are also phenomenological since they provide no means of calculating the coefficients in the constitutive relations (i.e., the "generalized susceptibilities"). Such approach is justified at distances that are much larger than any length scales corresponding to the underlying "microscopic" scattering processes, the condition that is very often satisfied in experiments.

The most common equation describing the long-wavelength dynamics in solids is the diffusion equation [1]. In the simplest example, spin diffusion [2,3] arises in a system of spin-1/2 particles with a velocity- and spin-independent interaction leaving the total magnetization conserved. This behavior has been observed experimentally (see, e.g., Ref. [4]) and is generally expected to be applicable to a wide variety of spin systems (with the possible exception of one-dimensional integrable models, see Refs. [5,6,7,8]).

Low-temperature charge transport is also often considered to be diffusive [9]. In the simplest case, charge carriers are assumed to be independent and noninteracting, so that their total number is a conserved quantity, while the dominant relaxation process is the electron-impurity scattering described by the transport mean free time, τ. The latter defines both the diffusion constant and electrical conductivity [10] and is still one of the most important quantities characterizing conductive properties of experimental samples. The diffusive behavior is commonly expected to take place in real metals and semiconductors as long as the sample size is large compared to the mean free path (typically, $\ell = v_F\tau$ with v_F being the Fermi velocity) [11] and at low temperatures, $T\tau \ll 1$ [12] (the units with $\hbar = k_B = 1$ are used throughout this paper).

A common feature of the above theories is the decaying (diffusive) nature of collective modes (defined as the normal modes of the set of linearized macroscopic equations). In contrast, the collective modes in conventional fluids, both classical (e.g., water [13,14]) and quantum (e.g., ^{3}He [15]), include also sound waves (with the linear dispersion). This crucial difference can be attributed to the momentum conservation. Indeed, the usual description of a fluid (or a gas, see [16]) assumes a system of "particles" (molecules or atoms) interacting by means of local collisions. In the simplest case (of a single-component, monoatomic fluid) the collisions preserve momentum, and hence overall there are three global conserved quantities – the number of particles, energy, and momentum. If, moreover, Galilean invariance is assumed, then the current is defined by the momentum, which is the key point ultimately leading to the existence of the sound-like collective mode.

The macroscopic theory describing flows of a conventional fluid – namely, hydrodynamics – can be derived in the several ways. One can follow the above prescription using the continuity equations and entropy [1], one can "guess" (or postulate) the constitutive relations based on the Galilean invariance (or, in the relativistic case, Lorentz invariance) [13], or one can use the "microscopic" kinetic theory [16]. The latter approach is justified, strictly speaking, in a dilute gas, but yields the same set of hydrodynamic equations as the more phenomenological methods. This fact is typically attributed to the *universality* of the hydrodynamic approach: the belief that long-distance properties are largely independent of the short-distance (microscopic) details. As a result, strongly interacting fluids (such as water) can be successfully described by the same hydrodynamic theory as an ideal gas [16].

In condensed matter context, hydrodynamic approaches have been applied to phonons [17] (also see the recent experiment [18] and references therein) and magnons [19], while applications to electronic systems [20,21,22] have only recently attracted widespread attention [23,24,25]. This may appear surprising, after all the Fermi Liquid theory originally developed for ^{3}He [15] has become a dominant paradigm in solid state physics. In ^{3}He, the Fermi Liquid theory can be used to derive the hydrodynamic equations [26], so why cannot the same be done in solids? Unlike helium atoms, electrons in solids exist in the environment created by a crystal lattice and can scatter off both lattice imperfections (or "disorder") and lattice vibrations (phonons). In both cases, their momentum is not conserved. As a result, the electron motion is typically diffusive [9], unless the sample size is smaller than the mean free path and the system is "ballistic" [11]. For most typical scattering mechanisms in solids, the mean free path is strongly temperature dependent. In conventional metals [10] electron-impurity scattering dominates at low temperatures, leading to, e.g., the residual resistance. At high temperatures, the main scattering mechanism is the electron–phonon interaction. In many materials, at least one of these two mechanisms is more effective than electron–electron interaction at any temperature, leaving no room for hydrodynamic behavior. In terms of the associated length scales, this statement can be formulated as $\ell_{ee} \gg \ell_{\rm dis}, \ell_{e-ph}$ (in the self-evident notation). If a material would exist, where the opposite condition were satisfied at least in some temperature range, then one would be justified in neglecting momentum non-conserving processes and applying the hydrodynamic theory. For a long time such a material was not known. In recent years, several extremely pure materials became available bringing electronic hydrodynamics within experimental reach [27,28,29,30,31,32,33,34,35, 36].

2 Experimental signatures of hydrodynamic behavior

The parameter regime supporting the hydrodynamic behavior can be readily found in systems where the temperature dependence of key length scales (ℓ_{ee}, $\ell_{\rm dis}$, ℓ_{e-ph}, etc.) is sufficiently different. This may happen, for example, in two-dimensional (2D) systems where the electron-electron scattering length varies with temperature as $\ell_{ee} \sim T^{-2}$ (within the typical Fermi Liquid description), while the contribution of acoustic phonon scattering to the electronic mean free path varies as $\ell_{e-ph} \sim T^{-1}$. At the same time, the low temperature values of ℓ_{ee} are easily surpassed by the mean free path $\ell_{\rm dis}$ in ultrapure samples. Hence, 2D systems may offer an intermediate temperature window [23,24,37,38], where electron-electron interaction is the dominant scattering process and hence appear to be plausible candidates to support the hydrodynamic behavior. It is then not surprising that many experiments on electronic hydrodynamics were focusing on 2D systems and especially on graphene. The latter is a particularly convenient material [27,28,39, 40,41,42,43] where the mean free path remains long up to room temperatures, $\max[\ell_{\rm dis}, \ell_{e-ph}] > 1\,\mu$m. At the same time, at $T \geqslant 150\,$K the electron-electron scattering length decreases to $\ell_{ee} \approx 0.1 \div 0.3\,\mu$m. Since the pioneering work on the nonlocal resistance [27] and Wiedemann-Franz law violation [28], several impressive experiments [39,41,42,43,44,45] aimed at uncovering the hydrodynamic behavior of the electronic system in graphene. In particular, it was suggested that a viscous hydrodynamic flow in electronic systems might exhibit enhanced, higher-than-ballistic conduction [39,44,45]. More recently, several breakthrough experiments [30,34,44,45,46,47,48,49,50,51,52,53,54] demonstrated various distinct imaging techniques making it possible to "observe" the electronic flow in graphene "directly".

Hydrodynamic flow of electrons in solids should be observable not only in graphene, but in any material that is clean enough to satisfy the condition that the electron-electron scattering length is much shorter than the disorder mean free path. In particular, modern semiconductor technology allows fabricating ultra-high-mobility heterostructures [30,32,36,55,56,57], a noticeable improvement since the original observation of the Gurzhi effect [58].

At the same time, the hydrodynamic behavior might be observable in a wide range of novel materials including the 2D metal palladium cobaltate [29], topological insulators (where the conducting surface states may exhibit hydrodynamic behavior), and Weyl semimetals [59,60]. The latter systems have attracted considerable attention since they exhibit a solid-state realization of the Adler-Bell-Jackiw chiral anomaly [61,62,63,64,65]. One of the hallmark manifestations of the anomaly in Weyl systems [59,66] is the recently observed negative magnetoresistance [64,67]. Observation of relativistic Weyl hydrodynamics in these systems is the next milestone in the field.

2.1 Gurzhi effect

In his pioneering work [17,20,21], Gurzhi considered an idealized problem of the electric current flowing in a thin, clean wire. In this case there are two competing scattering processes: the electron scattering off the walls of the wire (i.e., system boundaries) and the electron-electron interaction, either direct or effective (e.g., phonon-mediated). Assume that at the lowest temperatures the electron-electron scattering length is longer than the width of the wire, $\ell_{ee} \gg d$. Then boundary scattering will dominate leading to the approximately temperature-independent resistivity, $\rho \sim 1/d$. Now, the electron-electron scattering length ℓ_{ee} is inversely proportional to some power of temperature (for

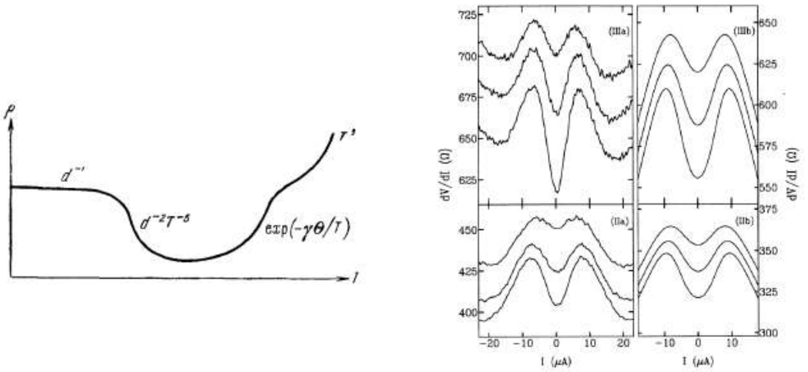


Fig. 1 Gurzhi effect. Left panel: a sketch of the theoretically predicted resistance minimum (reprinted with permission from Ref. [17]; copyright (1968) Uspekhi Fizicheskikh Nauk). Right panel: experimental (IIa and IIIa) and theoretical (IIb and IIIb) differential resistance dV/dI as a function of the current I at the lattice temperatures $T = 4.5$, 3.1, 1.8 K, from top to bottom (reprinted with permission from Ref. [58]; copyright (1995) by the American Physical Society).

the direct electron-electron interaction $\ell_{ee} \propto T^{-2}$ [20], while for the phonon-mediated interaction $\ell_{ee} \propto T^{-5}$ [21], see Fig. 1). As the temperature increases, ℓ_{ee} will eventually become smaller than d. In the limit $\ell_{ee} \ll d$, the resistivity will be determined by the electron-electron scattering, $\rho \sim \ell_{ee}/d^2$ [20,21] and hence will decrease with the increasing temperature. This effect can be seen as the electronic analogy of the crossover between the Knudsen (molecular) flow and the Poiseuille (viscous) flow in a rarefied gas driven through a tube [68].

The above conclusion crucially depends on the assumption that the effective mean free path d^2/ℓ_{ee} is much smaller than the length scale ℓ_{dis} describing bulk momentum-relaxing processes (i.e., electron-impurity or electron-phonon scattering). Then the electronic momentum is approximately conserved and one can introduce the hydrodynamic description (the expression for ρ follows from the standard expression for the kinematic viscosity, $\nu = v_F \ell_{ee}/3$ [20]).

Once the effective mean free path due to electron-electron interaction exceeds the disorder scattering length, $d^2/\ell_{ee} \gg \ell_{\text{dis}}$, the system becomes diffusive and the resistivity resumes its usual growth with temperature. Hence, $\rho(T)$ is expected to exhibit a minimum, see Fig. 1, the result now known as the Gurzhi effect.

A direct observation of the Gurzhi effect in metals is hindered by several factors: in addition to the electron-impurity and electron-phonon scattering, Umklapp scattering, nonspherical Fermi surface shapes, or Kondo effect may all contribute to the temperature dependence of the resistivity. An elegant way around these obstacles was suggested by de Jong and Molenkamp [58]. They used 2D wires defined electrostatically in the two-dimensional electron gas (2DEG) in semiconductor (GaAs/AlGaAs) heterostructures. Given the weakness of the electron-phonon coupling in this system, it was possible to control the electronic temperature selectively without changing the temperature of

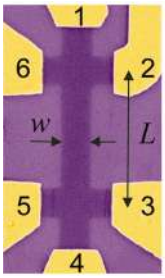


Fig. 2 Hall bar geometry for nonlocal transport measurements. Traditional four-terminal measurement involves passing a current between leads 1 and 4, while measuring the voltage drop between leads 2 and 3. The resulting resistance $R_{23,14} = V_{23}/I_{14}$ is related to the longitudinal resistivity, $\rho_{xx} = R_{23,14}W/L$, where W and L are the width and length of the Hall bar. In contrast, a nonlocal measurement consists of passing a current between, e.g., 2 and 6, while measuring the voltage between leads 3 and 5. In the case of usual diffusive transport such voltage should be exponentially suppressed [69], $R_{NL} = R_{35,26} \sim \rho_{xx} \exp(-\pi L/W)$ (From Ref. [70]. Reprinted with permission from AAAS).

the whole sample by passing a *dc* current. The resulting measurement exhibited a minimum in the differential resistance as a function of the current, see Fig. 1, which was argued to be equivalent to the Gurzhi effect. More recently, the observed decrease of resistivity with increasing temperature typical of the Gurzhi effect ($\rho \sim T^{-2}$) was reported in Ref. [32].

2.2 Nonlocal transport measurements

The "modern era" in electronic hydrodynamics was announced in the three back-to-back Science papers in 2016 reporting the negative vicinity resistance [27] and Wiedemann-Franz law violation [28] in graphene, as well as hints of the hydrodynamic behavior in [29] in $PdCoO_2$. These groundbreaking experiments opened the door for further studies focusing on unconventional aspects of electronic transport in ultra-pure materials.

Conventional experiments aimed at uncovering inner workings of solids often rely on transport measurements [10,16], the tool that proved to be indispensable throughout the history of condensed matter physics. In a traditional experiment one measures a current-voltage characteristic and extracts linear response functions determined by properties of the unperturbed system. A basic quantity that can be measured in this way is the Ohmic resistance R. At the simplest level, R can be described by the Drude theory [10,71], which essentially amounts to writing down classical equations of motion of charge carriers in applied electric and magnetic fields with a phenomenological friction term.

A more intricate question concerns the spatial distribution of the electric current density, which is most relevant in small samples (chips) with multiple leads. Here the current density may exhibit complex patterns depending on the external bias, electrostatic environment, chip geometry, and magnetic field. One way to detect such patterns is provided by nonlocal transport measurements [72,73,74,75,76,77,78], i.e., by measuring voltage drops between various leads that are spatially removed from the source and drain, see Fig. 2. These techniques were devised to study ballistic propagation of charge carriers in mesoscopic systems, but recently they were applied to investigate possible hydrodynamic behavior in ultra-pure conductors [23,24,27,41,42].

Nonlocal resistance measurements have also been used to study edge states accompanying the quantum Hall effect [70,79,80,81,82,83]. While the exact nature of the edge states has been a subject of debate, the nonlocal resistance, R_{NL}, appears to be an intuitively clear consequence of the fact that the electric current flows along the edges of the sample. Such a current would not be subject to exponential decay [69] exhibited by the bulk charge propagation leading to a much stronger nonlocal resistance.

2.2.1 Giant nonlocality in magnetic field

While traditional studies of electronic transport tended to focus on low temperatures, more recent experimental work has been gradually shifting towards measurements at nearly room temperatures [27,41,42,70,77]. A detailed analysis of the nonlocal resistance in a wide range of parameters (temperatures, carrier densities, and magnetic fields) was performed in Ref. [70] using graphene samples.

At low temperatures and in strong magnetic fields, graphene exhibits the quantum Hall effect (QHE) with well-defined plateaus in Hall resistivity corresponding to regions of the carrier density where $\rho_{xx} = 0$. At the same densities, the nonlocal resistance also remains zero, but in between the QHE zeros may reach values as high as $1\,\mathrm{k}\Omega$. At high temperatures, all but one such peaks disappear. The remaining peak at charge neutrality exhibits behavior that appears to be inconsistent with the QHE interpretation. In particular, the strong signal persists at near room temperatures, way beyond the QHE regime with the peak value $R_{NL} \approx 1.5\,\mathrm{k}\Omega$ at $B = 12\,\mathrm{T}$ and $T = 300\,\mathrm{K}$, three times higher than that at $T = 10\,\mathrm{K}$, see Fig. 3.

The unexpected "giant" nonlocality in neutral graphene was originally explained by diffusion of the mismatched spin-up and spin-down quasiparticles in the presence of the Zeeman splitting [70]. This interpretation was disputed in Ref. [84] where the effect was not observed in the nearly parallel field (the Zeeman splitting is independent of the field direction). Moreover, the magnitude of the effect proposed in Ref. [70] was disputed in Ref. [85], where the residual quasiparticle density due to Zeeman splitting (at $T = 0$ and $B = 10\,\mathrm{T}$) was estimated to be $\rho_Q \approx 2.2 \times 10^6\,\mathrm{cm}^{-2}$ leading to a nonlocal resistance that is much weaker than the data of Ref. [70].

The alternative explanation suggested in Ref. [85] was based on the "two band" phenomenology of the electronic system in neutral graphene [86,87,88]. Indeed, at the charge neutrality point, the conductance and valence bands in graphene touch. At finite temperatures, both bands contain mobile carriers leading to a two-component nature of the electronic system. Given the exact particle-hole symmetry at neutrality, this system is "compensated" and hence

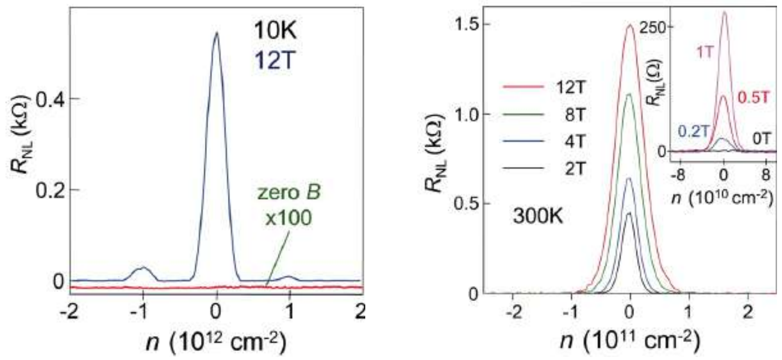


Fig. 3 Nonlocal resistance in graphene. Left panel: QHE regime at $T = 10\,\mathrm{K}$ and $B = 12\,\mathrm{T}$ (the red curve indicates that no signal could be detected at $B = 0$ within the experimental resolution; the curve is downshifted for clarity and magnified). Right panel: high-temperature regime, $T = 300\,\mathrm{K}$. (From Ref. [70]. Reprinted with permission from AAAS).

there is no classical Hall effect, such that the bulk Hall conductivity vanishes, $\rho_{xy} = 0$, and the longitudinal conductivity is unaffected by the magnetic field. In contrast, the same approach yields the nonlocal response that is strongly field dependent. Indeed, the presence of two types of carriers (electrons and holes) leads to the existence of two macroscopic currents: the electric current $\boldsymbol{J}$ and the total quasiparticle (or "imbalance" [89]) current $\boldsymbol{j}_I$,

$$\boldsymbol{j} = \boldsymbol{j}_e - \boldsymbol{j}_h, \qquad \boldsymbol{j}_I = \boldsymbol{j}_e + \boldsymbol{j}_h, \qquad \boldsymbol{J} = e\boldsymbol{j}, \tag{1}$$

where $\boldsymbol{j}_{e(h)}$ is the electron (hole) current and e is the electron charge. In the absence of the magnetic field, the neutral current $\boldsymbol{j}_I$ is decoupled from $\boldsymbol{J}$ and is practically undetectable (it does not couple to the electric field). The electrons and holes are drifting in parallel, but opposite directions. However, the magnetic field bends the quasiclassical trajectories of charge carriers coupling the two currents and turning $\boldsymbol{j}_I$ in the direction that is orthogonal to $\boldsymbol{J}$. Now the neutral current can transport charge carriers to distant parts of the sample, where a nonlocal response is induced, again, by the magnetic field, see Section 4 for more details.

The arguments of Ref. [85] yield the nonlocal response capturing the main qualitative features of the effect observed in Ref. [70]. Quantitatively, these results are consistent with the rapid decay of the nonlocal signal away from the neutrality point, but overestimate the magnitude of the effect. The latter discrepancy was attributed to the simplicity of the model that did not take into account the effects of electron-electron interaction contributing to resistivity of neutral graphene, the residual carrier population at neutrality due to fluctuations of the electrostatic potential [84], and viscous phenomena, all of which are expected to suppress R_{NL}.

Viscous effects are of particular interest in the context of electronic hydrodynamics and may also lead to nonlocality. However, these effects are expected to occur in the absence of magnetic field as well and in graphene are most pronounced away from charge neutrality.

2.2.2 Negative vicinity resistance

Away from charge neutrality, i.e., when the chemical potential exceeds the temperature, $\mu \gg T$, electrons in graphene are typically expected to behave similarly to 2DEG in semiconductor heterostructures. The contribution of the valence band is exponentially suppressed and the electronic system comprises only the single component. In that case, a Fermi liquid is expected to behave hydrodynamically [26], the issue with the electronic systems being whether the material is pure enough.

Assuming the hydrodynamic regime is possible, the single-component electronic system should obey the Navier-Stokes-like equation [13,90,91,92] with an additional damping due to disorder scattering [21], as well as the continuity equation. Within linear response and in the static limit, these equations can be written as (see, e.g., Section 3)

$$e\boldsymbol{E} = -m\nu\Delta\boldsymbol{u} + m\boldsymbol{u}/\tau_{\text{dis}}, \qquad \boldsymbol{\nabla u} = 0, \tag{2}$$

where $\boldsymbol{u}$ is the hydrodynamic velocity, ν is the kinematic viscosity, and m is the effective mass (in graphene this should be replaced by μ/v_g^2, with μ being the chemical potential and v_g the velocity of the Dirac spectrum). The electric current is expressed in terms of the hydrodynamic velocity as

$$\boldsymbol{j} = n\boldsymbol{u}, \tag{3}$$

where n is the carrier density, see also Eq. (1).

The resulting behavior of the current density is determined by the relative strength of the viscosity and disorder scattering, which can be expressed in terms of the dimensionless "Gurzhi number" (note that this definition is written in analogy to the Reynolds number [13] and is the inverse of the number defined in Ref. [93])

$$\text{Gu} = \frac{l^2}{\nu\tau_{\text{dis}}}, \tag{4}$$

where l is the typical length scale of the problem. Large values of Gu indicate that the disorder scattering dominates (such that the current density exhibits patterns typical to the traditional diffusive behavior), whereas small values of Gu correspond to the hydrodynamic viscous flow [94,95,96,93,97,98].

In confined geometries, viscous flows may be accompanied by vortices (or whirlpools) [94,95,96,93], which may be detected by observing *negative* nonlocal resistance by placing the leads on the opposite sides of a vortex. This idea was realized in the pioneering experiment of Ref. [27]. Here (unlike the measurement in Ref. [70]) the leads were placed close to each other (based on the expected vortex size), see Fig. 4, hence the measured quantity was referred to as "vicinity resistance".

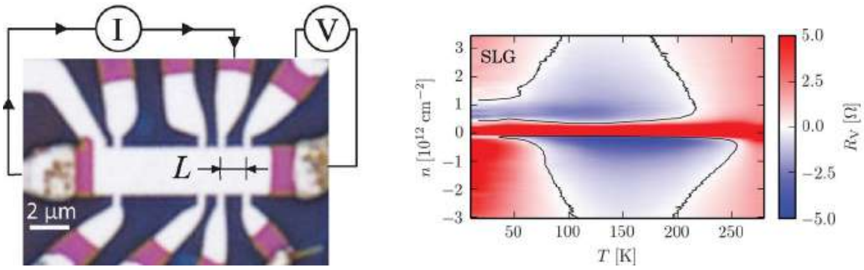


Fig. 4 Negative vicinity resistance in graphene. Left panel: multi-lead device with the measurement schematic. Right panel: color map showing a wide, intermediate temperature range where the vicinity resistance is negative (From Ref. [27]. Reprinted with permission from AAAS).

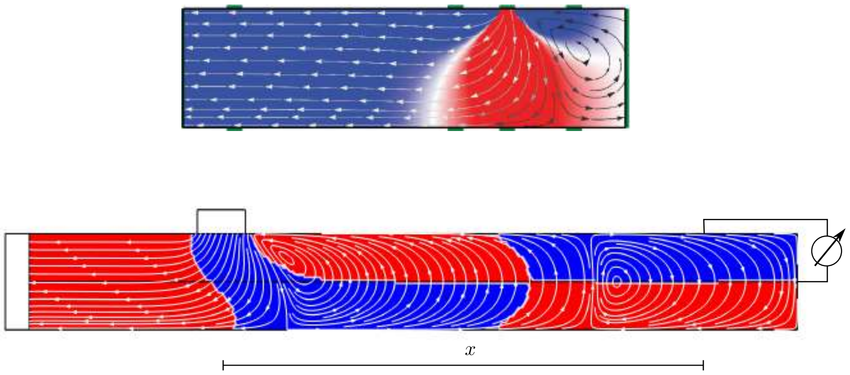


Fig. 5 Vorticity in electronic flows in graphene. Top panel: simulated flow in the experimental device shown in Fig. 4 (from Ref. [27]. Reprinted with permission from AAAS). Bottom panel: double vortex in a long device suggested in Ref. [93]. The red and blue colors indicate the alternating sign of the deviation of the electrochemical potential from its median value. (Reprinted with permission from IOP Publishing).

In agreement with the expectation that the hydrodynamic behavior should occur at intermediate temperatures, the measured vicinity resistance is negative roughly between 70 K and 250 K (with the actual range being density dependent), see Fig. 4. This observation was supported in Ref. [27] by a solution to the above equations (2) showing formation of a vortex close to the leads. Similar theoretical results were reported in Refs. [94,95,96,93], see also Ref. [99].

Despite the apparent agreement between theory and experiment, observation of the negative vicinity resistance does not represent the proverbial “smoking gun” proving that the system is in fact in the hydrodynamic regime. The reason is that ballistic systems may also exhibit negative nonlocal resistance [51] as has been shown both experimentally [77] and theoretically [100]. This issue has been specifically studied in Ref. [41], where it was shown that in addition to being negative, the vicinity resistance has to grow with temperature (the crossover from the ballistic to hydrodynamic behavior was identified with the minimum in the vicinity resistance as a function of temperature). More recently, Ref. [93] reported a numerical solution to the hydrodynamic equations (2) showing the existence of multiple vortices in long samples, see Fig. 5. Since the vorticity of the adjacent vortices has the opposite sign, placing multiple leads along the sample and measuring the voltage as a function of distance from the source electrode should yield a *sign-alternating nonlocal resistance* which should in principle distinguish the ballistic and hydrodynamic behavior. Alternatively, one could try to use one of the novel imaging techniques [50, 51, 52, 53] to observe vortices "directly".

2.3 Hydrodynamic flow around macroscopic obstacles

The collective hydrodynamic flow is expected to differ strongly from the single-particle ballistic motion in systems with macroscopic obstacles. Whereas particles tend to scatter off anything they may encounter – sample boundaries, other geometrical features, or long-range potentials, a viscous fluid tends to avoid obstacles by flowing around them. As a result, the collective flow maybe more efficient in carrying the constituent particles through the system in question. In the context of the traditional hydrodynamics of rarefied gases, this fact has been established already by Knudsen [68]. In the context of electronic hydrodynamics, this issue was first addressed theoretically in Ref. [101] and experimentally in Ref. [39].

2.3.1 Superballistic transport

One of the most common types of "obstacles" studied in the context of electronic transport is a constriction (or a point contact). This object was extensively studied in mesoscopic physics [11], with the conductance quantization [102, 103] being the hallmark effect. In particular, it was established that ballistic propagation of charge carriers through a point contact yields the conductance that is constrained by a fundamental upper bound [104].

Quantization of the point contact conductance can be understood by considering one-dimensional (1D) subbands in the constriction of the width W (corresponding to the quantized values of the transverse momentum, $k_y = \pm\pi n/W$). Each subband contributes equally to the conductance due to the cancellation of the group velocity and the 1D density of states (DoS) [11]. Observing that the number of the occupied subbands is naturally an integer, one finds that the total conductance is quantized, $G_b = 2Ne^2/h$. In the classical limit, the number of propagating (Landauer) channels in 2D can be estimated as $N = [k_F W/\pi]$ (square brackets indicate the integer value), yielding the upper bound known as the Sharvin limit [11, 104].

The above argument neglects electron-electron interaction and is justified when the corresponding scattering length is large compared to the width of the constriction, $\ell_{ee} \gg W$. In the hydrodynamic regime, $\ell_{ee} \ll W$, electrons move collectively avoiding the boundaries and thus may carry the charge through the

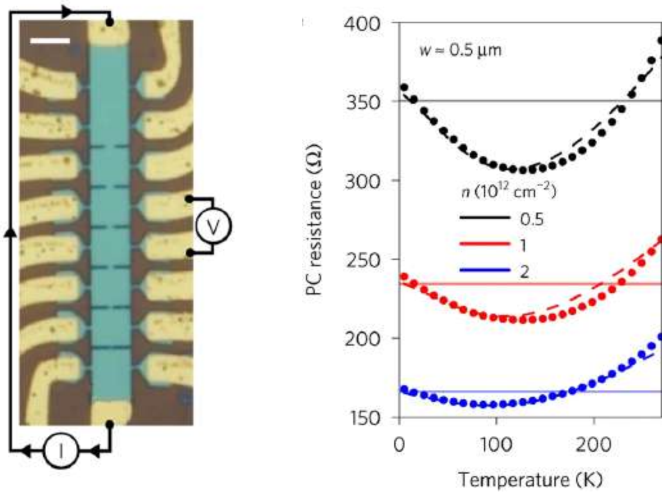


Fig. 6 Superballistic transport in graphene. Left panel: a typical measuring device showing multiple point contacts varying in width from $W = 0.1$ to $W = 1.2\,\mu$m. Right panel: point contact resistance for a $W = 0.5\,\mu$m constriction at representative carrier densities. The experimental data are represented by dots, while the horizontal lines indicate the Sharvin limit of the maximum classical ballistic conductance. Lower-than-the-limit resistance at intermediate temperatures is indicative of the collective, viscous flow of electrons. (From Ref. [39]. Reprinted with permission from Springer Nature).

point contact more effectively than free fermions (i.e., achieving conductance higher than G_b, see Fig. 6). Indeed, the solution to the hydrodynamic equations describing the electron flow through a simplest 2D constriction reported in Ref. [101] yields the conductance

$$G_h = \frac{\pi e^2 n^2 W^2}{32\eta}, \tag{5}$$

where η is the shear viscosity. Since G_h grows with width faster than G_b, there is a possibility for the “superballistic” conduction for wide enough channels.

The theoretical expectation ($G_h > G_b$) was first confirmed in the experiment of Ref. [39], see Fig. 6, and more recently corroborated in Ref. [44], where a novel imaging technique was applied to the point contact problem (see Section 2.4), see also Ref. [45]. The theory of Ref. [101] was revisited and expanded upon in Ref. [37], where the same hydrodynamic equation was solved for the current density profile. The authors of Ref. [37] also analyzed the intermediate parameter regime where hydrodynamic flows could be realistically observed. Heating effects in similar inhomogeneous flows ware analyzed in Ref. [105].

2.3.2 Flows around macroscopic obstacles

The transition from the Ohmic to hydrodynamic flow observed in the point contact geometry in Refs. [39, 44] is similar to the transition between the Knudsen and Poiseuille flows [17, 68, 58]. The tendency of the viscous flow to avoid

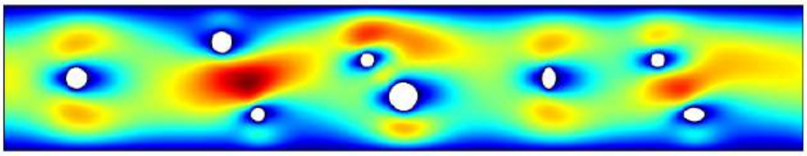


Fig. 7 Numerical simulation of the Poiseuille flow in a 2D channel with randomly placed macroscopic obstacles (represented by white shapes). The color map indicates the magnitude of the flow velocity (ranging from zero shown in blue to the maximum shown in dark red).

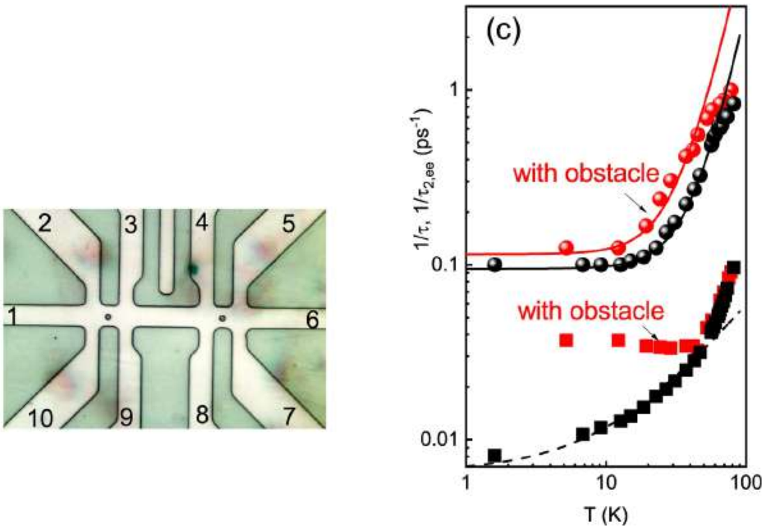


Fig. 8 Stokes flow around an obstacle in GaAs. Left panel: image of the Hall bar with two anti-dots used in the experiment [32]. Right panel: the disorder (squares) and electron-electron interaction (circles) scattering rates obtained from the experimental data measured in sample with (red) and without (black) the obstacle (From Ref. [32]).

obstacles is well known in hydrodynamics and is illustrated in Fig. 7. However, a naive solution of the hydrodynamic equations in a 2D system with macroscopic obstacles within linear response leads to the so-called "Stokes paradox" [13, 14, 32, 106, 107].

The problem of a motion of a spherical object through an otherwise stationary viscous fluid (or equivalently, viscous flow around a stationary sphere subject to the condition of constant flow velocity at infinity) is a classic problem in hydrodynamics [13, 14, 92]. For flows characterized by small Reynolds numbers, one may neglect the nonlinear term in the Navier-Stokes equation [13, 90, 91] and solve the resulting system of linear equations. In 3D the problem can be solved analytically not only for the sphere but also for several other simple shapes [14], where one typically calculates the "drag force" acting on the obstacle.

The above simple solution of the linearized hydrodynamic equations appears to fail if the obstacle has the form of an infinitely long cylinder (or equivalently, in 2D), the issue known as the "Stokes paradox". The reason for the apparent paradox lies in the approximation used to linearize the Navier-Stokes equation: the Reynolds number (i.e., the quantitative expression for the relative strength of the nonlinear and viscous terms) is scale-dependent and cannot be assumed small at arbitrary large distances [13,14]. Instead of simply neglecting the nonlinear term, one should linearize it following Oseen [108], whose modified equation yields a consistent solution (as well as the corrected expression for the drag force).

In contrast to traditional hydrodynamics, in solid state physics one is typically interested in linear response properties and has to take into account momentum relaxation due to weak impurity scattering. The latter allows one to stabilize the solution, while keeping it within linear response [106,107]. Indeed, in ultra-pure electronic systems the Gurzhi number (4) may be much larger than the Reynolds number

$$\frac{\mathrm{Gu}}{\mathrm{Re}} = \frac{l^2/(\nu\tau_{\mathrm{dis}})}{ul/\nu} = \frac{l}{u\tau_{\mathrm{dis}}}, \tag{6}$$

justifying the Stokes approximation in the presence of the impurity scattering.

Stokes flow in the 2D electron system with a circular obstacle was observed in Ref. [32]. The experiment was performed in a GaAs heterostructure with the role of the obstacle played by an anti-dot (or a micro-hole) in the middle of the Hall bar, see Fig. 8. The measured resistivity was interpreted using the macroscopic approach of Refs. [109,110]. The two scattering mechanisms (one due to impurity scattering and another due to viscosity) were treated as two parallel channels of momentum relaxation (based on the fact that the corresponding relaxation rates can be attributed to the first and second moments of the semiclassical distribution function). The two contributions can be separated since they have a different temperature dependence, in particular, the viscous contribution should exhibit the Gurzhi-like $\rho \sim T^{-2}$ behavior. Now, the obstacle does not seem to affect the latter, while the disorder contribution at low temperatures is significantly enhanced, see Fig. 8, which is consistent with the expectation of the viscous fluid avoiding the obstacle (as opposed to individual electrons scattering off it).

2.4 Imaging of electronic flows

Although traditional linear response measurements may be strongly affected by the collective, hydrodynamic behavior, interpretation of such experiments is not straightforward [41]. It would be much easier if one could simply "watch" the flow (in a close analogy to the usual hydrodynamics). Fortunately, in recent years several "scanning" or "imaging" techniques were suggested allowing one to do just that even if indirectly.

The basic requirement for any imaging technique is that it should be non-invasive, i.e., it should not disrupt the flow itself. When trying to image the flow of electrons, one can rely either on detecting spatial variation of electric potential [51,52] or on detecting the local magnetic field induced by the charge motion [50].

2.4.1 Scanning carbon nanotube single-electron transistor

Electric current flowing through a conductor is known to generate a local change in electrostatic potential (or "voltage drop"). This potential can be detected using the capacitive coupling to a local probe such as the scanning single-electron transistor (SET), see Fig. 9. In particular, a nanotube SET may exhibit extreme voltage sensitivity, while the planar probe design could help minimizing the back action on (or gating) the sample [52]. Moreover, by applying weak perpendicular magnetic field, the same probe is able to resolve the Hall voltage associated with the flow, yielding a direct measure of the local current density.

Applying the nanotube SET technique to doped graphene in the hydrodynamic regime allowed to image the Poiseuille flow of charge carriers [51]. Similarly to the case of the Gurzhi effect, see Sec. 2.1, the main goal of the experiment was to distinguish the collective (hydrodynamic) motion from the single-particle (ballistic) behavior (assuming $\ell_{\rm dis}$ is the largest length scale in the problem). However, instead of contrasting the temperature dependence of the sample's resistance [41], here one has to compare the spatial profile of the current density. In the channel geometry, see Fig. 9, one studies its dependence on the lateral coordinate, $\boldsymbol{j} = j(y)\boldsymbol{e}_x$ (where $\boldsymbol{e}_x$ is the unit vector directed along the channel and y is the coordinate across the channel). The difficulty is that in contrast to the textbook diffusive behavior, where the current density is uniform (except in the narrow regions close to the sample boundaries), both in the ballistic and hydrodynamic cases $j(y)$ is characterized by a non-uniform profile with the maximum at the center of the channel [11,112], making it difficult to distinguish the two regimes experimentally.

The hydrodynamic Poiseuille flow in a narrow channel is a textbook problem [13]. Taking into account weak impurity scattering and making the common assumption of the no-slip boundary conditions, one finds for the electric current density in doped graphene in the channel geometry

$$J_x = \sigma E_x \left[1 - \frac{\cosh(y/\ell_G)}{\cosh[W/(2\ell_G)]} \right], \tag{7}$$

where σ in the bulk longitudinal conductivity and ℓ_G is the Gurzhi length [93, 110,113,114,115]

$$\ell_G = \sqrt{\nu \tau_{\rm dis}}. \tag{8}$$

Here ν is the kinematic viscosity, see Eq. (2). The parabolic current density profile typical of the standard Poiseuille flow [13,116] can be recovered by assuming a large Gurzhi length, $\ell_G \gg W$. In this limit, the sample resistance is proportional to the shear viscosity [114], a manifestation of the Gurzhi effect [17].

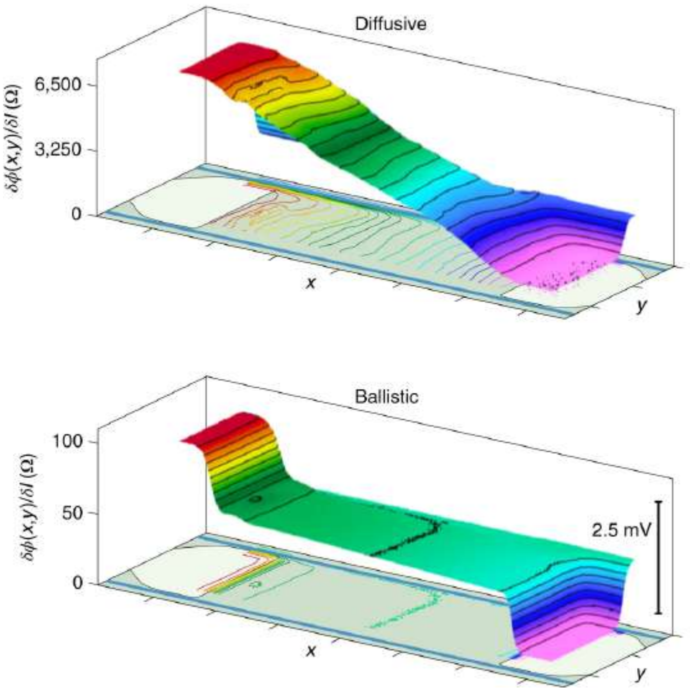


Fig. 9 Spatial imaging of the voltage drop of flowing electrons in the diffusive (top) and ballistic (bottom) regimes [52]. Both plots show the imaged electrostatic potential normalized by the total current (yielding a quantity with the units of resistance). The data were measured at $T = 4\,\mathrm{K}$. The diffusive flow was observed at charge neutrality (determined by the sharp maximum in the two-terminal resistance of the sample), while the ballistic behavior was imaged at the hole density of $1 \times 10^{12}\,\mathrm{cm}^{-2}$. In the latter case, most of the voltage drop occurs at the contacts, with the contact resistance approaching the ideal Sharvin value [104, 111]. The bottom plane shows the equipotential contours superimposed on the schematic of the graphene channel and contacts, indicating the gradual voltage drop in the diffusive case contrasted to the flat potential typical of the ballistic motion [112] (From Ref. [52]. Reprinted with permission from Springer Nature).

Introducing more realistic (Maxwell's) boundary conditions with nonzero slip length [117] effectively sets the coordinates where the catenary curve (7) reaches zero outside of the channel, but does not significantly affect the current density in the bulk of the sample. From the experimental viewpoint, however, the resulting curve is difficult to distinguish from the non-uniform current density in the ballistic regime, see the bottom panel in Fig. 10 and Sec. 3.2. As a result, one has to perform other measurements (e.g., the Hall field, see Fig. 10) to distinguish the two regimes [51].

2.4.2 Quantum spin magnetometry

An alternative technique for imaging the electric current density is based on the idea of measuring the associated stray magnetic field [50]. A sensitive

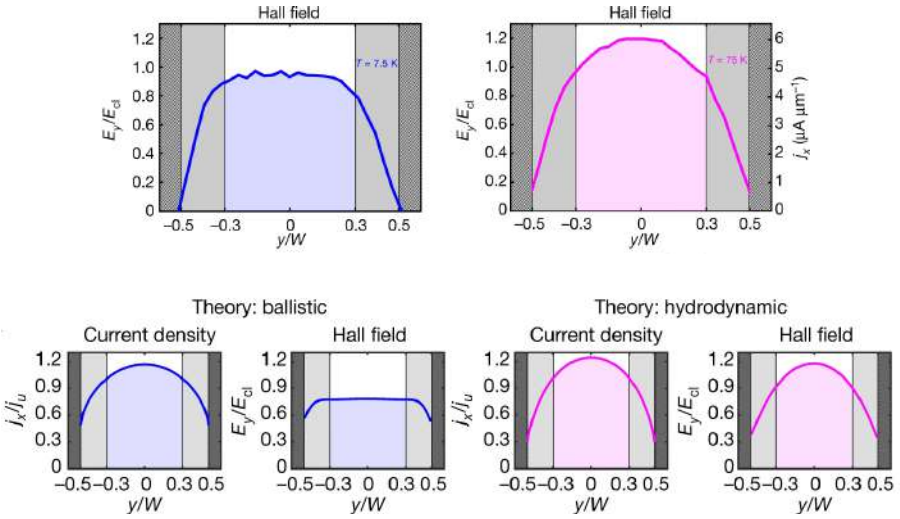


Fig. 10 Spatial imaging of the hydrodynamic flow of electrons in doped graphene [51]. Top: the Hall field E_y as obtained by numerical differentiation of the measured Hall voltage with respect to y, normalized by the classical value, $E_{\rm cl} = BJ/(neW)$. The top left panel shows data taken at $T = 7.5\,$K, $B = \pm 12.5\,$mT, and $E_{\rm cl} = 91\,{\rm Vm}^{-1}$. The right top panel shows data in the presumed hydrodynamic regime at $T = 75\,$K, $B = \pm 18\,$mT, and $E_{\rm cl} = 162\,{\rm Vm}^{-1}$. The right vertical axis converts the field into the units of the current density by scaling with ne/B. Bottom: calculated current density J_x/J_u (with $J_u = J/W$) and Hall field $E_y/E_{\rm cl}$. The numerical values were obtained using the parameters corresponding to the experimental data in the top panels. (From Ref. [51]. Reprinted with permission from Springer Nature).

quantum spin magnetometer was realized using nitrogen vacancy (NV) centers in diamonds [118]. In contrast to Ref. [51], the experiment of Ref. [50] targeted the so-called Dirac fluid in neutral graphene and contrasted the presumed hydrodynamic regime with the diffusive behavior in low-mobility devices. The latter measurements served as a benchmark and yielded the standard picture of nearly uniform current (exhibiting a sharp decay near the channel boundaries, see also Sec. 3.2) shown in Fig. 11.

The main result of Ref. [50] is the observation (by means of the scanning NV magnetometry) of a Poiseuille-like flow of the electric current in neutral graphene described by a catenary curve (7). Comparing the data to Eq. (7) the authors have extracted the kinematic viscosity of the Dirac fluid in graphene (see the right panel in Fig. 11) showing a good quantitative agreement with the theoretical calculations of Ref. [119] (without any fitting procedure). Nevertheless, the results of Ref. [50] remain controversial. Within the existing theory of electronic hydrodynamics, the electric current is related to the hydrodynamic velocity by Eq. (3) up to an Ohmic correction. Precisely at charge neutrality, $n = 0$, and Eq. (3) yields zero, implying that any electric current at charge neutrality is not hydrodynamic, but is rather given by the Ohmic correction [120, 121] with the corresponding bulk conductivity determined by electron-electron interaction [122]. The situation is more involved if the system

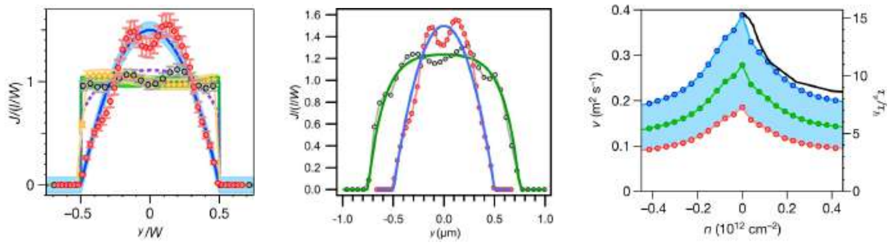


Fig. 11 Spatial imaging of the electric current in neutral graphene [50]. Left: reconstructed current density as a function of the lateral coordinate. The current is normalized by the average charge carrier flux I/W, where I is the total flux and $W = 1\,\mu$m is the width of the channel. The spatial coordinate y is normalized by W and centered on the channel. Red points show data measured in neutral graphene, gray points – in palladium channel, orange points – low-mobility graphene. The curves correspond to idealized theoretical expectations: blue – ideal viscous flow, green – uniform current, purple dashed – the current profile of non-interacting electrons with diffusive boundary condition. Center: similar measurement for $W = 1.5\,\mu$m compared to the data on the left. Solid lines are fit to Eq. (7). Right: bounds on kinematic viscosity obtained from fitting the data to Eq. (7). The black curve is the result of a theoretical calculation of Ref. [119] at $T = 300\,$K and no adjustable parameters (From Ref. [50]. Reprinted with permission from Springer Nature).

is subjected to the external magnetic field. In that case, the Ohmic correction acquires an additional dependence on the hydrodynamic velocity [120], which in particular leads to positive magnetoresistance [123,124]. However, a recent theoretical calculation of the electronic flow in a channel geometry in neutral graphene based on the direct solution of hydrodynamic equations (see Sec. 4) yields the so-called “anti-Poiseuille” flow [125], with the current density exhibiting a minimum in the center of the channel – in contrast to the maximum in Eq. (7), see Sec. 5.

Another feature of the data shown in Fig. 11 not accounted for by the existing theory is that the electric current vanishes at the channel boundaries. Indeed, the boundary conditions for the Ohmic correction to Eq. (3) should be derived from the kinetic theory similarly to those describing ballistic propagation of electrons [11]. In that case, one has to solve the kinetic equation imposing boundary conditions on the electronic distribution function. Both extreme limits typically considered in literature, namely the diffusive and specular boundary conditions, do not lead to the current vanishing at the boundary. Moreover, the kinetic theory derivation of the hydrodynamic equations yields the Maxwell’s boundary conditions for the hydrodynamic velocity [117]. Finally, there is strong experimental evidence [53] for the existence of classical edge currents in graphene that are not taken into account in existing theories but casting further doubts on the results shown in Fig. 11.

The importance of edge physics is further highlighted by the experiment of Ref. [44], where the NV magnetometry was used to image the flow of charge through a constriction (or a slit) in neutral graphene. The authors performed measurements in a channel geometry as well with somewhat contradicting results, see Fig. 12. While the channel measurement at nearly room temperature ($T = 298\,$K) yielded the current density profile similar to that reported in

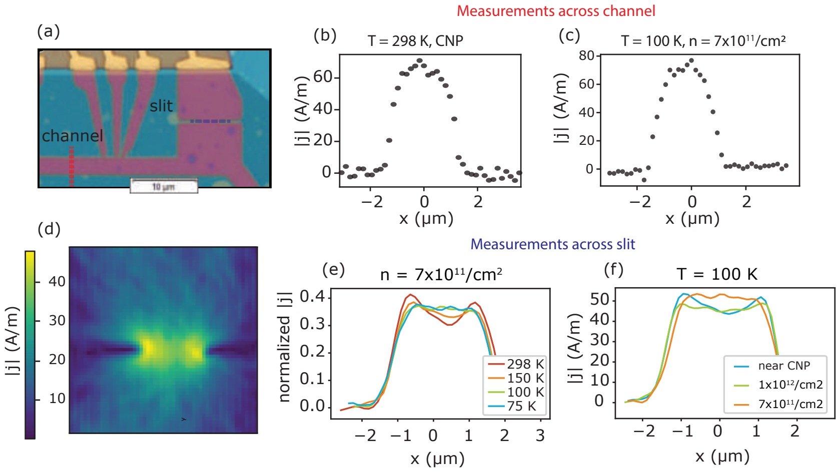


Fig. 12 Spatial imaging of the electric current in neutral graphene [44]. (a) Optical image of the graphene device showing the locations used in obtaining the current density measurements for the channel geometry (b-c), $W = 2.7\,\mu$m, and the slit geometry (d-f). (b) Current density profile in the channel near the charge neutrality point (CNP) at $T = 298\,$K. The black dots are the reconstructed current density. (c) Measurement of the current density profile of the channel at the same position as in (b), but at $T = 100\,$K and $n = 7 \times 10^{11}\,\mathrm{cm}^{-2}$. (d) Reconstructed current density magnitude at $T = 298\,$K, near the CNP, showing the characteristic double peaks of Ohmic flow. (d) Temperature dependence of the reconstructed j_y at fixed carrier density $n = 7 \times 10^{11}\,\mathrm{cm}^{-2}$ in a line cut through the constriction. (e) Carrier density dependence of j_y at fixed temperature $T = 100\,$K. (From Ref. [44]. Reprinted with permission from the authors).

Ref. [50], see Fig. 11, the same profile was observed at $T = 100\,$K implying that the charge flow in the channel is not very sensitive to the variation of the scattering length. In contrast, the current density measured in the slit geometry exhibited Ohmic behavior at room temperature, while at lower temperatures and finite charge densities the Ohmic double peaks disappeared indicating the crossover into the hydrodynamic regime. The authors of Ref. [44] explained the contradiction between the results in the channel and slit geometries by fact that the latter is not affected by the boundary conditions as much as the former. They conclude that while the edge physics is poorly understood the slit geometry is better suited to observe the Ohmic-viscous crossover.

2.4.3 Non-topological edge currents

Sample edges play a crucial role in all of the experiments discussed so far. Yet, understanding of the physics of the edges themselves has proven somewhat challenging. In traditional condensed matter physics [10], the focus is typically on bulk behavior and hence a system is modeled to be infinite. Sample geometry and edge scattering becomes important in mesoscopic physics [11, 112], but most details are encoded in the boundary conditions. Finally, edge states are being actively researched in the context of the Quantum Hall Effect

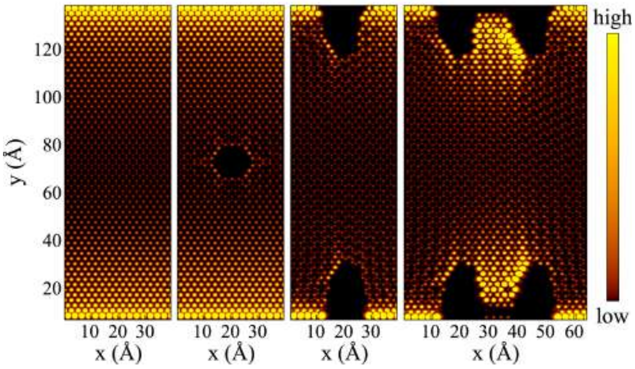


Fig. 13 DFT calculation of the local density of states (LDoS) in a graphene flake [131]. The enhanced LDoS at the edges appears regardless of the shape of the edge and the presence of macroscopic defects in the bulk (From Ref. [131]).

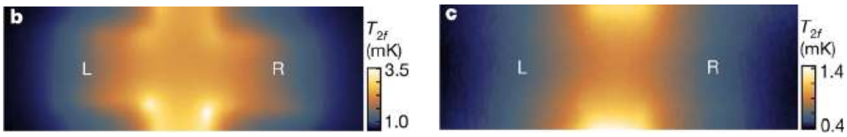


Fig. 14 Thermal imaging of a graphene sample [53]. Both images show the local temperature distribution obtained using the scanning SOT at the background temperature $T = 4.2\,\mathrm{K}$ and $B = 0$. Left: enhanced nonlocality in neutral graphene – heat dissipation is extended into the left and right arms of the Hall bar. Right: Ohmic behavior – heat dissipation is confined to the central region of the sample between the source and drain electrodes. (From Ref. [53]. Reprinted with permission from Springer Nature).

(QHE) [126,127] and more generally in the field of topological insulators [128]. But even in the latter case, the edge behavior is dictated by the topological properties of the bulk. At the same time, experiments show that sample edges (in particular, in graphene, see Fig. 13) may exhibit charge accumulation [48, 129,130,131] and carry non-topological currents [53,132].

Charge accumulation at the surface is a known phenomenon in semiconductors [133] and is a key feature in the traditional theory of the Schottky barrier [134]. Typically, these effects are associated with "band bending" or local, position-dependent changes in quasiparticle energy levels in the vicinity of the sample surface (or an interface). The band bending can also occur in 2D systems. In particular, it has been suggested that in graphene band bending leads to p-doping of the edges, due to either intrinsic mechanisms or charged impurities (or defects) [48,132,135]. The resulting hole accumulation at the sample edges has been used in Ref. [53] to interpret the highly unusual nonlocal transport observed by means of SQUID-on-tip (SOT) thermal imaging and scanning gate microscopy [47,48,136] (for applications of scanning gate microscopy to 2D electron systems in semiconductor heterostructures see Ref. [137]).

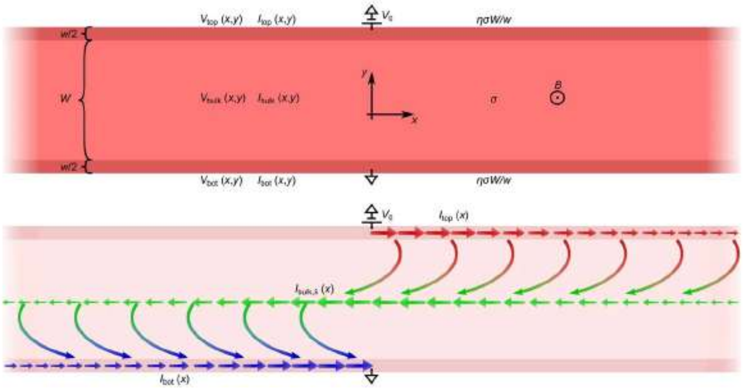


Fig. 15 Classical model mimicking the effects of charge accumulation at the sample edges [53]. Top: the setup – a strip-like sample of width W with bulk conductivity σ and narrow edge regions (width $w/2$) with the conductivity $\eta\sigma W/w$ with η being the phenomenological measure of charge accumulation. Bottom: nonuniform current density in the presence of the magnetic field B featuring the bulk flow in the direction opposite to the applied electric field. (From Ref. [53]. Reprinted with permission from Springer Nature).

The experiment of Ref. [53] provided a deeper insight into the giant nonlocality observed in neutral graphene subjected to magnetic field in Ref. [70], see Sec. 2.2.1. While confirming the giant enhancement of the nonlocal resistance at charge neutrality and in magnetic field, the new data show a number of novel features: (i) the nonlocality exists even in the absence of magnetic field; although the observed R_{NL} is much smaller than in the presence of the field, it is still an order of magnitude stronger that the Ohmic expectation; (ii) the observed nonlocality is asymmetric with respect to electron and hole doping; (iii) in magnetic field, the system exhibits the Hall voltage of the opposite sign (as compared to the naive expectation); and most importantly, (iv) the observed nonlocality can be suppressed by applying a potential at the sample edges. The latter observation represents the key evidence in support of the interpretation of the data offered in Ref. [53]. The authors argue that the sample edges may carry electric current which in turn leads to nonlocal resistance. The fact that this current can be suppressed by a local potential points towards its non-topological origin (a topological current tends to flow around obstacles [48] such that applying a potential would just "redefine" the edge). The existence of the edge current is further corroborated by the thermal imaging, see Fig. 14.

The authors of Ref. [53] offer a simple theoretical model to account for the experimental data. Consider a sample that is infinite in x direction, while having a width W in the y direction. Without charge accumulation at the edges, the sample can be assumed to host a uniform charge density, while the current density can be found using the Ohm's law and the continuity equation. Consider now a different situation, where the charge density in narrow regions close the sample edge exceeds the bulk density. Now, the same equations have

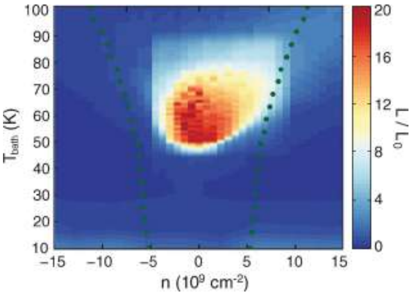


Fig. 16 Wiedemann-Franz law violation in neutral graphene [28]. The color scheme shows the Lorenz number as a function of the charge density and bath temperature. The unusually large Lorenz number is observed in the vicinity of charge neutrality and in a temperature window above the disorder-dominated regime, but below the onset of electron-phonon coupling (From Ref. [28]. Reprinted with permission from AAAS).

to be solved separately in the edge and bulk regions leading to the complicated behavior shown in Fig. 15.

The classical model accounts for the unexpected inversion of the Hall voltage and edge currents observed in the experiment, but does not explain the physical origin of these effects at a microscopic level. Some of these features appear to be rather general for the usual transport equations in the strip geometry. For example, current flows against the direction of the applied electric fields have also been reported in Ref. [115], where the hydrodynamics-like phenomenology was used to define distinct edge regions where charge carriers react to the applied magnetic field differently than carriers in the bulk of the sample, see also Ref. [87].

Although implications of the results of Ref. [53] are not fully understood at the time of writing, it is clear that the boundary effects play a very important role in the observed behavior of small graphene samples. This presents a clear challenge for the theory which so far was focusing on bulk systems, see Sec. 3. In particular, the existing solutions of the hydrodynamic equations in the strip geometry (similar to Fig. 15) were found under the simplest model assumptions of either the no-slip or Maxwell's boundary conditions, see Sec. 3.

One could try to avoid the issue of the boundary conditions (except for the boundaries with the source and drain electrodes [138]) by utilizing the Corbino disk geometry [139]. Due to inherently inhomogeneous current flow (even in the Ohmic regime), the Corbino disk was suggested as a potential device to measure electronic viscosity [140]. More recently, hydrodynamic behavior in this setting was reported in the imaging experiment of Ref. [54].

2.5 Wiedemann-Franz law violation

Unconventional charge transport properties exhibited by electronic systems presumed to be in the hydrodynamic regime may be accompanied by unusual heat transport leading to strong violations of the Wiedemann-Franz law

[141,142]. Initially an empirical observation, the Wiedemann-Franz law can be readily understood within the standard, single-particle transport theory [10]. Qualitatively, if both charge and heat are carried by the same excitations and affected by the same scattering mechanisms (as is the case for noninteracting electron models), then the only difference between the electric and thermal conductivities is the dimensionality, leading to the famous expression

$$\frac{\kappa}{\sigma} = \mathcal{L}T, \qquad \mathcal{L} = \mathcal{L}_0 = \frac{\pi^2}{3e^2}. \tag{9}$$

Here σ and κ are the electric and thermal conductivities and the coefficient $\mathcal{L}$ is known as the Lorenz number, while the "universal" value $\mathcal{L}_0$ corresponds to free electrons. Now, electrons in solids are typically not free and hence there is no reason for Eq. (9) to be universally valid. In conventional metals, the Wiedemann-Franz law is approximately obeyed, for example the Lorenz number in copper exhibits deviations from $\mathcal{L}_0$ up to a factor of 2 at intermediate temperatures (depending on sample purity) [143]. Consequently, a strong violation of the Wiedemann-Franz law almost certainly an indication of unconventional physics, that in the context of electronic systems may include hydrodynamic behavior.

2.5.1 Large Lorenz number in neutral graphene

Unconventional thermal transport in neutral graphene was reported already in early experiments of Refs. [144,145]. The Wiedemann-Franz law was then studied in detail in Ref. [28] where it was interpreted as evidence for the hydrodynamic "Dirac fluid". An observation of the related phenomenon of giant thermal diffusivity in a Dirac fluid was reported in Ref. [146].

In hindsight, strong violation of the Wiedemann-Franz law in graphene should have been expected on the basis of the two celebrated features (see also Ref. [147]) – the linear spectrum [148,149,150] and "quantum" conductivity [122,151,152,153,154,155]. The latter indicates that the unusual feature of the electrical conductivity at charge neutrality is not its value, but rather the scattering mechanism behind it – electron-electron interaction. In contrast, the former ensures that the electron-electron interaction does not relax the energy current (since it is equivalent to the momentum flux, see Sec. 3), which implies that the thermal conductivity is determined by disorder scattering. As a result, the Lorenz number is expected to be proportional to the ratio of the disorder mean free time to the electron-electron scattering time, which in the hydrodynamic regime (or otherwise in ultra-clean graphene in the appropriate temperature interval) is assumed to be large, $\mathcal{L} \propto \tau_{\mathrm{dis}}/\tau_{ee} \gg 1$, see Fig. 16.

The intermediate nature of the hydrodynamic regime suggested by the data in Fig. 16 is corroborated by the results of the experiments on the thermoelectric power [40]. Here it manifested itself in the failure to uncover the ideal hydrodynamic limit where (in the absence of disorder) the thermopower

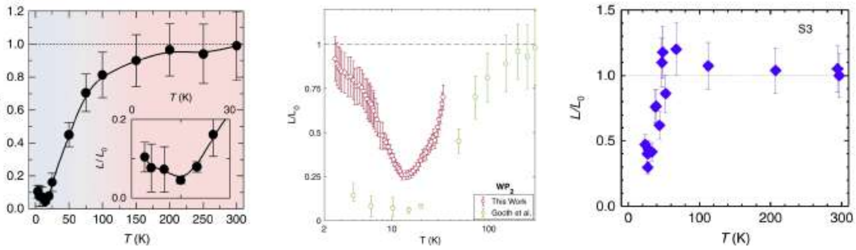


Fig. 17 Wiedemann-Franz law violation in topological materials. Left: the Lorenz number extracted from measurements of the electrical and thermal conductivities in a WP_2 micro-ribbon (width $2.5\,\mu$m) (Reprinted from Ref. [156]). The inset shows the zoomed-in low-temperature region of the same data. Central: the Lorenz number in bulk (mm-sized) single crystals of WP_2 (Reprinted from Ref. [31]). Green dots show the data from the left plot. Right: the Lorenz number in MoP (Reprinted from Ref. [157]).

equals the thermodynamic entropy per carrier charge [89,158]. Still, the observed thermopower at relatively high temperatures significantly exceeded the standard Mott relation indicating the hydrodynamic behavior [40].

Interestingly, the hydrodynamic theory predicts the Wiedemann-Franz law violation even in doped graphene (in the Fermi-liquid regime) [24,159] (for a detailed discussion of the Wiedemann-Franz law violation in Fermi liquids in general see Ref. [160]), but now the Lorenz number is predicted to be small (and in fact vanish in the limit of large densities, see Sec. 3). The effect can not be clearly seen in Fig. 16, presumably due to relatively low densities explored in the data shown. This prediction suggests a possible relation with the small Lorenz number observed in topological materials, which has not been fully addressed so far.

2.5.2 Small Lorenz number in topological materials

Recently, hints of electronic hydrodynamics have been observed in the topological material WP_2 [31,156], where the Wiedemann-Franz law is also strongly violated, see Fig. 17. The measured thermal and electrical conductivities in WP_2 exhibit features that are significantly different from those observed in graphene. The presumed hydrodynamic regime is limited to temperatures below 20 K (as determined by the electron-phonon scattering dominating transport properties at higher temperatures). Here, the measured Lorenz number turns out to be small, $\mathcal{L} \ll \mathcal{L}_0$, the result that was attributed to the existence of the hydrodynamic regime (confirmed by the extremely large measured values of the typical length scale describing momentum-relaxing scattering properties). Interestingly enough, similar effects have been observed in a different topological material, MoP [157].

The precise microscopic nature of the proposed hydrodynamic state and especially its relation to the hydrodynamic regime in graphene remains unclear. An interesting proposal on the experimental measurement of one of the relevant length scales, the "momentum-relaxing" length (e.g., $\ell_{\rm dis}$), which together with the "momentum-conserving" length ℓ_{ee} determines whether the sample is in the hydrodynamic, ballistic, or Ohmic regime, was suggested in Ref. [161]. The authors used Sondheimer oscillations [162] to extract $\ell_{\rm dis}$ even in the ballistic case $\ell_{\rm dis} \gg L$ (where L is the typical system size) and suggested that this effect can be used as an effective quantitative probe for identifying scattering processes in ultra-clean materials.

3 Electronic hydrodynamics

Hydrodynamic description of interacting particles (or excitations) has long been part of the theoretical toolbox used (in addition to traditional fluid mechanics [13]) in a wide range of fields including many-body theory [3], superfluids [15,26], quark-gluon plasma [163], or interstellar matter [164]. The underlying general idea allowing to develop the hydrodynamic theory suitable to such different circumstances is the universality of the long-time, long-wavelength behavior, i.e., the assumption that macroscopic (long-distance) physics is independent of microscopic details and is governed by symmetries, which can be expressed in terms of continuity equations.

The most common symmetry assumed in physics is time translation invariance leading to energy conservation. The corresponding continuity equation reads

$$\partial_t n_E + \boldsymbol{\nabla}\cdot\boldsymbol{j}_E = 0, \tag{10a}$$

where n_E is the energy density and $\boldsymbol{j}_E$ is the energy current.

The second conservation law typically assumed in the context of electronic systems is the particle number (or charge) conservation (manifesting gauge invariance) described by the continuity equation

$$\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0. \tag{10b}$$

Here n and $\boldsymbol{j}$ are the particle number and current densities while the charge and electric current densities differ by a factor of the electric charge, see also Eq. (1).

Supplementing equations (10a) and (10b) by the thermodynamic equation of state and the entropy balance equation [1] one may arrive at the macroscopic theory describing the long-distance properties of the system and find the spectrum of the collective modes. The resulting behavior is diffusive (i.e., equivalent to the standard Drude-like approach to electronic transport [10]).

In contrast, conventional fluids are additionally assumed to be translationally invariant which implies momentum conservation described by the continuity-like equation for the momentum density, $\boldsymbol{n_k}$,

$$\partial_t n^\alpha_{\boldsymbol{k}} + \nabla^\beta \varPi^{\alpha\beta}_E = 0. \tag{10c}$$

Here $\varPi^{\alpha\beta}_E$ is the momentum flux (or stress-energy) tensor. Introducing momentum conservation has a drastic effect on the collective modes of the system leading to the appearance of a mode with the linear dispersion, i.e., the sound mode [1]. The existence of the latter is the crucial distinction between hydrodynamics and other macroscopic, long-wavelength theories (although a more general interpretation of the term “hydrodynamics” is also used in literature, see, e.g., Ref. [1]).

The explicit form of the hydrodynamic equations can be obtained by supplementing the continuity equations (10) by the so-called “constitutive relations” reducing the amount of independent variables and turning Eqs. (10) into a closed set. This is typically done under the assumption of local equilibrium [13]. Moreover, the form of the stress-energy tensor in the moving fluid is often obtained by relating to the properties of the stationary fluid (that are assumed to be known). To do that, one needs to change the reference frame to the rest frame of the fluid. Consequently, traditional hydrodynamics [13] distinguishes the two cases of Galilean- and Lorentz-invariant fluids, i.e., the classical and relativistic hydrodynamics. While early applications of the hydrodynamic approach to electronic transport were based on the classical theory [17,22,165], it is the possibility of realization of relativistic hydrodynamics in graphene that ignited the current interest in the field.

3.1 Relativistic hydrodynamics in a solid state laboratory

The discovery of graphene and Dirac fermions in it [151] has provided a unique opportunity to study relativistic effects in a solid state laboratory [166]. In particular, early work on collective electronic flows attempted to adapt relativistic hydrodynamics in $(2+1)$ dimensions to Dirac fermions in graphene [147,167].

3.1.1 Ideal relativistic fluid

Standard equations of relativistic hydrodynamics [13] are encoded in the relation

$$\frac{\partial T_i^k}{\partial x^k} = 0, \tag{11}$$

where T^{ik} is the relativistic stress-energy tensor (in graphene, this is a 3×3 tensor in the $(2+1)$-dimensional space-time)

$$T^{ik} = w u^i u^k - p g^{ik}, \tag{12a}$$

with w and p being the enthalpy and pressure, respectively, in the local rest frame.

For the purposes of this review, it will be instructive to write down the explicit form of the individual components of T^{ik}: the energy density

$$T^{00} = \frac{w}{1 - u^2/v_g^2} - p, \tag{12b}$$

the momentum density (here we adopt the usual practice of denoting the space components by Greek indices, while the Roman indices refer to the space-time)

$$T^{0\alpha} = \frac{w u_\alpha}{v_g \left(1 - u^2/v_g^2\right)}, \tag{12c}$$

and finally the momentum flux density

$$T^{\alpha\beta} = \frac{w u_\alpha u_\beta}{v_g^2 \left(1 - u^2/v_g^2\right)} + p\delta_{\alpha\beta}. \tag{12d}$$

The energy flux density is proportional to the momentum density and is given by $v_g T^{0\alpha}$. This fact will be explored in more detail below.

The relativistic generalization of the Euler equation [168] can be obtained by projecting Eq. (11) onto the direction perpendicular to the 3-velocity u^i [13]. This yields

$$\frac{w}{1 - u^2/v_g^2} \left[\frac{\partial}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\right] \boldsymbol{u} + v_g^2 \boldsymbol{\nabla} p + \boldsymbol{u}\frac{\partial p}{\partial t} = 0. \tag{13a}$$

Supplementing the Euler equation (13a) by the relativistic continuity equation

$$\frac{\partial \left(n u^k\right)}{\partial x^k} = 0, \tag{13b}$$

and the thermodynamic equation of state

$$w = 3p, \tag{13c}$$

one can quickly convince oneself that the ideal flow described by Eq. (13a) is isentropic

$$\frac{\partial \left(s u^k\right)}{\partial x^k} = 0. \tag{13d}$$

Equations (13) represent the closed set of hydrodynamic equations describing an ideal (non-dissipative) flow of a single-component relativistic fluid in a (2 + 1)-dimensional space-time with the velocity v_g playing the role of the speed of light. This theory possesses a collective mode [24,121,122,169,170, 171,172,173,174,175,176] with the linear dispersion relation

$$\omega = \frac{v_g q}{\sqrt{2}}. \tag{14}$$

In the literature, this mode has been referred to as the "cosmic sound" [169] or the "second sound" [174].

3.1.2 Electronic fluid in graphene

The ideal hydrodynamic theory outlined in the previous Section can be considered a purely phenomenological since it is based on an implicit assumption of equilibrium in the local rest frame without discussing the physical processes responsible for the equilibration. In the case of graphene, that has to be electron-electron interaction, which is the *classical, three-dimensional* Coulomb interaction. The latter point refers to the fact that although graphene is atomically thin so that the electron motion is restricted to two dimensions, the electric field induced by the electron charges is not. The former point refers to the orders of magnitude difference between the electron velocity and the speed of light, $v_g \ll c$, preventing the above hydrodynamic theory and electromagnetic fields to be transformed by the same Lorentz transformation. This issue was addressed in detail in Ref. [147].

Coulomb interaction can be included in the hydrodynamic description by re-writing the relativistic Euler equation (11) in the form

$$\frac{\partial T_i^k}{\partial x^k} = \frac{e}{c} F_{ik} j^k. \tag{15}$$

Notice, that in the right-hand side of this equation one has to write the speed of light, which is inconsistent with the use of the velocity v_g in the stress-energy tensor (12). A possibility to resolve this issue was suggested in Ref. [147]. Indeed, redefining the electromagnetic field tensor F_{ik} and the current j^k as

$$F_{ik} = \begin{pmatrix} 0 & (c/v_g)E_x & (c/v_g)E_y \\ -(c/v_g)E_x & 0 & -B \\ -(c/v_g)E_y & B & 0 \end{pmatrix}, \tag{16}$$

$$j^k = \big(v_g n, \boldsymbol{j}\big), \tag{17}$$

one may remove the inconsistency from Eq. (15) turning it into the standard form of the relativistic Euler equation. However, this is only a partial solution since the redefined field tensor (16) leaves only two Maxwell's equations intact,

$$\boldsymbol{\nabla} \times \boldsymbol{B} = 0, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{E} = 4\pi e n, \tag{18}$$

while the other two are violated leaving the above approach questionable.

Even if the modified equation (15) can be accepted for those problems that do not involve the two violated Maxwell's equations (e.g., a description of stationary currents), there are other issues that prevent one from treating electronic flows in graphene as truly relativistic. As already mentioned above, there are other scattering processes in graphene (and in any other solid) affecting the behavior of charge carriers. These may include electron-phonon and disorder scattering, Auger processes, and three-particle collisions, none of which are Lorentz-invariant. Moreover, typical currents studied in present-day experiments are small enough, such that the hydrodynamic velocity is small as well, $u \ll v_g$. As a result, one would be interested in the non-relativistic limit of the hydrodynamic equation (15) anyways. Now, the non-relativistic form of hydrodynamics can also be derived within the kinetic theory approach (see the next Section), where all of the above issues can be consistently taken into account. In the absence of dissipative processes, the generalized Euler equation for the hydrodynamic electronic flows in graphene obtained from the kinetic theory does indeed closely resemble Eq. (13a), while containing additional terms taking into account scattering processes that were not considered so far. In addition, introducing dissipative processes within the phenomenological approach involves defining new parameters, such as electrical conductivity and viscosity, that can only be determined in an experiment. While the kinetic theory provides a method to "calculate" these parameters, the accuracy of these calculations may be limited depending on the initial assumptions allowing one to formulate the kinetic equation in the first place. The form of the dissipative corrections remains the same in both approaches providing a useful checkpoint.

3.2 Kinetic theory approach

Kinetic approach has been used to describe electronic transport in solids for decades [10]. While applicability of the kinetic theory to quantum many-body systems remains an active area of research [177], it is often assumed that at least at high enough temperatures electrons behave semiclassically such that the kinetic theory is applicable. At the same time, this implies that quasiparticle excitations are long-lived, the assumption that might not be valid in strongly correlated or hydrodynamic regimes. Strictly speaking, the kinetic equation can only be applicable in weakly interacting electronic systems. This might be a problem in graphene, where the effective coupling constant in an idealized model is $\alpha_g = e^2/v_g \approx 2.2$ (which may be reached in suspended graphene) and while an insulating substrate may reduce this value (by a factor of the dielectric constant), the resulting α_g is not small (typically, $\alpha_g \approx 0.2 \div 0.3$ [43,178]). Consequently, derivation of the hydrodynamic equations has to rely on universality: one assumes that the form of the equations is independent of the interaction strength (similarly to how the Navier-Stokes equation derived from the kinetic theory of rarefied gases [16] can be used to describe properties of water, where the kinetic equation is not applicable). Calculation of kinetic coefficients then has to rely on the renormalization group procedure [179,119] treating α_g as a running coupling constant [180,181,182,183]. One renormalizes the theory to the parameter regime, where the coupling constant is small, solves the kinetic equations, and then renormalizes back to realistic parameter values. For a more microscopic approach to deriving the hydrodynamic equations based on the nonequilibrium Keldysh technique, see Ref. [184]. This paper provides a proper microscopic treatment of inelastic electron-electron scattering that is responsible for establishing the local equilibrium that is the central assumption of the kinetic approach discussed below.

3.2.1 Quasiclassical Boltzmann equation

At high enough temperatures (where the hydrodynamic behavior is observed [27,39,41,42]), the quasiparticle spectrum in monolayer graphene [185] comprises two bands of carriers (the "conductance" and "valence" bands) that touch in the corners of the hexagonal Brillouin zone, i.e., at the "Dirac points" (multilayer graphene was discussed in Ref. [186]). In the vicinity of the Dirac points the spectrum can be approximately considered to be linear (logarithmic renormalization due to electron-electron interaction [181], see also Ref. [187], is observed at much lower temperatures [188]). The linearity of the Dirac spectrum leads to two important kinematic effects: (i) the suppression of Auger processes [89,189] and hence approximate conservation of the number of particles in each band independently [23,24,89,190]; and (ii) the so-called "collinear scattering singularity" [122,123,124,153,180,190,191,192]. The former represents an additional conservation law that is not taken into account in the above phenomenological hydrodynamics. The latter is justified by the smallness of the effective coupling constant and allows for a nonperturbative solution of the Boltzmann equation (recall that the Boltzmann approach itself is justified in the weak coupling limit, $\alpha_g \to 0$).

Consider now the two-band model of low-energy quasiparticles in graphene. Within the kinetic approach, the quasiparticles can be described by a distribution function, $f_{\lambda\boldsymbol{k}}$, where each quasiparticle state is characterized by the band index (or chirality), $\lambda = \pm 1$, and 2D momentum, $\boldsymbol{k}$. The spectrum is assumed to be linear,

$$\epsilon_{\lambda\boldsymbol{k}} = \lambda v_g k, \tag{19a}$$

with the straightforward relation between velocities and momenta,

$$\boldsymbol{v}_{\lambda\boldsymbol{k}} = \lambda v_g \frac{\boldsymbol{k}}{k}, \qquad \boldsymbol{k} = \frac{\lambda k}{v_g} \boldsymbol{v}_{\lambda\boldsymbol{k}} = \frac{\epsilon_{\lambda\boldsymbol{k}} \boldsymbol{v}_{\lambda\boldsymbol{k}}}{v_g^2}. \tag{19b}$$

The distribution function satisfies the kinetic (Boltzmann) equation

$$\mathcal{L} f_{\lambda\boldsymbol{k}} = \mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}] + \mathrm{St}_R[f_{\lambda\boldsymbol{k}}] + \mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}], \tag{20a}$$

where the left-hand side (LHS) is defined by the Liouville's operator

$$\mathcal{L} = \partial_t + \boldsymbol{v}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}} + \left(e\boldsymbol{E} + \frac{e}{c}\boldsymbol{v}\times\boldsymbol{B}\right)\cdot\boldsymbol{\nabla}_{\boldsymbol{k}}, \tag{20b}$$

and the right-hand side (RHS) represents the collision integral.

In the simplest Golden-Rule-like approximation, different scattering processes contribute to the collision integral in the additive fashion; hence, the form of the RHS in Eq. (20a). In the hydrodynamic regime, the electron-electron interaction (described by St_{ee}) is the dominant scattering process responsible for equilibration of the system. Consequently, *local equilibrium* is described by the distribution function that nullifies St_{ee} [16]

$$\mathrm{St}_{ee}\left[f^{(le)}_{\lambda\boldsymbol{k}}\right] = 0, \qquad f^{(le)}_{\lambda\boldsymbol{k}} = \left\{1 + \exp\left[\frac{\epsilon_{\lambda\boldsymbol{k}} - \mu_\lambda(\boldsymbol{r}) - \boldsymbol{u}(\boldsymbol{r})\cdot\boldsymbol{k}}{T(\boldsymbol{r})}\right]\right\}^{-1}, \tag{21}$$

where $\mu_\lambda(\boldsymbol{r})$ is the local chemical potential and $\boldsymbol{u}(\boldsymbol{r})$ is the hydrodynamic (or "drift") velocity. The local equilibrium distribution function (21) allows for independent chemical potentials in the two bands, which can be expressed in terms of the "thermodynamic" and "imbalance" [89] chemical potentials

$$\mu_\lambda = \mu + \lambda \mu_I. \tag{22}$$

In global equilibrium (i.e., for stationary fluid)

$$f^{(0)} = f^{(le)}_{\lambda\boldsymbol{k}}(\mu_I = 0, \boldsymbol{u} = 0). \tag{23}$$

In addition, two more scattering processes need to be taken into account. Even ultra-pure graphene samples contain some degree of (weak) disorder. Scattering on impurities violates momentum conservation leading to a weak decay term in the generalized Euler equation [23,24,120,121]. This process (as well as other momentum-relaxing processes) is described in Eq. (20a) by $\mathrm{St}_{\mathrm{dis}}$. At the same time, electron-phonon interaction may lead not only to the loss of electronic momentum (which is already taken into account in $\mathrm{St}_{\mathrm{dis}}$), but also to the loss of energy. Consequently, despite being subdominant in the hydrodynamic regime the electron-phonon interaction should be taken into account as one of the dissipative processes. However, due to the linearity of the Dirac spectrum, lowest order scattering on acoustic phonons is kinematically suppressed. Instead, it is a higher order process, the so-called disorder-assisted electron-phonon scattering [193] or "supercollisions" [194,195,196,197], that plays the most important role in the hydrodynamic regime. Indeed, supercollisions violate not only energy conservation, but also conservation of the number of particles in each band. As a result, continuity equations for energy and single-band particle numbers also acquire weak decay terms. In the kinetic equation (20a), these effects are described by St_R (the subscript "R" here stands for "recombination", see Refs. [86,87,88,89,114,198,199,200]).

Within the kinetic theory, conservation laws are manifested in the sum rules for the collision integrals. There are four conservation laws to consider: energy, momentum, and particle number in the two bands. The latter can be expressed in terms of the "charge" and "total quasiparticle" (or imbalance) numbers similarly to Eq. (22)

$$n_\lambda = \frac{1}{2}\left(\lambda n + n_I\right). \tag{24}$$

The continuity equation (10b) representing global charge conservation can be obtained by summing the kinetic equation (20a) over all quasiparticle states. During this procedure, all three collision integrals in Eq. (20a) vanish [16]

$$N\sum_\lambda \int \frac{d^2k}{(2\pi)^2} \mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}] = N\sum_\lambda \int \frac{d^2k}{(2\pi)^2} \mathrm{St}_R[f_{\lambda\boldsymbol{k}}] = N\sum_\lambda \int \frac{d^2k}{(2\pi)^2} \mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}] = 0. \tag{25a}$$

Moreover, electron-electron and disorder scattering also conserve the number of particles in each band, such that

$$N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\lambda\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\lambda\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=0, \tag{25b}$$

whereas supercollisions lead to a decay term in the continuity equation for the imbalance density

$$N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\lambda\,\mathrm{St}_R[f_{\lambda\boldsymbol{k}}]\approx-\mu_I n_{I,0}\lambda_Q\approx-\frac{n_I-n_{I,0}}{\tau_R}. \tag{25c}$$

Here $n_{I,0}$ is the imbalance density at global equilibrium, see Eq. (23), i.e., for $\mu_I=0$ and $\boldsymbol{u}=0$. The first equality in Eq. (25c) was suggested in Ref. [89] and serves as the definition of the dimensionless coefficient λ_Q, while the second (valid to the leading order) was suggested in Refs. [87,120] and offers the definition of the "recombination time" τ_R (see also Ref. [198]). The two expressions are equivalent since $n_I-n_{I,0}\propto\mu_I$.

Similarly, both electron-electron and disorder scattering conserve energy, hence the corresponding collision integrals vanish upon summation over all quasiparticle states with an extra factor of energy

$$N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=0. \tag{25d}$$

Integrating the "recombination" collision integral one finds [201]

$$N\sum_{\boldsymbol{k}}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_R[f_{\lambda\boldsymbol{k}}]=-\mu_I n_{E,0}\lambda_{QE}\approx-\frac{n_E-n_{E,0}}{\tau_{RE}}. \tag{25e}$$

The equivalence of the two forms of the decay term stems from $n_E-n_{E,0}\propto\mu_I$ assuming the electrons and holes are characterized by the same temperature.

Supercollisions contribute differently to recombination and energy relaxation. Recombination typically implies scattering between the quasiparticle states in different bands only. At the same time, supercollisions may also take place within a single band [193]. This process does not affect the number of particles in the band, but is accompanied by the energy loss as the electron scatters from a higher energy state into a lower energy state (losing its momentum to the impurity). Consequently, this process provides an additional contribution to energy relaxation. Thus, the time scales τ_R and τ_{RE} should be quantitatively different, although of the same order of magnitude (at least at charge neutrality and in the hydrodynamics regime).

Now, other processes may contribute to τ_R and τ_{RE}, including direct electron-phonon scattering [86,89,124,192,193,202,203], scattering on optical phonons [158,204], three-particle collisions [24,204], and Auger processes [23, 24,89,190]. Taking into account these effects does not change the functional form of the continuity equations leaving the integrated collision integrals (25c) and (25e) intact, but may affect the theoretical estimates of the values of τ_R

and τ_{RE} (see Refs. [193,201]). Given the approximate nature of such calculations, one may treat these parameters as phenomenological taking into account all relevant scattering processes.

Finally, electron-electron interaction conserves momentum and hence

$$N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\boldsymbol{k}\,\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=0. \tag{25f}$$

On the other hand, weak disorder scattering leads to a weak decay term that should be included in Eq. (10c). Within the simplest τ-approximation [16,120]

$$N\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\boldsymbol{k}\,\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=\frac{\boldsymbol{n}_{\boldsymbol{k}}}{\tau_{\mathrm{dis}}}. \tag{25g}$$

The remaining collision integral St_R also does not conserve momentum, but given the phenomenological nature of τ_{dis} [43] (a better version of the disorder collision integral in graphene should involve the Dirac factors suppressing backscattering [205] which would lead to the similar approximation but with the transport scattering time, which in graphene differs by a factor of 2), the contribution of the next-order supercollisions (involving both disorder and phonons) may be considered to be included in τ_{dis} (similarly to the above discussion of τ_R and τ_{RE}).

3.2.2 Continuity equations in graphene

Using the above properties of the collision integrals, one can easily obtain the continuity equations in graphene [23,24,120] by integrating the kinetic equation (20a). In comparison to the "phenomenological" continuity equations (10), the resulting equation will contain extra terms due to the weak decay processes (discussed in the previous Section) and external electromagnetic fields. Hence the only true symmetry of the electronic fluid in a solid is gauge invariance that manifests itself by means of the continuity equation (10b)

$$\partial_t n+\boldsymbol{\nabla}\cdot\boldsymbol{j}=0, \tag{26a}$$

where the kinetic definitions of the "charge" density and current are [cf. Eq. (24)]

$$n=n_+-n_-,\qquad n_+=N\int\frac{d^2k}{(2\pi)^2}f_{+,\boldsymbol{k}},\qquad n_-=N\int\frac{d^2k}{(2\pi)^2}\left(1-f_{-,\boldsymbol{k}}\right), \tag{26b}$$

and [cf. Eq. (1)]

$$\boldsymbol{j}=\boldsymbol{j}_+-\boldsymbol{j}_-=N\int\frac{d^2k}{(2\pi)^2}\left[\boldsymbol{v}_{+,\boldsymbol{k}}f_{+,\boldsymbol{k}}-\boldsymbol{v}_{-,\boldsymbol{k}}\left(1-f_{-,\boldsymbol{k}}\right)\right]. \tag{26c}$$

In the two-band model of graphene, the number of particles in each band is approximately conserved (see above). Hence, in addition to Eq. (26a), one finds a continuity equation for the "imbalance density", see Eq. (24),

$$\partial_t n_I + \boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\frac{n_I - n_{I,0}}{\tau_R}, \tag{26d}$$

where

$$n_I = n_+ - n_-, \qquad \boldsymbol{j} = \boldsymbol{j}_+ + \boldsymbol{j}_-, \tag{26e}$$

and the RHS in Eq. (26d) comes from integrating the collision integral, see Eq. (25c).

The continuity equation for the energy density is obtained by multiplying the kinetic equation (20a) by $\epsilon_{\lambda\boldsymbol{k}}$ and summing over all quasiparticle states,

$$\partial_t n_E + \boldsymbol{\nabla}\cdot\boldsymbol{j}_E = e\boldsymbol{E}\cdot\boldsymbol{j} - \frac{n_E - n_{E,0}}{\tau_{RE}}, \tag{26f}$$

where n_E and $\boldsymbol{j}_E$ are defined as

$$n_E = N\sum_\lambda \int \frac{d^2k}{(2\pi)^2}\epsilon_{\lambda\boldsymbol{k}} f_{\lambda\boldsymbol{k}} \tag{26g}$$

and

$$\boldsymbol{j}_E = N v_g^2 \sum_\lambda \int \frac{d^2k}{(2\pi)^2}\boldsymbol{k} f_{\lambda\boldsymbol{k}} = v_g^2 \boldsymbol{n}_{\boldsymbol{k}}. \tag{26h}$$

The last equality represents the fact that in graphene the momentum density is proportional to the energy density [due to the properties of the Dirac spectrum Eq. (19)]. The two terms in the RHS in Eq. (26f) come from the Lorentz term in the Liouville's operator (20b) and the integrated collision integral, see Eq. (25e). The former physically represents Joule's heat.

Finally, the continuity equation representing momentum conservation is obtained by multiplying the kinetic equation (20a) by $\boldsymbol{k}$ and summing over all states. In contrast to the "phenomenological" equation (10c), the resulting equations contains extra terms stemming from the effect of the electromagnetic field and weak disorder (25g)

$$\partial_t n^\alpha_{\boldsymbol{k}} + \boldsymbol{\nabla}^\beta \mathit{\Pi}^{\alpha\beta}_E - enE^\alpha - \frac{e}{c}\left[\boldsymbol{j}\times\boldsymbol{B}\right]^\alpha = -\frac{n^\alpha_{\boldsymbol{k}}}{\tau_{\rm dis}}. \tag{26i}$$

Here $\boldsymbol{n}_{\boldsymbol{k}}$ is given by Eq. (26h) and the momentum flux tensor is defined as

$$\mathit{\Pi}^{\alpha\beta}_E = N\sum_\lambda \int \frac{d^2k}{(2\pi)^2} k^\alpha v^\beta_{\lambda\boldsymbol{k}} f_{\lambda\boldsymbol{k}}. \tag{26j}$$

3.2.3 Constitutive relations

Continuity equations represent the global conservation laws and are valid without any further assumptions. Hydrodynamics, however, assumes that the set of continuity equations can be closed by expressing the vector and tensor quantities (i.e., the currents and stress-energy tensor) in terms of the "velocity field" $\boldsymbol{u}(\boldsymbol{r})$. Such expressions are known as "constitutive relations". Phenomenologically, they can be derived using the Galilean or (in the relativistic case) Lorentz invariance [13]. However, neither is valid for Dirac fermions in graphene (the former due to the linear spectrum and the latter due to the classical nature of the Coulomb interaction, see Sec. 3.1.2). Instead, one can derive the constitutive relations from the kinetic theory under the assumption of local equilibrium [23, 120]. Indeed, substituting the local equilibrium distribution function into the definitions of the three currents (26c), (26e), and (26h) yields the expected relations

$$\boldsymbol{j} = n\boldsymbol{u}, \qquad \boldsymbol{j}_I = n_I\boldsymbol{u}, \qquad \boldsymbol{j}_E = \mathcal{W}\boldsymbol{u}, \tag{27a}$$

where $\mathcal{W}$ is the enthalpy density. This thermodynamic quantity can also be evaluated using the local equilibrium distribution function, which yields the "equation of state"

$$\mathcal{W} = n_E + P = \frac{3n_E}{2+u^2/v_g^2}, \tag{27b}$$

where P is the thermodynamic pressure. Both of these quantities appear in the explicit expression of the momentum flux tensor

$$\Pi_E^{\alpha\beta} = P\delta^{\alpha\beta} + \frac{\mathcal{W}}{v_g^2}u^\alpha u^\beta. \tag{27c}$$

Combining Eqs. (27) with the continuity equation for momentum density (26i), one may generalize the Euler equation [168] to Dirac quasiparticles in graphene

$$\mathcal{W}(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} + v_g^2\boldsymbol{\nabla}P + \boldsymbol{u}\partial_t P + e(\boldsymbol{E}\cdot\boldsymbol{j})\boldsymbol{u} = v_g^2\left[en\boldsymbol{E} + \frac{e}{c}\boldsymbol{j}\times\boldsymbol{B}\right] - \frac{\mathcal{W}\boldsymbol{u}}{\tau_{\mathrm{dis}}}. \tag{28}$$

It is instructive to compare Eq. (28) to the relativistic version of the Euler equation, Eq. (13a). Formally, the first three terms in the LHS of Eq. (28) coincide with the three terms of Eq. (13a). The rest of the terms – the Joule's heat, Lorentz force, and weak decay due to disorder – have not been considered in the relativistic theory and are explicitly not Lorentz-invariant. Even though the first three terms in Eq. (28) have the same form as Eq. (13a), there is a subtle difference: the pressure p in Eq. (13a) is the thermodynamic pressure in the local rest frame, while P in Eq. (28) is the pressure in the laboratory frame. The latter is evaluated with the distribution function (21) and hence is a function of the velocity $\boldsymbol{u}$, while $p = P(\boldsymbol{u} = 0)$. This point is the only difference between the relativistic equation of state (13c) and Eq. (27b) as well.

The generalized Euler equation (28) together with the continuity equations (26a), (26d), and (26f) describe the “ideal” flow of the electronic fluid. In conventional hydrodynamics “ideal” means “in the absence of dissipation”, which is not quite the case here, since weak disorder scattering, quasiparticle recombination, and energy relaxation are already taken into account. However, none of these processes are due to electron-electron interaction and hence are absent in the conventional theory [13].

3.2.4 Dissipative corrections

In its simplest form, conventional hydrodynamics [13,16] considers a system of particles (atoms, molecules, etc.) with the contact (short-range) interaction, such that individual scattering processes are almost literally “collisions”. These collisions represent the physical process responsible for equilibration: if the system is driven out of equilibrium, they tend to restore it. In the process the system is bound to lose energy, hence the collisions are responsible for dissipation.

In graphene (and other solids, see below), the situation is slightly more involved, but the main idea remain the same – physical processes responsible for equilibration lead to dissipation that is described by “kinetic coefficients”. This can be described as follows [13,16]. Nonequilibrium states are characterized by nonzero macroscopic current. In the process of equilibration the currents relax (their values are being reduced towards zero). Hence, the quasiparticle currents (27a) acquire additional terms – the dissipative corrections [23,24, 120,176]

$$\boldsymbol{j} = n\boldsymbol{u} + \delta\boldsymbol{j}, \qquad \boldsymbol{j}_I = n_I\boldsymbol{u} + \delta\boldsymbol{j}_I. \tag{29a}$$

In the absence of magnetic field, the dissipative corrections are related to external bias by means of a “conductivity matrix” [120,89,204]

$$\begin{pmatrix} \delta\boldsymbol{j} \\ \delta\boldsymbol{j}_I \end{pmatrix} = \widehat{\Sigma} \begin{pmatrix} e\boldsymbol{E} - T\boldsymbol{\nabla}(\mu/T) \\ -T\boldsymbol{\nabla}(\mu_I/T) \end{pmatrix}. \tag{29b}$$

At charge neutrality $\mu = \mu_I = 0$ the matrix $\widehat{\Sigma}$ is diagonal. In the absence of disorder, the upper diagonal element defines the “quantum” or “intrinsic” conductivity [23,24,89,120,204]

$$\sigma_Q = e^2\Sigma_{11}(0). \tag{29c}$$

The third current $\boldsymbol{j}_E$ does not acquire a dissipative correction since it is proportional to the momentum density, see Eq. (26h), and electron-electron interaction conserves momentum. This point represents the key difference between electronic hydrodynamics in graphene (or any semimetal with linear spectrum) from conventional fluid mechanics of systems with parabolic (Galilean-invariant) spectrum. In the latter case, it is the particle number (or mass) current $\boldsymbol{j}$ that is proportional to the momentum density. As a result, the energy current gets a dissipative correction described by the thermal conductivity

$\varkappa$ that is determined by interparticle collisions. In the hydrodynamic theory of graphene, the role that is equivalent to that of $\varkappa$ is played by the elements of the matrix $\widehat{\Sigma}$. The matrix nature of $\widehat{\Sigma}$ reflects the band structure of graphene. In the case of strong recombination, the imbalance mode becomes irrelevant and one is left with the single dissipative coefficient σ_Q, see Ref. [24]. Now, the thermal conductivity in graphene arises purely due to weak disorder scattering that is already taken into account in the Euler equation (28). This is the reason for the strong violation of the Wiedemann-Franz law in neutral graphene, see Sec. 2.5.1.

The kinetic coefficients $\widehat{\Sigma}$ can be found by solving the kinetic equation (20a) perturbatively using the standard procedure [16,24,121,147]. In a bulk system and in the absence of magnetic field, this calculation was performed in detail in Ref. [120], where a 3×3 matrix was considered [i.e., adding the energy current and its relaxation due to weak disorder to Eq. (29b)]. The following 2×2 matrix was introduced in Ref. [176]. In both cases, one expresses the matrix $\widehat{\Sigma}$ as a linear combination of the interaction and disorder contributions

$$\widehat{\Sigma}=\widehat{\mathfrak{M}}\widehat{\mathfrak{S}}_{xx}^{-1}\widehat{\mathfrak{M}}, \qquad \widehat{\mathfrak{S}}_{xx}=\frac{\alpha_g^2T^2}{2\mathcal{T}^2}\widehat{\mathfrak{T}}+\frac{\pi}{\mathcal{T}\tau_{\rm dis}}\widehat{\mathfrak{M}}, \tag{30a}$$

where (following the 2×2 notation)

$$\widehat{\mathfrak{M}}=\begin{pmatrix} 1-\frac{2\tilde{n}^2}{3\tilde{n}_E}\frac{T}{\mathcal{T}} & \frac{xT}{\mathcal{T}}-\frac{2\tilde{n}\tilde{n}_I}{3\tilde{n}_E}\frac{T}{\mathcal{T}} \\ \frac{xT}{\mathcal{T}}-\frac{2\tilde{n}\tilde{n}_I}{3\tilde{n}_E}\frac{T}{\mathcal{T}} & 1-\frac{2\tilde{n}_I^2}{3\tilde{n}_E}\frac{T}{\mathcal{T}} \end{pmatrix}, \tag{30b}$$

with dimensionless densities (in self-evident notation; $Li_n(z)$ is the polylogarithm)

$$\tilde{n}=\mathrm{Li}_2\left(-e^{-x}\right)-\mathrm{Li}_2\left(-e^{x}\right), \quad \tilde{n}_I=\frac{x^2}{2}+\frac{\pi^2}{6}, \quad \tilde{n}_E=-\mathrm{Li}_3\left(-e^{x}\right)-\mathrm{Li}_3\left(-e^{-x}\right),$$

$$x=\mu/T, \quad \mathcal{T}=2T\ln\left[2\cosh(x/2)\right], \tag{30c}$$

and dimensionless scattering rates

$$\widehat{\mathfrak{T}}=\begin{pmatrix} t_{11}^{-1} & t_{12}^{-1} \\ t_{12}^{-1} & t_{22}^{-1} \end{pmatrix}, \qquad t_{ij}^{-1}=\frac{8\pi\mathcal{T}}{\alpha_g^2NT^2}\tau_{ij}^{-1}. \tag{30d}$$

Here τ_{ij}^{-1} represent the integrated collision integral appearing while solving the kinetic equation within the three-mode approximation [121,120,153,174]. The fact that the collision integrals can be represented by the effective scattering rates τ_{ij}^{-1} is not equivalent to the simplest τ approximation that was employed above for the collision integrals $\mathrm{St}_{\rm dis}$ and St_R. Instead, this is simply a manifestation of the dimensionality of a collision integral (that is inverse time).

Numerical values of the scattering rates (30d) were discussed in Ref. [206]. In particular, at charge neutrality the off-diagonal elements vanish, $t_{12}^{-1}(0)=0$.

The diagonal element $t_{11}^{-1}(0)$ determines the "intrinsic" or "quantum" conductivity matrix, σ_Q. For small $x \ll 1$ the dimensionless "scattering rates" t_{ij} have the form [206]

$$\frac{1}{t_{11}} = \frac{1}{t_{11}^{(0)}} + x^2 \left(\frac{1}{t_{11}^{(2)}} - \frac{1}{8\ln 2} \frac{1}{t_{11}^{(0)}} \right) + \mathcal{O}(x^3), \tag{31a}$$

$$\frac{1}{t_{12}} = \frac{x}{t_{12}^{(1)}} + \mathcal{O}(x^3), \tag{31b}$$

$$\frac{1}{t_{22}} = \frac{1}{t_{22}^{(0)}} + x^2 \left(\frac{1}{t_{22}^{(2)}} - \frac{1}{8\ln 2} \frac{1}{t_{22}^{(0)}} \right) + \mathcal{O}(x^3). \tag{31c}$$

For unscreened Coulomb interaction, the dimensionless quantities $t_{ij}^{(0,1,2)}$ are independent on any physical parameter. Numerically, one finds the values [176] (neglecting the small exchange contribution [191]):

$$\left(t_{11}^{(0)}\right)^{-1} \approx 34.63, \qquad \left(t_{11}^{(2)}\right)^{-1} \approx 5.45,$$

$$\left(t_{12}^{(1)}\right)^{-1} \approx 5.72, \qquad \left(t_{22}^{(0)}\right)^{-1} \approx 19.73, \qquad \left(t_{22}^{(2)}\right)^{-1} \approx 5.65.$$

In the case of screened interaction, the quantities $t_{ij}^{(0,1,2)}$ depend on the screening length.

The above values for the effective scattering rates yield the following value for the intrinsic conductivity

$$\sigma_Q = \mathcal{A} e^2/\alpha_g^2, \qquad \mathcal{A} \approx 0.12. \tag{32}$$

The quantity σ_Q was studied by multiple authors [24,120,121,122,123,147,152,153,191] and is a temperature-dependent constant. This temperature dependence appears due to the logarithmic renormalization of the coupling constant α_g [179].

The above theoretical values can be related to the experimental data of Ref. [43]. Using the value of the coupling constant $\alpha_g \approx 0.23$ that is consistent with measurements at charge neutrality, the dimensionfull scattering rates at a typical temperature $T = 267\,\mathrm{K}$ have the following values

$$\tau_{11}^{-1} \approx 7.35\,\mathrm{THz}, \qquad \tau_{22}^{-1} \approx 4.17\,\mathrm{THz}.$$

The disorder scattering rate at $T = 267\,\mathrm{K}$ can be estimated as

$$\tau_{\mathrm{dis}}^{-1} \approx 0.8\,\mathrm{THz}.$$

In the opposite limit of strongly doped graphene, $x \gg 1$, all elements of the matrix (30d) coincide approaching the value [120,176,206]

$$t_{ij}^{-1}(\mu \gg T) \to \frac{8\pi^2}{3}. \tag{33a}$$

The reason for this is the exponentially small contribution of the second band in which case the two currents $\boldsymbol{j}$ and $\boldsymbol{j}_I$ coincide. In this limit, the corresponding dimensionfull rate vanishes

$$\tau_{11}^{-1} \approx \frac{\pi N \alpha_g^2 T^2}{3\mu}, \tag{33b}$$

leading to the vanishing dissipative corrections to the quasiparticle currents

$$\delta\boldsymbol{j} = \delta\boldsymbol{j}_I \to 0. \tag{34}$$

As a result, electric current has the hydrodynamic form (3) leading to the use of the hydrodynamic approach to electronic transport in doped graphene, both theoretically [101,95,96,207,100] and experimentally [27,41,42,51]

In the presence of magnetic field or in confined geometries the dissipative corrections to quasiparticle currents are more complicated. External magnetic field entangles all three modes and hence the corrections to quasiparticle currents acquire a dependence on the hydrodynamic velocity $\boldsymbol{u}$ [120]. In confined geometries, the coordinate dependence of the distribution function becomes important and as a result the dissipative corrections (29) become non-uniform [125]. In that case, the usual local conductivity may become poorly defined [53,125], but the issue remains insufficiently explored.

3.2.5 Electronic viscosity

Dissipative processes also contribute a correction to the momentum flux (or stress-energy) tensor (26j). In the non-relativistic limit, one writes the dissipative correction to $\varPi_E^{\alpha\beta}$ [here $\varPi_{E,0}^{\alpha\beta}$ denotes the tensor given in Eq. (27c)]

$$\varPi_E^{\alpha\beta} = \varPi_{E,0}^{\alpha\beta} + \delta\varPi_E^{\alpha\beta}, \tag{35a}$$

to the leading order in gradient expansion as

$$\delta\varPi_E^{\alpha\beta} = \eta^{\alpha\beta\gamma\delta}\nabla^\gamma u^\delta, \tag{35b}$$

where $\eta^{\alpha\beta\gamma\delta}$ is the rank-four viscosity tensor [13]. In a fully rotationally-invariant system the explicit form of the viscosity tensor is dictated by symmetry and in 2D is given by

$$\eta^{\alpha\beta\gamma\delta} = \eta\left(\delta^{\alpha\gamma}\delta^{\beta\delta} + \delta^{\alpha\delta}\delta^{\beta\gamma}\right) + (\zeta - \eta)\delta^{\alpha\beta}\delta^{\gamma\delta}, \tag{35c}$$

where η and ζ are the shear and bulk viscosity, respectively.

In graphene, the bulk viscosity vanishes, at least to the leading approximation [23,24,113,121,152], similarly to the situation in ultrarelativistic systems [16,208] and Fermi liquids [26,209] (although it may appear in disordered systems in magnetic field [210]). As a result, the leading term of the gradient expansion of the dissipative stress tensor has the form [13,119,120]

$$\delta\varPi_E^{\alpha\beta} = -\eta\boldsymbol{\mathfrak{D}}^{\alpha\beta}, \tag{35d}$$

where

$$\boldsymbol{\mathfrak{D}}^{\alpha\beta} = \nabla^\alpha u^\beta + \nabla^\beta u^\alpha - \delta^{\alpha\beta} \boldsymbol{\nabla}\cdot\boldsymbol{u}. \tag{35e}$$

In the presence of magnetic field, the shear viscosity acquires a field dependence [109,110,165] and the correction to the stress tensor gains an additional contribution

$$\delta \mathit{\Pi}_E^{\alpha\beta} = -\eta(B)\boldsymbol{\mathfrak{D}}^{\alpha\beta} + \eta_H(B)\epsilon^{\alpha ij}\boldsymbol{\mathfrak{D}}^{i\beta} e_B^j, \tag{35f}$$

where $\boldsymbol{e}_B = \boldsymbol{B}/B$ and $\eta_H(B)$ is the Hall [109,110,113,120,165,211,212,213, 214] viscosity. While the sign of η is fixed by thermodynamics [13,16], the sign of η_H is not. Equation (35f) follows Ref. [42]: the Hall viscosity is positive for electrons [120] (and negative for holes).

Electronic viscosity can be calculated in two different ways. As a linear response function relating stress to strain [211,215], the viscosity tensor can be found using a Kubo formula [211,215,216] (that can be related to the usual Kubo formula for conductivity [211]). Such calculations are mostly perturbative and were used to evaluate viscosity in strongly doped graphene [216] and in the high-frequency (collisionless) regime [215], as well as in disordered 2D electron systems beyond the hydrodynamic regime [217]. A further extension of this approach yields higher order corrections, such as "drag viscosity" [218] (by analogy to Coulomb drag [219]). Alternatively, one can proceed with the solution of the kinetic equation (20a) following the standard procedure [16, 120,121]. For arbitrary carrier density this yields a somewhat cumbersome expression that can only be analyzed numerically [119], but simplifies in the limiting cases of neutral and strongly doped graphene.

At charge neutrality and in the absence of magnetic field, the only energy scale in the problem is the temperature T and hence the shear viscosity has the form [180]

$$\eta(\mu{=}0, B{=}0) = \mathcal{B}\frac{T^2}{\alpha_g^2 v_g^2}. \tag{36}$$

The coefficient $\mathcal{B}$ has been evaluated in Ref. [180] to have the value $\mathcal{B} \approx 0.45$. This result was later confirmed in Ref. [120]. In both cases, the numerical value was obtained with the simplest model of unscreened Coulomb interaction, which is valid for small α_g, i.e., in the regime of formal validity of the kinetic approach (as well as the three-mode approximation allowing for nonperturbative results). At realistic parameter values one has to supplement kinetic calculations by the renormalization group (RG) approach treating α_g as a running coupling constant [179,180,181,182,183]. However, the product $\alpha_g v_g$ remains constant along the RG flow [180,191], such that Eq. (36) represents the correct form of shear viscosity in graphene at low temperatures and $B = 0$ [179].

Experimentally, a measurement of the shear viscosity is nontrivial [140]. However, nonlocal resistance measurements [27] yield an estimate of a related quantity, the kinematic viscosity, see Eq. (2). In graphene, the kinematic viscosity is defined as

$$\nu = \frac{v_g^2 \eta}{\mathcal{W}}. \tag{37}$$

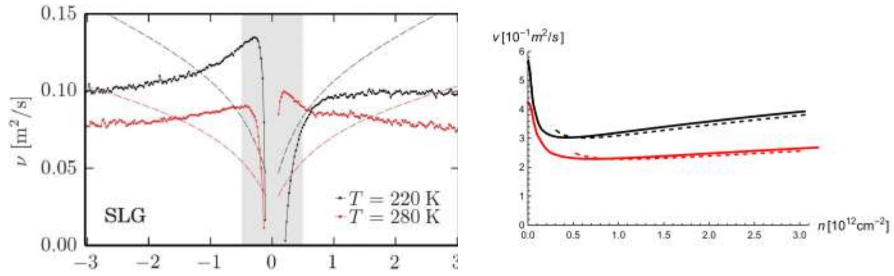


Fig. 18 Kinematic viscosity in monolayer graphene. Left: experimental data of Ref. [27] obtained by means of vicinity resistance measurements, see Sec. 2.2.2 (From Ref. [27]. Reprinted with permission from AAAS). Right: theoretical result of Ref. [119] obtained using the kinetic theory and renormalization group techniques (Reprinted with permission from Ref. [119]. Copyright (2019) by the American Physical Society).

The appearance of the enthalpy density in this definition is a manifestation of the fact that the hydrodynamic flow in graphene is the energy flow, see Eq. (26h). At charge neutrality, the kinematic viscosity is determined by the ratio of the velocity and coupling constant rather than their product [119]

$$\nu(\mu{=}0, B{=}0) \propto \frac{v_g^2}{\alpha_g^2 T}, \tag{38}$$

and hence is renormalized along the RG flow. In doped graphene, the dominant temperature dependence of the kinematic viscosity can be estimated as [119]

$$\nu(\mu \gg 1, B{=}0) \propto \frac{v_g^2 \mu}{\alpha_g^2 T^2} \frac{1}{1{+}T^2/\mu^2}. \tag{39}$$

This expression disregards additional temperature dependence arising from the RG and extra logarithmic factors [24, 119, 216].

Taking into account renormalization and screening effects, one can reach a quantitative estimate of the kinematic viscosity that is of the same order of magnitude as the experimental data reported in Ref. [27], see Fig. 18. Close to charge neutrality, the theoretical results show excellent agreement with the data reported in Ref. [50] as shown in Fig. 11 (see, however, Sec. 2.4.2 for the discussion of the controversial nature of that data).

The field dependence of the shear viscosity was discussed semiclassically in Refs. [42, 109, 110, 165] in the context of a single-component Fermi liquid or strongly doped graphene (where only one band contributes to low-energy physical properties). The resulting behavior is similar to the conventional magnetoconductivity [10]

$$\eta(B; \mu \gg T) = \frac{\eta(B=0; \mu \gg T)}{1 + \mathit{\Gamma}_B^2}, \tag{40a}$$

$$\eta_H(B; \mu \gg T) = \eta(B=0; \mu \gg T) \frac{\mathit{\Gamma}_B}{1 + \mathit{\Gamma}_B^2}, \tag{40b}$$

where

$$\varGamma_B = 2\omega_B\tilde{\tau}_{11}, \quad \omega_B = |e|v_g^2B/(\mu c). \tag{40c}$$

The kinetic approach [119] allows one to identify the scattering rate $\tilde{\tau}_{11}$ appearing in Eqs. (40). Indeed, this rate should be distinguished [121,216] from the transport scattering rate [122,220] that determines the electrical conductivity and the "quantum" scattering rate [122] that determines the quasiparticle lifetime. At the same time, the kinetic theory yields the field dependence of the shear viscosity at charge neutrality as well [119]

$$\eta(B;\mu=0) = \frac{T^2}{\alpha_g^2v_g^2}\frac{\mathcal{B}+\mathcal{B}_1\gamma_B^2}{1+\mathcal{B}_2\gamma_B^2}, \tag{41}$$

where

$$\gamma_B = \frac{|e|v_g^2B}{\alpha_g^2cT^2}, \tag{42}$$

where $\mathcal{B}_1 \approx 0.0037$ and $\mathcal{B}_2 \approx 0.0274$. In contrast to the Fermi liquid results, the shear viscosity at $\mu = 0$ does not vanish in the limit of classically strong field.

Frequency-dependent viscosity was analyzed in Refs. [211,216,221]. In particular, Ref. [221] suggested an existence of a resonance in strong magnetic fields (as well as the corresponding plasmon damping). Momentum-dependent viscosity in Fermi liquids (due to head on collisions [222,223]) was suggested in Ref. [224] (for an alternative approach to viscosity in Fermi liquids see Ref. [225]).

Beyond graphene, in anisotropic Dirac systems [183,212] one has to consider the full viscosity tensor (these are the systems where two Dirac cones merge in momentum space [226]; this may be relevant to the organic conductor α-(BEDT-TTF)$_2$I$_3$ under pressure [227], the heterostructure of the 5/3 TiO$_2$/VO$_2$ supercell [228,229], surface modes of topological crystalline insulators with unpinned surface Dirac cones [230], and quadratic double Weyl fermions [231]). In the absence of magnetic field, the viscosity matrix contains six independent components (in accordance with the Onsager reciprocity [13, 16]), which scale differently with temperature [183]. In particular, one of the six components vanishes at lowest temperatures violating the famous (conjectured) bound for the shear viscosity to entropy density ratio [232]. As a result, the authors of Ref. [183] proposed a generalization of the bound to anisotropic 2D systems, see Sec. 6. An alternative view on anisotropic Dirac semimetals taking into account spectrum topology (i.e., the Berry curvature) has been developed in Ref. [233]. Hall viscosity in the quantum Hall regime in such systems was discussed in Ref. [234]. More complicated spectra can be encountered in 3D Luttinger semimetals [235] where the long-screened nature of the Coulomb interaction leads to a scale-invariant, non-Fermi-liquid ground state [236]. Hydrodynamic behavior in such systems was considered in Ref. [237].

3.2.6 Hydrodynamic equations in graphene

Taking into account the dissipative corrections in the continuity equations (26), one finds the generalization of the Navier-Stokes equation [13, 90, 91] in graphene

$$\mathcal{W}(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} + v_g^2\boldsymbol{\nabla}P + \boldsymbol{u}\partial_t P + e(\boldsymbol{E}\cdot\boldsymbol{j})\boldsymbol{u} = \tag{43a}$$

$$= v_g^2\left[\eta\varDelta\boldsymbol{u} - \eta_H\varDelta\boldsymbol{u}\times\boldsymbol{e_B} + en\boldsymbol{E} + \frac{e}{c}\boldsymbol{j}\times\boldsymbol{B}\right] - \frac{\mathcal{W}\boldsymbol{u}}{\tau_{\mathrm{dis}}}.$$

The full set of the hydrodynamic equations contains also the continuity equations

$$\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0, \tag{43b}$$

and

$$\partial_t n_I + \boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\frac{n_I - n_{I,0}}{\tau_R}, \tag{43c}$$

and the thermal transport equation [201]

$$T\left[\frac{\partial s}{\partial t} + \boldsymbol{\nabla}\cdot\left(s\boldsymbol{u} - \delta\boldsymbol{j}\frac{\mu}{T} - \delta\boldsymbol{j}_I\frac{\mu_I}{T}\right)\right] = \delta\boldsymbol{j}\cdot\left[e\boldsymbol{E} + \frac{e}{c}\boldsymbol{u}\times\boldsymbol{B} - T\boldsymbol{\nabla}\frac{\mu}{T}\right] -$$

$$-T\delta\boldsymbol{j}_I\cdot\boldsymbol{\nabla}\frac{\mu_I}{T} + \frac{\eta}{2}\left(\nabla^\alpha u^\beta + \nabla^\beta u^\alpha - \delta^{\alpha\beta}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right)^2 -$$

$$-\frac{n_E - n_{E,0}}{\tau_{RE}} + \mu_I\frac{n_I - n_{I,0}}{\tau_R} + \frac{\mathcal{W}\boldsymbol{u}^2}{v_g^2\tau_{\mathrm{dis}}}, \tag{43d}$$

where s denotes the entropy density. The equation (43d) replaces the continuity equation for the energy density (26f) as is common in hydrodynamics [13]. The hydrodynamic equations are supplemented by the constitutive equations for the quasiparticle currents (29) and the generalized conductivity matrix $\widehat{\varSigma}$, as well as Maxwell's equations for the electromagnetic field, in other words, Vlasov self-consistency [16, 23, 120, 121].

3.2.7 Boundary conditions

The state of a conventional fluid is described by the velocity vector and two thermodynamic quantities, such as density and pressure. The hydrodynamic equations are differential equations containing spatial and time derivatives of these variables. Hence, to find a solution to these equations one has to specify the boundary conditions.

The conventional Navier-Stokes equation [13, 90, 91] greatly simplifies for an incompressible fluid. In this case, the fluid density is a constant, while the pressure gradient can be excluded by applying the curl operation to the equation. The resulting equation is a differential equation for the velocity only.

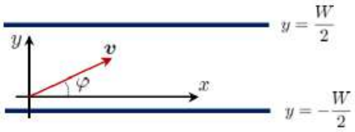


Fig. 19 Channel geometry: the electronic fluid is confined to a channel (along the x-direction) of the width W; $\boldsymbol{v}$ is the quasiparticle velocity directed at the angle φ to the channel boundary.

If a viscous fluid is flowing near a solid, stationary boundary, a simple "no-slip" boundary condition is often assumed [13] (due to the molecular forces acting between the fluid and the boundary). On the other hand, a boundary between a fluid and a gas can be characterized by the "no-stress" boundary condition, where the tangential stress is continuous at the interface. The two conditions can be "unified" as limiting cases of a more general condition due to Maxwell [238]

$$u_t^\alpha \Big|_S = \ell_S \, e_n^\beta \frac{\partial u_t^\alpha}{\partial x^\beta}\Big|_S, \tag{44}$$

where $\boldsymbol{e}_n$ is the unit vector normal to the surface, $\boldsymbol{u}_t = \boldsymbol{u} - (\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n}$ is the tangential velocity, and ℓ_S is the so-called "slip length". The no-slip boundary condition, $\boldsymbol{u} = 0$ (the normal component of the velocity has to vanish at any solid boundary by obvious reasons) corresponds to $\ell_S = 0$, while the limit $\ell_S \to \infty$ describes the no-stress case.

In electronic systems, the boundary condition (44) was studied in detail in Ref. [117] based on the kinetic approach. Solving the kinetic equation in the presence of a boundary requires boundary conditions for the distribution function. The latter are well studied [239], especially in the context of mesoscopic physics [11]. Analytic calculations are possible in the two limiting cases of specular and diffusive scattering at the boundary. Boundary conditions in the presence of magnetic field were studied in Ref. [240]. Recently, the issue of the boundary conditions and the slip length in the magnetic field was discussed in Ref. [241].

Specular scattering refers to ideally smooth boundaries such that the incidence and reflection angles (of the quasiparticle velocity) coincide. In that case, the distribution function obeys the simple boundary condition

$$f(\varphi)\Big|_S = f(-\varphi)\Big|_S, \tag{45}$$

where φ is the angle between the quasiparticle (microscopic) velocity $\boldsymbol{v}$ and the boundary. Experimental feasibility of smooth boundaries was recently explored in Ref. [242].

In the diffusive case, the boundary is assumed to be sufficiently rough, such that the incoming quasiparticle can scatter off the boundary in any direction with equal probability (independent of the incidence angle). This can be expressed by a more complex condition. In a channel geometry (see Fig. 19) the corresponding condition has the form [11]

$$f(W/2,-\pi<\varphi<0) \;=\frac{1}{2}\int\limits_0^\pi d\varphi'\sin\varphi' f(W/2,\varphi'), \tag{46a}$$

$$f(-W/2,0<\varphi<\pi)|_S=\frac{1}{2}\int\limits_{-\pi}^0 d\varphi'\sin\varphi' f(-W/2,\varphi'). \tag{46b}$$

The resulting slip length is strongly influenced by the choice of the boundary conditions for the distribution function [117]. The authors of Ref. [117] express ℓ_S in terms of the electron-electron scattering length

$$\ell_S=g(\kappa)\ell_{ee},\qquad \kappa=\frac{h_1^2h_2^{d-1}}{\lambda^{d+1}},\qquad g(\kappa)\to\begin{cases} g_0/\kappa, & \kappa\ll 1,\\ g_\infty, & \kappa\to\infty,\end{cases} \tag{47}$$

where h_1 and h_2 are the mean height and correlation length describing the boundary roughness [239], λ is the (temperature dependent) electron wavelength, and d is the spatial dimensionality. The precise value of $g(\kappa)$ varies dramatically, but at experimentally relevant temperatures one finds $\ell_S\approx 0.5\,\mu$m, the value that agrees with experimental observations, see Ref. [51].

Full solution to the hydrodynamic equations in electronic systems requires also boundary conditions for thermodynamic quantities. In electronic systems, these are most conveniently expressed in terms of electrochemical potentials.

Traditional transport theory is based on a single-electron approach, where the main mechanism of electrical resistance – and hence, dissipation – is the electron-impurity and electron-phonon scattering. In this case, the bulk system is characterized by a local conductivity, while contact interfaces – by the contact resistance. The latter appears due to equilibration of (originally mismatched) electrochemical potentials in the two interfacing materials [243]. The bulk and contact resistances could be seen as independent parts of the overall electrical circuit. If the bulk system is diffusive, the contribution of the contacts is typically negligible. On the contrary, in ballistic systems there is almost no dissipation in the bulk, such that most of the voltage drop occurring in the contacts, see Fig. 9.

In the context of ideal (inviscid) hydrodynamics in nearly neutral graphene, boundary conditions taking into account contact resistance were considered in Ref. [89]. Assuming the leads are represented by a disordered, particle-hole symmetric metal, the electron and hole currents are given by the difference of the electrochemical potentials across the interface divided by the contact resistance. If no electric current is allowed in the system (as is appropriate for measurements of thermal conductivity [24,89]), this leads to a boundary condition relating the imbalance chemical potential μ_I and current $\boldsymbol{j}_I$.

An alternative situation was considered in Ref. [138]. In this paper the authors have considered an idealized situation where a clean (disorder-free), but viscous electron fluid is contacted by an ideal conductor with an ideal interface characterized by the vanishing reflection coefficient [244]. The absence of disorder implies the lack of Ohmic dissipation in the bulk, while the ideal contacts do not provide any contact resistance. In that case the bulk dissipation due to viscosity has to be compensated by the work done by current source. If both the bulk and the contacts are disorder-free, then the electric potential exhibits a sharp inhomogeneity (on the hydrodynamic scale - a jump) in a narrow region close to the interface, which translates into a viscosity-dependent contribution to the contact resistance that can be positive or negative depending on the contact curvature sign.

Real samples are likely to exhibit all of the above effects and moreover may host additional localized charges at the sample edges leading to classical (nontopological) edge currents [53], see Sec. 2.4.3. The appropriate boundary conditions then strongly depend on sample geometry and the specific measurement scheme. For example, the authors of Ref. [140] suggest using the Corbino disk geometry to measure electronic viscosity. In their setup, the outer edge of the Corbino disk is isolated, implying the vanishing radial component of the electric current. In addition, they required the azimuthal momentum component to diffuse radially, such that the off-diagonal component of the viscous stress tensor vanishes at both edges of the disk. Interestingly enough, the authors of Ref. [140] considered the no-slip boundary conditions as well and found no qualitative difference with the above approach.

3.3 Hydrodynamic collective modes and plasmons

Hydrodynamic collective modes have been considered by many authors [24, 121, 122, 169, 170, 171, 172, 173, 174, 175, 176, 245, 246]. The point of consensus is that the ideal (neglecting dissipative processes) electronic fluid in neutral graphene is characterized by a sound-like collective mode (sometimes referred to as the "cosmic sound" [169] or the "second sound" [174]) with the linear dispersion relation

$$\omega = v_g q/\sqrt{2}. \tag{48}$$

In a way, this result justifies the claim that the electronic fluid behave hydrodynamically, see Sec. 1.

Dissipative processes damp the sound mode (48). In contrast to traditional hydrodynamics this happens since dissipation due to "external" scattering (e.g., disorder and electron-phonon scattering) appears already in the description of an "ideal" (i.e., inviscid) electronic fluid, see Eqs. (26d), (26f), and (28). Another issue is the regime of applicability of the dispersion relation (48) or its damped counterparts. The point is that hydrodynamics is based on the gradient expansion valid at length scales that are much larger than ℓ_{ee} (representing the energy and momentum conserving interaction responsible for equilibration). At smaller length scales other, more conventional collective excitations, such as plasmons [121, 170, 171, 173, 174, 175, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262], may be identified.

3.3.1 Electronic "sound" in neutral graphene

Collective excitations in the electronic system in graphene have been recently studied in detail in Ref. [176]. At charge neutrality and in the absence of magnetic field, the sound mode (48) damped by the dissipative processes has the dispersion relation

$$\omega = \sqrt{\frac{v_g^2 q^2}{2} - \frac{1}{4}\left(\frac{1+q^2\ell_G^2}{\tau_{\rm dis}} - \frac{1}{\tau_{RE}}\right)^2} - i\frac{1+q^2\ell_G^2}{2\tau_{\rm dis}} - \frac{i}{2\tau_{RE}}, \tag{49}$$

where ℓ_G is the Gurzhi length (8). Although Eq. (49) can be straightforwardly derived by linearizing the hydrodynamic equations (43), the damping in Eq. (49) can be seen as exceeding the accuracy of the hydrodynamic regime. Indeed, the gradient expansion in neutral graphene is justified for momenta smaller than a certain scale defined by the electron-electron interaction

$$q\ell_{\rm hydro} \ll 1, \quad \ell_{\rm hydro} \sim \frac{v_g}{\alpha_g^2 T}. \tag{50}$$

Assuming a clean system $\tau_{\rm dis} \to \infty$ (energy relaxation due to supercollisions [201] may be also neglected, $\tau_{RE} \gg \tau_{\rm dis}$), the expression under the square root in Eq. (49) can be expanded for small q as

$$\frac{v_g^2 q^2}{2} - \frac{\left(1+q^2\ell_G^2\right)^2}{4\tau_{\rm dis}^2} \to \frac{v_g^2 q^2}{2}\left[1 - Aq^2\ell_{\rm hydro}^2 - \mathcal{O}(\tau_{\rm dis}^{-1})\right],$$

where A is a numerical coefficient. Hence, within the hydrodynamic approach, the viscous contribution to damping should be neglected, leaving one with the simpler dispersion [121]

$$\omega = \sqrt{\frac{v_g^2 q^2}{2} - \frac{1}{4\tau_{\rm dis}^2}} - \frac{i}{2\tau_{\rm dis}}. \tag{51}$$

Now, the peculiar nature of the Dirac spectrum in graphene leads to the fact that the linearized version of the hydrodynamic equations is justified in a wider parameter region than Eqs. (43) themselves [121,124,192,206] (due to the "collinear scattering singularity" [23,24,121,153]). In the weak coupling limit, the linear response theory is valid at much larger momenta

$$q\ell_{\rm coll} \ll 1, \quad \ell_{\rm coll} \sim \frac{v_g}{\alpha_g^2 T|\ln\alpha_g|} \ll \ell_{\rm hydro}, \tag{52}$$

formally providing one with a justification to extend Eq. (49) beyond the hydrodynamic regime. However, already at $q\ell_{\rm hydro} \sim 1$ the imaginary part of the sound dispersion becomes comparable to the real part, at which point the dispersion is no longer observable.

The nature of the sound mode (48) [or Eq. (49)] becomes clear if one takes into account the fact that in neutral graphene in the absence of magnetic field the electric charge is decoupled from the hydrodynamic energy flow. Indeed, at charge neutrality $n = 0$ so that the electric field does not enter the linearized Navier-Stokes equation (43a), while the "conductivity matrix" in Eqs. (29) is diagonal. Hence, the energy flow is described by the Navier-Stokes equation (43a), while charge transport is described by the Ohmic relation (29b), together with the Vlasov self-consistency. The latter can be expressed using the Poisson's equation

$$\boldsymbol{E}_V = -e\boldsymbol{\nabla}\int d^2r' \frac{\delta n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{53a}$$

In gated structures [87,263], this can be simplified to

$$\boldsymbol{E}_V = -\frac{e}{C}\boldsymbol{\nabla}\delta n(\boldsymbol{r}), \tag{53b}$$

where $C = \varepsilon/(4\pi d)$ is the gate-to-channel capacitance per unit area, d is the distance to the gate, and ε is the dielectric constant. This approximation neglects the long-ranged (dipole-type) part of the screened Coulomb interaction and is justified while the charge density $n(\boldsymbol{r})$ varies on length scales exceeding d.

The charge sector of the theory is characterized by an overdamped collective mode with the dispersion

$$\omega = -iD_0q^2\left[1+eV_s(q)\frac{\partial n}{\partial \mu}\right], \qquad D_0 = \frac{1}{2}\frac{v_g^2\tau_{11}\tau_{\text{dis}}}{\tau_{11}+\tau_{\text{dis}}}. \tag{54}$$

In a gated structure, the mode is diffusive (with the Vlasov self-consistent potential $V_s = e/C$ providing a correction to the diffusion coefficient). For long-range Coulomb interaction (here $V_s = 2\pi e/q$), the dispersion remains purely imaginary with $\omega \sim iq$ at small q.

3.3.2 Electronic "sound" in doped graphene

In doped graphene, the charge and energy modes are coupled by the Vlasov self-consistency [176]. To the leading order in (weak) energy relaxation this leads to a sound mode similar to Eq. (49) and a diffusive mode that in a gated structure has the dispersion

$$\omega = -\frac{i}{\tau_{RE}}\frac{\varkappa v_g^2q^2}{(\varkappa+2\pi C)v_g^2q^2+4\pi C\tau_{RE}^{-1}\tau_{\text{dis}}^{-1}}, \tag{55}$$

where the Thomas-Fermi screening length is given by

$$\varkappa = N\alpha_g k_F = Ne^2\mu/v_g^2. \tag{56}$$

For long-range Coulomb interaction, the factor $2\pi C$ should be replaced with q. Physically, the mode (55) describes energy diffusion appearing due coupling of the charge and energy fluctuations by Vlasov self-consistency.

For a gated structure, the sound mode coincides with the "cosmic sound" (48) at the lowest momenta, albeit with the sound velocity modified by screening. In the case of long-range Coulomb interaction the dispersion is no longer sound-like. In the limit $q \to 0$ (and $\mu \gg T$), one finds the spectrum similar to the usual 2D plasmon [12,121]

$$\omega(q \ll \varkappa) = -\frac{i}{2\tau_{\text{dis}}} + \sqrt{\frac{1}{2} v_g^2 q \varkappa - \frac{1}{4\tau_{\text{dis}}^2}}. \tag{57}$$

The expression (57) is valid when

$$q\ell_G \ll 1, \quad q \ll \varkappa, \quad v_g^2 \varkappa q \tau_{\text{dis}}^2 \gg 1.$$

These conditions are consistent with the applicability condition of the hydrodynamic approach if

$$v_g \varkappa \tau_{\text{dis}} \gg 1 \quad \Rightarrow \quad N\alpha_g \mu \tau_{\text{dis}} \gg 1,$$
$$\ell_G \ll v_g^2 \varkappa \tau_{\text{dis}}^2 \quad \Rightarrow \quad N^2 \alpha_g^4 \mu \tau_{\text{dis}} (\bar{T} \tau_{\text{dis}})^2 \gg 1.$$

The above conditions provide a possibility to observe the dispersion (57) in a parametrically defined range of wavevectors.

3.3.3 Hydrodynamic modes and plasmons

The above sound-like modes have to be distinguished from plasmonic excitations in electronic systems. The latter are well studied, also in graphene [121, 170,173,174,175,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261, 262]. In a degenerate electron gas in 2D, the plasmon dispersion (neglecting impurity scattering, i.e., $\tau_{\text{dis}} \to \infty$) has the form [247]

$$\omega = \sqrt{2e^2 \mu q} \left(1 + \gamma \frac{q}{\varkappa}\right), \tag{58}$$

where γ is a numerical coefficient, that can be evaluated either within the random phase approximation (i.e., by computing the Lindhard function; this leads to $\gamma = 3/4$ [247]), or using a macroscopic (hydrodynamic-like) theory. The latter approach yields a different value of γ which is typically attributed to the fact that hydrodynamics is applicable at small momenta ($q\ell_{\text{hydro}} \ll 1$) and frequencies, while plasmons are nonequilibrium excitations that belong to higher momenta [247]. Based on this argument one might expect that the hydrodynamic collective modes and plasmons simply have nothing to do with each other [171]. Yet, given the same leading momentum dependence in Eqs. (57) and (58), the relation between the two is worth investigating.

In graphene, the possibility of discussing momenta exceeding $1/\ell_{\text{hydro}}$ is afforded by the collinear scattering singularity [23,24,120,121,122,124,152,153, 192] which leads to the existence of two parametrically different length scales, see Eq. (52), and hence of an intermediate momentum range, $\ell_{\text{hydro}}^{-1} \ll q \ll \ell_{\text{coll}}^{-1}$. Here a linear response theory of Ref. [124] can be used to find the collective modes. Remarkably, macroscopic equations of this theory coincide with the linearized hydrodynamic equations [176] such that the resulting dispersions should be valid in the hydrodynamic regime as well and can be compared with the above results.

In doped graphene, the electron system is degenerate and the linear response theory of Ref. [124] can be expressed in terms of a single equation

$$\frac{\partial \boldsymbol{J}}{\partial t}+\frac{v_g^2}{2}\boldsymbol{\nabla}\rho-\nu\Delta\boldsymbol{J}-\frac{v_g^2}{2}\frac{\partial n}{\partial\mu}e^2\boldsymbol{E}=-\frac{\boldsymbol{J}}{\tau_{\rm dis}}, \tag{59}$$

where $\boldsymbol{J}$ is the electric current, see Eq. (1), and ρ denotes the charge density. Taking into account the Vlasov field (53) and continuity equation, one finds the collective mode with the spectrum

$$\omega=\sqrt{2e^2\mu q\left(1+\frac{q}{\varkappa}\right)-\frac{(1+q^2\ell_G^2)^2}{4\tau_{\rm dis}^2}}-\frac{i(1+q^2\ell_G^2)}{2\tau_{\rm dis}}, \tag{60}$$

where $D=v_g^2\tau_{\rm dis}/2$ and $\sigma=v_g^2(\partial n/\partial\mu)\tau_{\rm dis}/2$ are the diffusion coefficient and the Drude conductivity.

The spectrum (60) is exactly the same as the screened sound mode leading to Eq. (57). In the limit $\tau_{\rm dis}\to\infty$, one may expand Eq. (60) in small $q\to 0$. This yields Eq. (58) with the "wrong" coefficient, $\gamma=1/2$. At the same time, the leading term (at $q\ll\varkappa$) agrees with the Fermi liquid result in the presence of disorder [12] (in the absence of viscosity). The dispersion (60) is valid for $q\ell_{\rm coll}\ll 1$, however, becomes overdamped already at $q\sim\ell_{\rm hydro}^{-1}$. For $q\gg\ell_{\rm coll}^{-1}$, the quasi-equilibrium description leading to Eq. (59) breaks down and true plasmons with the dispersion (58) emerge. At these momenta the spectrum (60) is purely imaginary. Based on this argument, the authors of Ref. [176] argue that the two modes are not connected. Similar conclusions were reached in Ref. [173], where it was argued that Coulomb interaction precludes the appearance of hydrodynamic sound in Fermi liquids.

In graphene at charge neutrality, the "true" plasmon dispersion was established in Ref. [122] on the basis of microscopic theory. The leading behavior of the plasmon dispersion is given by

$$\omega=\sqrt{(4\ln 2)e^2Tq}. \tag{61}$$

This expression can be compared to the results of the linear response theory in graphene [124,176]. The linear response theory of Ref. [124] is based on the same three-mode approximation as the hydrodynamics discussed in Sec. 3.2. Similarly to the discussion in Sec. 3.3.1, at charge neutrality the charge sector decouples from the rest of the theory and can be described by the equation

$$\frac{\partial \boldsymbol{j}}{\partial t}+\frac{v_g^2}{2}\boldsymbol{\nabla}n-\frac{2\ln 2}{\pi}e^2T\boldsymbol{E}=-\frac{\boldsymbol{j}}{\tau_{\rm dis}}-\frac{\boldsymbol{j}}{\tau_{11}}, \tag{62}$$

where τ_{11} determines the quantum conductivity (32), see also Eqs. (29) and (30). Combining Eq. (62) with the continuity equation one finds

$$\omega^2 + i\omega\left(\frac{1}{\tau_{\mathrm{dis}}} + \frac{1}{\tau_{11}}\right) = \frac{v_g^2}{2}q^2 + (4\ln 2)e^2Tq, \tag{63a}$$

leading to a plasmon-like spectrum that can be expressed similarly to Eq. (54)

$$\omega = -i\frac{\sigma(\omega)q^2}{e^2\partial n/\partial\mu}\left[1 + eV_s(q)\frac{\partial n}{\partial\mu}\right], \tag{63b}$$

where $\sigma(\omega)$ is the optical conductivity [206] [in contrast to the static conductivity (32) in Eq. (54]

$$\sigma(\omega) = \frac{2e^2T\ln 2}{\pi}\frac{1}{-i\omega + \tau_{11}^{-1} + \tau_{\mathrm{dis}}^{-1}}. \tag{63c}$$

In the hydrodynamic regime of small frequencies, $\sigma(\omega \to 0) \to \sigma_0$, the mode (63b) is purely diffusive recovering Eq. (54).

Resolving Eq. (63a) one finds the plasmon dispersion in the form

$$\omega = -i\frac{\tau_{\mathrm{dis}}+\tau_{11}}{2\tau_{\mathrm{dis}}\tau_{11}} + \sqrt{(4\ln 2)e^2Tq + \frac{v_g^2}{2}q^2 - \frac{(\tau_{\mathrm{dis}}+\tau_{11})^2}{4\tau_{\mathrm{dis}}^2\tau_{11}^2}}\,. \tag{63d}$$

For $\omega \gg \tau_{11}^{-1} \gg \tau_{\mathrm{dis}}^{-1}$ and $q \to 0$, the leading behavior in Eq. (63d) coincides with Eq. (61). At large momenta the first term in the RHS of Eq. (63a) dominates and the dispersion resembles the hydrodynamic sound, Eq. (48). This contradicts the results of Ref. [122]: although at large q the true dispersion also becomes linear, the coefficient (analogous to the speed of sound) is different (there is no factor of $\sqrt{2}$).

To summarize, the plasmon mode (63d) should be contrasted with the diffusive charge mode (54), and not the sound mode (49). The plasmon and the sound belong largely to different frequency regimes [171], but most importantly, stem from the two different, decoupled sectors of the theory (the sound mode can also be obtained from the linear response theory hence one can extend its region of applicability beyond the hydrodynamic regime). The latter fact is the reason why the plasmon dispersion is independent of viscosity, while the sound mode (49) is unaffected by screening effects (which are essentially responsible for plasmon excitations). Formally, the two modes coexist but are characterized by different frequencies that are much higher for the plasmon mode. Approximately at $q \sim \ell_{\mathrm{coll}}^{-1}$, i.e., at the applicability limit of the linear response theory, the sound mode becomes overdamped, which does not happen to the plasmon. At that point the plasmon dispersion is almost linear albeit with the coefficient that disagrees with the microscopic theory [122], as pointed out above.

An alternative approach to plasmons is to consider the electromagnetic response of the 2D electron fluid to the high-frequency field generated by a Hertzian dipole [264]. For small enough frequencies ($\omega\tau_{ee} \ll 1$) the electron system responds hydrodynamically. Coupling the hydrodynamic equations with the 3D Maxwell's equations one can define a boundary value problem yielding the full description of the spatial structure of the electromagnetic field. In particular, the numerical analysis of Ref. [264] suggests co-existence between the plasmon and diffusive modes in a way that is somewhat different from the above solution of the purely hydrodynamic problem (where the electromagnetic field was assumed to be static). For analytic analysis of edge magnetoplasmons (using the Wiener-Hopf technique) see Ref. [265].

4 Known solutions to hydrodynamic equations in electronic systems

Once equipped with the hydrodynamic equations and boundary conditions, one may embark on finding solutions in an attempt to either explain or predict experimental observations. Since most transport measurements in solids are performed within linear response, many authors consider solutions to linearized hydrodynamic equations.

Hydrodynamic charge flow in doped graphene (more generally, in hydrodynamic Fermi liquids) was considered analytically in Refs. [101,94,95,96,37, 109,110,207,266] and numerically in Refs. [93,97,98]. Neutral graphene (more generally, compensated semimetals) was analyzed in Refs. [85,87,114,115,125, 199,267].

Nonlocal transport properties observed in doped graphene [27,41,42] were studied in Refs. [94,95] focusing on the appearance of vortices (or "whirlpools") in viscous flows in confined geometries, the effect that is responsible for the observed negative nonlocal resistance [27]. A purely analytic approach to that problem (albeit in an idealized geometry) was offered in Refs. [96,207]. The authors of Ref. [96] hinted on the possibility to observe multiple vortices, the effect that was further explored numerically in Ref. [93] (see Fig. 5), where a *sign-alternating* nonlocal resistance was suggested as a consequence. The latter is especially important given that negative nonlocal resistance is not a unique characteristic of the viscous flow and can be observed in ballistic systems [41,100]. Interestingly enough, complicated patterns of multiple vortices may arise also in nearly neutral graphene with long-ranged disorder [267]. Further complications with the hydrodynamic interpretation of the observed nonlocal resistance and the associated vorticity were discussed in Ref. [268], where it was argued that nonlocal (i.e., momentum-dependent) conductivity in disordered electron systems may mimic the hydrodynamic effects even in the absence of electron-electron interaction [the idea is to interpret Eq. (2) as the Ohm's law with nonlocal conductivity]. However, extracting the viscosity from the nonlocal conductivity obtained by means of the Kubo formula [211] might not be straightforward in disordered systems [217]. Moreover, it is unclear why should one use the hydrodynamic "no-slip" boundary conditions [which are needed to obtain Poiseuille-like solutions from Eq. (2)] in conventional disordered systems outside of the hydrodynamic regime.

An alternative measurement providing indirect evidence of hydrodynamic behavior, namely superballistic transport through a point contact [39] was discussed theoretically in Refs. [37,101]. Reference [37] provided a detailed analysis of the hydrodynamic theory in the slit geometry comparing the results to those of the ballistic and diffusive (Ohmic) behavior. The authors of Ref. [37] concluded that the hydrodynamic regime represents a relatively narrow intermediate parameter region between the two more conventional regimes (namely, the diffusive and ballistic). Further analysis of a viscous flow through a constriction and the related enhancement of conductivity was reported in Ref. [269].

Now, one of the most popular geometries to consider hydrodynamic effects is the channel (or slab) geometry, see Figs. 9 and 19. The reason for this is the wide spread of the Hall bar geometry of the experimental samples, see Figs. 2 and 4, as well as simplicity of theoretical solution, since assuming a long channel all physical quantities depend only on the coordinate along the channel (the x-coordinate in the notations adopted in Fig. 19). Assuming the no-slip boundary conditions, one finds the solution to the Navier-Stokes equation in the form of the catenary curve, which reduces to the standard Poiseuille flow [13,116,270] in the limit of the large Gurzhi length, $\ell_G \gg W$ (where W is the channel width).

In doped graphene, the electric current is hydrodynamic and is expected to exhibit this behavior [94], with

$$J_x = \sigma_0 E_x \left[1 - \frac{\cosh y/\ell_G}{\cosh W/[2\ell_G]} \right], \tag{64}$$

where J_x and E_x are the components of the current density and electric field along the channel and σ_0 is the Drude conductivity (due to, e.g., disorder). This effect was later observed in the imaging experiment of Ref. [51]. If the system is subjected to the magnetic field, then increasing the field decreases the viscosity, see Eqs. (40), and hence the Gurzhi length (8) leading to *negative* magnetoresistance (suggested theoretically in Refs. [109,110] and observed experimentally in Ref. [42]). These effects were also considered within the two-fluid hydrodynamic model in Ref. [114]. For an alternative theory of the electronic flows in narrow channels in magnetic fields describing the interplay of electron-electron interactions, disorder, and boundary conditions that goes beyond the hydrodynamic description, see Ref. [240]. For a detailed discussion of the Hall voltage and more generally the role of Hall viscosity in 2D Fermi liquids see Ref. [271]. The case of long-range disorder (or general inhomogeneity of the medium) was considered in Refs. [266,272,273], where a positive bulk magnetoresistance was found due to the absence of the Hall voltage [273]. The latter point is reminiscent of the situation in graphene at charge neutrality (other than the boundary effects).

In neutral graphene, the picture is more complicated due to decoupling of the charge and energy flows in the absence of magnetic field. In that case, the

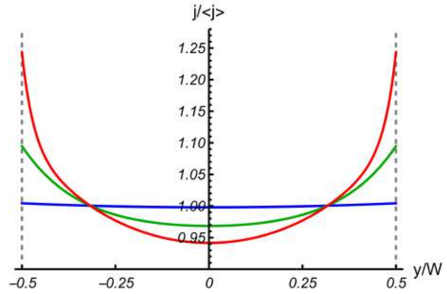


Fig. 20 Anti-Poiseuille flow in narrow channels in graphene in perpendicular magnetic field [125]. The curves represent the inhomogeneous current density in narrow channels of width $W = 0.1$, 1, $5\,\mu$m (blue, green, and red curves, respectively). Calculations wee performed for typical parameter values $\tau_{\rm dis} \approx 0.8\,$THz [43], $\alpha_g \approx 0.2$ [43, 178], $\nu \approx 0.4\,{\rm m}^2/{\rm s}$ [50, 119], $B = 0.1\,$T, $T = 250\,$K (Reprinted with permission from Ref. [125]. Copyright (2021) by the American Physical Society).

hydrodynamic, Poiseuille-like flow is expected for the energy current [183], while the charge transport exhibits the usual diffusion with the quantum conductivity (32) due to electron-electron interaction instead of the standard Drude conductivity due to disorder. Applying external magnetic field naively leads to a positive, parabolic magnetoresistance. This is because the bulk electric current in neutral graphene is accompanied by the lateral quasiparticle (and energy) current (which in turn leads to the geometric magnetoresistance). However, due to the compensated Hall effect and quasiparticle recombination, see Sec. 3.2.1, there is a strong boundary effect changing that behavior and leading to nonsaturating, linear magnetoresistance (at charge neutrality) [87] that is somewhat similar to the edge effects considered in Ref. [86, 274]. The key point is that the above bulk effect is incompatible with finite size geometry: assuming that the bulk current is flowing along the channel, the lateral quasiparticle current must flow across the channel and hence must vanish at both boundaries. The resulting inhomogeneity of the individual electron and hole currents is inconsistent with the standard geometric magnetoresistance. Moreover, this inhomogeneity is only compatible with the continuity equation for the total quasiparticle density, Eq. (26d), if one takes into account recombination. The resulting quasiparticle density is practically uniform in the bulk (characterized by the parabolic geometrical magnetoresistance), but is strongly inhomogeneous in boundary regions of the width of the recombination length, $\ell_R(B) = \ell_R(B=0)/\sqrt{1+\mu^2B^2}$ (here μ stands for carrier mobility). The edge contribution to the overall resistance is linear in magnetic field [87] and can dominate in classically strong fields. This effect is not specific to Dirac fermions. Theoretically similar phenomena were considered in Refs. [114, 115, 199]. Experimentally, linear magnetoresistance due to recombination was studied in bilayer graphene in Ref. [88].

Electric current in a neutral graphene channel also becomes inhomogeneous in magnetic field (where all three modes in the “three mode approximation” discussed in Sec. 3 are coupled). However, unlike the situation in doped graphene, the current does not exhibit the Poiseuille-like flow (64) [125], see Fig. 20. One of the reasons for that is the boundary conditions: the Poiseuille flow is the solution of the hydrodynamic equations with the no-slip boundary conditions (which can be generalized to the Maxwell's boundary conditions with a relatively small slip length). The electric current in neutral graphene is not related to any solution of the Navier-Stokes equation and, moreover, there is no reason to assume that the current vanishes at the channel boundaries. In fact, for specular boundary conditions the opposite happens [125]: quasiparticle recombination leads to a minimum of the current density in the center of the channel, while the maximum value occurs at the boundaries. More general boundary conditions (see Sec. 3.2.7) require a numerical solution of the kinetic equation, which has not yet been carried out in this context.

An alternative geometry to study hydrodynamic flows is offered by the Corbino disk [139]. Here the electric current is inhomogeneous even in the simplest case of the Ohmic flow in the absence of magnetic field ($\boldsymbol{j} \propto (1/r)\boldsymbol{e}_r$, where $\boldsymbol{e}_r$ is the unit vector in the radial direction). Applying an external magnetic field that is orthogonal to the disk one can induce an azimuthal, nondissipative Hall current (that is not compensated by the Hall voltage due to the absence of boundaries). The resulting inhomogeneous flows represent an excellent opportunity to study viscous effects [140]. The Corbino disk with specular boundaries was analyzed in Ref. [214]. Assuming small momentum relaxation, the authors of Ref. [214] concluded that the Hall angle (that can be determined by the ratio of the azimuthal and radial components of the current) is directly related to the ratio of the Hall and shear viscosities such that the resistive Hall angle approaches the viscous Hall angle. Anomalous thermoelectric response (i.e., violating the Matthiessen's rule, Wiedemann-Franz law, and Mott relation) exhibited by hydrodynamic flows in the absence of Galilean invariance was reported in Ref. [275].

Recently, the Corbino geometry was used to demonstrate the "superballistic conduction" both experimentally [54] and theoretically [276, 277]. Both theories focused on the boundary effects. Reference [277] analyzed the radial electric current (in the absence of magnetic field). In the hydrodynamic regime, the interface between the lead (assumed to be a perfect conductor) and the Corbino disk is characterized by the finite Knudsen layer [138] with the boundary conductance that can exceed the Sharvin conductance [104]. Ref. [276] came to similar conclusions arguing that if the number of conducting channels varies along the current flow (using either a wormhole or Corbino geometries as examples), the Landauer-Sharvin resistance is detached from the leads and is spread throughout the bulk of the system. If the length scale characterizing the spread is larger than ℓ_{ee} then the resistance is reduced leading to superballistic conductance.

More complicated flow patterns can be achieved by considering curved boundaries or adding artificial obstacles to engineer boundary conditions [97, 98]. In particular, on the basis of numerical analysis it was shown [98] that additional barriers on the channel walls may lead to the effective "no-slip" boundary conditions that are commonly assumed in theoretical calculations.

5 Nonlinear phenomena in electronic hydrodynamics

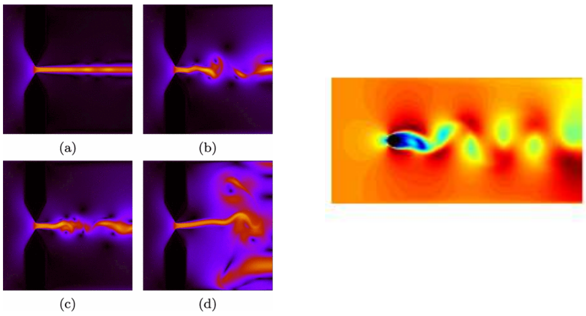
Nonlinear hydrodynamic effects in electronic systems remain largely unexplored both theoretically and experimentally. Early numerical work [278] suggests that electron flows with the high enough Reynolds numbers (for samples of the size of $5\,\mu$m and macroscopic speeds $u \sim 10^5$ m/s [279], the authors of Ref. [278] estimate $\mathrm{Re} \sim 100$) may exhibit pre-turbulent phenomena such as vortex shedding.

A representative example of nonlinear phenomena in graphene — hot spot relaxation — was considered in Ref. [121]. A hot spot is a particular non-equilibrium state of the system that is characterized by a locally elevated energy density. This state can be prepared with the help of a local probe or focused laser radiation [250, 251]. As expected [250, 251], the hot spot loses energy by emitting plasmon-like waves. At charge neutrality, these are in fact acoustic energy waves analogous to the long-wavelength oscillations in interacting systems of relativistic particles [sometimes called the "cosmic sound", see Eq. (48)]. However, a nonzero excess energy remains at the hot spot due to compensation between the thermodynamic pressure and the self-consistent (Vlasov) electric field. Dissipation tends to destroy the thus achieved quasi-equilibrium, but the resulting decay is characterized by a longer time scale as compared to the initial emission of plasmons. At the same time, the plasmons appear to be damped by viscous effects, see Sec. 3.3.3. The plasmon emission can also be expected in the in the high-frequency regime, where it has been linked to the Cherenkov effect [280, 281, 282, 283].

The above quasi-equilibrium solution [121] may be viewed as an example of a soliton-like stationary nonlinear wave where charge and energy fluctuations (otherwise distinct at charge neutrality) are coupled by nonlinearity of the hydrodynamic theory. Away from charge neutrality, solitons were considered in the inviscid limit in Ref. [284] and more generally in Ref. [285]. In particular, the authors of Ref. [285] focused on hydrodynamic flows in graphene, where the decay of solitonic solutions was suggested as a possible experimental measure of electronic viscosity.

One of the most important consequences of nonlinearity of the hydrodynamic equations – turbulence [13] – is currently regarded as unlikely to occur in electronic systems, e.g., in graphene. In conventional fluids, turbulence can be reached when the Reynolds number characterizing the flow becomes large, $\mathrm{Re} \gtrsim 1000$ [13]. In contrast, typical Reynolds numbers characterizing existing experiments in graphene are rather small. Indeed, assuming one of the highest reported values of the drift velocity graphene, $u \sim 10^5$ m/s (based on the "saturation velocity measurements" [286]), the experimental estimate for the kinematic viscosity $\nu \sim 0.1\,\mathrm{m}^2/\mathrm{s}$ [27], and a typical sample size $L \sim 1\,\mu$m, one can estimate the Reynolds number as

$$\mathrm{Re} = \frac{uL}{\nu} \sim \frac{10^5\,\frac{\mathrm{m}}{\mathrm{s}} \times 10^{-6}\mathrm{m}}{0.1\,\frac{\mathrm{m}^2}{\mathrm{s}}} = 1.$$



(a) (b) (c) (d)

Fig. 21 Preturbulent hydrodynamic phenomena [278]. Left: microscale impurities in graphene can trigger coherent vorticity patterns that closely resemble classical 2D turbulence. The color represents the magnitude of the velocity. Calculations wee performed for Re = 25. Right: Vortex shedding in graphene at Re = 100 (Reprinted with permission from Ref. [278]. Copyright (2011) by the American Physical Society).

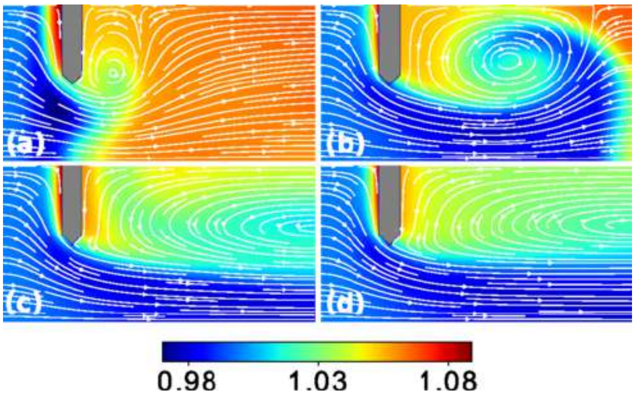


Fig. 22 Kelvin-Helmholz instability in graphene [288]. The color represents density fluctuations relative to the initial density. The streamlines show the direction of the hydrodynamic velocity. The gray object is the stationary obstacle. The four images are respective snapshots of the fluid motion taken at different times. Calculations wee performed for Re = 53 (Reprinted with permission from Ref. [288]. Copyright (2017) by the American Physical Society).

At such values of the Reynolds number, one may observe "pre-turbulent" phenomena, such as vortex shedding, as can be seen by solving the hydrodynamic equations numerically [278] (although at somewhat higher Re, see Fig. 21). For a possibility to achieve turbulence in electronic systems other than graphene, see Ref. [287].

Nonlinearity of the Navier-Stokes equation also leads to a number of known instabilities, arising in particular in systems with nontrivial boundary conditions [13]. One of these instabilities, the Kelvin-Helmholz instability [290, 291], was studied numerically in Ref. [288], see Fig. 22. In conventional fluids this effect (actually visible in the atmosphere as a specific cloud pattern, the "fluctus") occurs in the case of velocity shear within a continuous fluid or at the

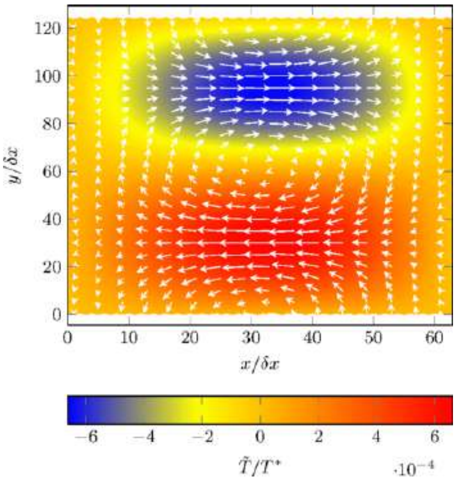


Fig. 23 Rayleigh-Bénard instability in graphene [289]. The color represents the temperature perturbation field with $T^* = 100\,\mathrm{K}$. The streamlines show the electron velocity. The image shows the formation of convection cells and the cosine-shaped temperature perturbation vanishing at the thermal contacts. (Reprinted with permission from Ref. [289]. Copyright (2015) by the American Physical Society).

interface between two fluids. In an electronic system this can be achieved by directing a charge flow through a macroscopic obstacle beyond which one observes vortex formation [288] that is reminiscent of the "whirlpools" that have been argued to be at the core of the nonlocal resistance experiments [27,93, 94,95,96]. Similarly, numerical simulations demonstrate the Rayleigh-Bénard instability [292,293], see Ref. [289] and Fig. 23. Note that the simulations of Ref. [288] were performed using a lattice Boltzmann method for relativistic gases. For more recent work on that method see Ref. [294].

In addition to the "conventional" instabilities of the hydrodynamic equations, there is another instability that is predicted to occur in a ballistic field effect transistor [22] or, in other words, in a gated 2D electron systems. There are two key observations leading to the appearance of this instability. Firstly, the carrier density in gated structures is determined by the same electric field (or voltage), see Eq. (53b), that represents the driving term in the Navier-Stokes equation (43a). In that case, the simplified Navier-Stokes equation (i.e., in the absence of magnetic field, neglecting Joule heating and weak disorder scattering) together with the continuity equation closely resemble the standard hydrodynamic equations for "shallow water" [13]. Secondly, one requires somewhat unusual (but experimentally feasible) boundary conditions: by connecting the source and drain of the device to a current source and the gate, while at the same time connecting the source to a voltage source, one arrives at the setup with a constant value of the voltage at the source together with the constant value of the current at the drain. In that case the wave velocities (shallow water waves in hydrodynamics or plasma waves in the heterostructure) describing propagation in the opposite directions are different leading to

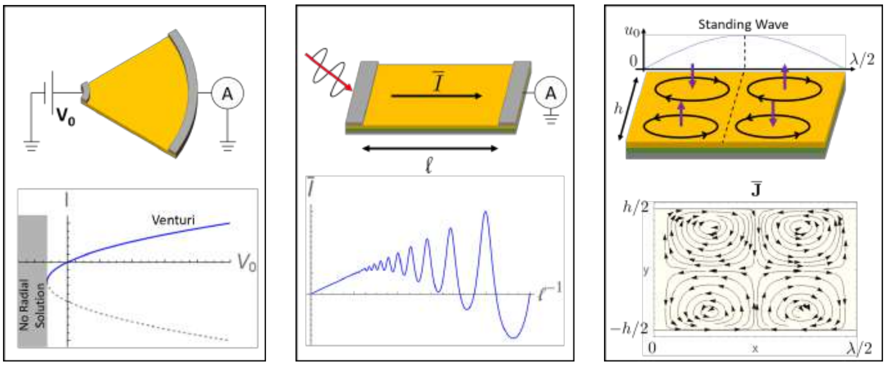


Fig. 24 Nonlinear hydrodynamic phenomena suggested in Ref. [300]. Left: the Venturi geometry and the expected nonlinear $I-V$ characteristic with $I \sim \sqrt{V}$ (the gray dashed line represents and unstable solution, while the gray area corresponds to the parameter regime of a possible instability towards turbulence). Center: Eckart streaming and the rectification effect. Right: Rayleigh streaming. (Reprinted with permission from Ref. [300]. Copyright (2021) by the American Physical Society).

the instability with respect to plasmon generation. Known as the "Dyakonov-Shur" instability, this effect has attracted considerable attention in literature, including that on hydrodynamic behavior in graphene [295, 296]; however, a definitive experimental observation of the effect is still lacking. For a detailed numerical analysis of a similar instability in GaAs MESFETs see Ref. [297]. An alternative suggestion for using viscous electrons as a source of terahertz radiation was proposed in Ref. [298]. Dyakonov-Shur instability in the Corbino geometry was discussed in Ref. [299].

Further nonlinear phenomena were discussed in Ref. [300] where three distinct hydrodynamic effects, namely the Bernoulli effect [301], Eckart streaming [302], and Rayleigh streaming [303], were suggested as possible experiments revealing nonlinear electron fluid dynamics, see Fig. 24. The suggested electronic analog of the Bernoulli effect yields a nonlinear term in the I-V characteristic ($V \propto I^2$) in the "Venturi geometry" (named after the Venturi tube, the standard device used for demonstrating the Bernoulli effect), which is essentially a finite-angle sector of the Corbino disk. The proposed effect is strongly dependent on sample geometry (e.g., it is expected to vanish in rectangular samples) and hence the boundary conditions. While the stationary Bernoulli effect is expected to occur in the ideal (inviscid) fluid, the dynamic nonlinear phenomena, such as the Eckart and Rayleigh streaming, are expected to occur in the presence of dissipation. Applying an oscillatory voltage to one of the sources while grounding the drain, the authors of Ref. [300] find a dc current (via the down-conversion). The two effects are distinguished by whether the dominant dissipation occurs in the bulk (Eckart streaming) or at the boundaries (Rayleigh streaming).

6 Theoretical conjectures of hydrodynamic behavior in strongly correlated systems

A (relatively) recent discovery of gauge-gravity duality (or AdS/CFT correspondence) [304,305,306] offers a new alternative theoretical tool to study strongly correlated systems by relating strongly coupled quantum field theories to gravity theories in one additional dimension. The best-known result of this approach is the conjectured lower bound for the shear viscosity to entropy density ratio [232] that has been found to be satisfied in quark-gluon plasma [307], cold atoms in the unitary limit [308], and intrinsic graphene [180]. The same physics can also be expressed in terms of the diffusivity bound [309]. Such bounds reflect not only the interaction strength, but also the symmetry properties of the system. In particular, in anisotropic systems the proposed bounds should be modified [183,310].

In the condensed matter context, the duality has also been applied to the by now perennial issue of the linear resistivity [311] in "strange metals" [312,313] (cuprates [314,315], iron-based superconductors [316,317,318,319], twisted bilayer graphene at magic angles [320,321], etc). The main premise of this approach is that excitations in strongly correlated systems are predominantly of the "collective" nature unlike the quasiparticles in conventional metals [306]. In that case, the system is described by "hydrodynamic-like" currents, with their relation to the external fields provided by the standard linear response theory. This way one can suggest universal bounds on the diffusion coefficient and conductivity (related by the Einstein relation) of a strange metal, as well as their scaling with temperature [310,309]. The concept of diffusion appears through a particular collective mode (the so-called quasinormal mode [322]). Such modes essentially replace quasiparticles in the qualitative interpretation of the resulting theory [306]. Linear response transport properties can then be obtained by means of either solving the hydrodynamic equations or using the memory matrix formalism [323]. The latter has the advantage of being independent of the concept of quasiparticles and extending beyond the hydrodynamic regime.

The holographic duality can also be used in the opposite direction, where solutions of hydrodynamic theories can provide insight into physical properties of gravitational objects [324].

While they might appear too abstract, the holographic methods can be put to test by studying the typical condensed matter experiment: optical pump-probe spectroscopy [325]. The idea is to test one of the characteristic predictions of the bulk (gravity) side of the duality – instantaneous thermalization [326]. This feature (impossible in the usual semiclassical description of transport) is the natural consequence of causality and is related to the "eigenstate thermalization hypothesis" [327,328]. As a result, measuring the optical conductivity in a strange metal excited by a short, intense laser pulse that does not contain a zero-frequency component one should obtain the exact same results as in the same system at equilibrium (characterized by the final temperature) at all times after the pulse.

The linear resistivity has also been interpreted as a signature of “Planckian dissipation” [329, 330] (which is also related to the above proposed bounds). The idea comes from the fact that the observed optical conductivity in strange metals often allows for a good fit with the standard Drude expression [311, 315, 331, 332] which is described by a timescale typically referred to as the “transport scattering time”, τ_{tr}, [10]. The linear temperature dependence of the resistivity thus translates into the τ_{tr} being inverse proportional to temperature or, in other words, proportional to the “Planckian” timescale

$$\tau_{\mathrm{tr}} \propto \tau_P = \frac{\hbar}{k_B T}, \tag{65}$$

where the Planck’s and Boltzmann’s constants ($\hbar$ and k_B) are restored for clarity. While completely natural in neutral graphene, see Eqs. (30), where the temperature dependence (65) follows already from dimensional analysis (in graphene at charge neutrality, T is the only energy scale), application of the concept of the scattering time to strongly correlated systems is more problematic. One possibility is that one can trace the decay of correlation functions (which can be characterized through a “transport” time scale) to the decay of local operators, as suggested in Ref. [330].

The hypothesis of the near-hydrodynamic behavior in strange metals (at least, at low temperatures where the measured optical conductivity has a Drude form) might sound attractive, but it certainly does not solve all the problems [329]. At higher temperatures, there appears the state of a “bad metal”, where the optical conductivity is no longer of the Drude form [333], while the temperature dependence of the resistivity is still linear. Quantum Monte Carlo simulations [334] suggest that this state is accompanied by hints of spin-stripe correlations [335]. While there might be a way to include that physics into holographic modeling [336], the role of electron-phonon coupling, quantum criticality, and their relation to the seemingly “universal” linear resistivity across several distinct families of materials remains to be understood.

The above ideas on applying holographic methods to strange metals (in particular in cuprates) remain controversial. For a recent critique of this approach, see Ref. [337].

A detailed discussion of relativistic hydrodynamics on the basis of the AdS/CFT correspondence was offered in Ref. [338]. In the 2D wire (channel) geometry with no-slip boundary conditions, this theory yields the Poiseuille behavior (see Sec. 2) for all velocities up to the ultrarelativistic limit $u \to v_g$. In the latter case, however, the differential resistance of the channel vanishes as a consequence of the kinematics of special relativity. The theory of Ref. [338] also offers further insights into the importance of the shear viscosity to entropy density ratio, η/s. Firstly, the channel resistance strongly scales with η/s, such that “holographic strongly coupled fluids” (either at or near the proposed bound $\eta/s \simeq 1/(4\pi)$ [232]) are characterized by smaller resistance in comparison to conventional fluids. Secondly, the boundary relaxation time (i.e., the timescale describing the rate of the loss of momentum at the channel boundaries with no-slip boundary conditions) is inverse proportional to η/s.

7 Open questions and perspectives

The scope of this review was mostly limited to observable effects that can be interpreted as evidence of electronic hydrodynamics in graphene and other 2D materials as well as theoretical work exploring hydrodynamic phenomena in electronic systems. Several important topics were purposefully left out, most notably the hydrodynamic behavior of non-electronic excitations in solids, topological hydrodynamics, and generalized hydrodynamics in 1D systems.

The initial argument for electronic hydrodynamics requiring the electron-electron interaction to be the dominant scattering mechanism implies the existence of scale separation between electronic thermalization and energy relaxation due to, e.g., electron-phonon interaction. The latter typically assumes that the phonons are in thermal equilibrium. However, the current-carrying distribution of electrons is generally nonequilibrium and hence electron-phonon coupling can drive the phonons out of equilibrium as well [339]. The resulting phenomenon of phonon drag is well studied [340,341] and in particular allows for a hydrodynamic description [342]. Recently, evidence of the coupled electron-phonon fluid was reported in the Dirac semimetal $PtSn_4$ [343] (for the theory see Ref. [344]), the material characterized by very low resistivity as well as showing a pronounced phonon drag peak [156] at low temperatures. Moreover, it was argued [345] that near-hydrodynamic behavior of electronic transport in the delafossite metals $PdCoO_2$ and $PtCoO_2$ [29,346] should be understood in the context of phonon drag.

Another aspect of the strong coupling between the electronic system and the crystal lattice is the interplay between electronic viscosity and elasticity of the crystal [211,212,215,347,348]. Moreover, static deformations in graphene are known to lead to the appearance of giant pseudomagnetic fields [349]. From a general perspective, elasticity and hydrodynamics belong to a broader class of tensor-field theories that also includes gravitation theories and the theory of critical phenomena in spaces with nontrivial metrics [350].

Observations of viscous hydrodynamics in electronic transport raised the question of whether other excitations in solids might behave hydrodynamically as well. In particular, the classic proposal for the hydrodynamic behavior of spin waves [19] recently came under intense scrutiny both experimentally [351] and theoretically [352,353]. Emergent hydrodynamics in a strongly interacting dipolar spin ensemble (consisting of substitutional nitrogen defects – P1 centers – and nitrogen-vacancy centers in diamonds) was studied experimentally in Ref. [354].

Generally speaking, hydrodynamic flows represent a macroscopic, long wavelength motion governed by global conservation laws [1]. In conventional fluids, these include the particle number, energy, and momentum conservation allowing for a statistical description of the system based on traditional Gibbs approach [16]. In two- (and three-) dimensional electronic systems energy and momentum are conserved only approximately, which limits the applicability of the hydrodynamic approach to a relatively narrow temperature interval [23, 24,176] as well as leads to unconventional behavior [87,114,85]. In one spatial dimension, in particular in the context of integrable (exactly solvable) models, the situation is different: here the system is characterized by a large number of integrals of motion leading to the concept of *generalized* Gibbs ensembles [355, 356]. Applying the hydrodynamic approach to the generalized Gibbs statistics yields generalized hydrodynamics offering new possibilities in describing quantum transport in systems with predominantly ballistic behavior (due to the large number of conservation laws). This approach was introduced in the context of integrable field theories [357] and quantum spin chains [358] and was successfully applied to a number of other integrable systems [359]. The resulting framework was used to describe one-dimensional cold atomic gases at large wavelengths [360] and has been observed experimentally [361]. Generalized hydrodynamics in nonintegrable systems was studied in Ref. [362].

Another topic outside of the scope of this review is topological hydrodynamics, see Ref. [363] and references therein. Recently, an optical topological invariant (measurable via the evanescent magneto-optic Kerr effect [364]) was proposed to describe properties of the viscous Hall fluid [365] suggesting that graphene with the "repulsive" Hall viscosity (i.e., $\omega_c \nu_H > 0$) may be used to create a topological electromagnetic phase of matter. Especially interesting in this context is the interplay of topological band structure and electron-electron interactions (responsible for establishing the local equilibrium underlying the usual hydrodynamic theory). A related issue is quantum hydrodynamics of vorticity [366] describing vortex-antivortex dynamics in 2D bosonic lattices pertaining to the superfluid-insulator transition.

While some experimental work on hydrodynamics in topological materials was addressed in Sec. 2.5.2, the theoretical discussion of Sec. 3 focused on the well-studied cases of the 2D Dirac and Fermi liquids. In contrast, a hydrodynamic theory of topological materials (including Weyl semimetals [59,60] and conducting surface states of topological insulators) has not been hammered out yet. For recent literature on this subject see Refs. [66,367,368,369,370, 371]. The effects of band topology on the shear viscosity were considered in Ref. [372].

Despite the impressive amount of recent work on the subject, electronic hydrodynamics remains a young field with many unanswered questions. So far, the main focus of the community was on Fermi-liquid-type materials (including doped graphene), where the hydrodynamic equations are basically equivalent to the standard Navier-Stokes equation and the velocity field completely determines the electric current. Even in this simplest setting, the question of boundary conditions remains largely unresolved, especially in view of the experiment of Ref. [53]. Furthermore, practical applications of hydrodynamic equations require reliable tools for their numerical solution. Although there exists a massive amount of literature devoted to solution of differential equations (as well as commercial and open source software packages dealing with their numerical solutions), equations of electronic hydrodynamics have to be solved together with the equations describing the electrostatic environment of the system, the electronic circuit into which the system is integrated, and in the case of spintronics applications – the magnetic environment. Combining all these aspects of the problem with the realistic boundary conditions and specific symmetries represents a formidable computational problem that is rather difficult to solve using the available "canned" solvers [373,374].

Transport properties addressed with the hydrodynamic approach so far remain at the semiclassical "Drude" level (which is not surprising given that the hydrodynamic equations, see Sec. 3.2.6, were derived from the semiclassical Boltzmann equation, see Sec. 3. In contrast, the traditional transport theory considers also "higher order" processes leading to the so-called "quantum corrections" [9,12,375]. While typically discussed using field-theoretic methods, these results can also be obtained within the kinetic approach (for the corresponding "quantum kinetic equation" see Ref. [12]). It remains to be seen, whether this physics can be included in a macroscopic, hydrodynamic-type description. Moreover, it is unclear whether one can establish any relation between the well-known hydrodynamic fluctuations [13] and mesoscopic fluctuations in conventional diffusive conductors [376,377,378].

One of the most intriguing promises of the hydrodynamic approach is its supposed ability to describe properties of more complicated systems, including the "strange metals" [312]. This direction of research is still in its infancy. Many novel materials (including van der Waals heterostructures [320,379], conducting surface states of topological insulators, and Weyl semimetals) are characterized by strong spin-orbit coupling. Up until now, a coherent kinetic theory for electrons with spin-orbit interaction has not been established, see Refs. [177,380,381,382]. An advance in this direction could provide a substantial contribution to the application-oriented field of spintronics [383,384], which has been under active development in the last two decades.

The conjecture of Planckian dissipation does not explain why does the observed resistivity remains linear (i.e., does not vanish faster) in different materials where different scattering mechanisms are presumed to be relevant in different temperature regimes [311]. This could indicate an existence of a universal principle limiting the decay rate of the longest lived modes in strongly correlated systems (similarly to the phase space limitations on the quasiparticle properties in Fermi liquids). Such principle has not been established yet.

Finally, the issues raised in the course of the rapid development of electronic hydrodynamics are of fundamental importance for the physics of novel electronic systems necessary for the future development of functional materials. Future advances in this field will have far-reaching implications beyond the scope of particular systems considered in this review. They will substantially improve our understanding of interrelation of macroscopic transport properties (of charge, spin, and heat) and microscopic structure (symmetry properties, band structure, electronic correlations) of materials allowing for material engineering and functionalization.

Acknowledgements The author is grateful to I. Aleiner, P. Alekseev, U. Briskot, I. Burmistrov, A. Dmitriev, I. Gornyi, V. Kachorovskii, E. Kiselev, A. Mirlin, J. Schmalian, M. Schütt, A. Shnirman, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG project NA 1114/5-1, by the German Research

Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 program (Project 873028 HYDROTRONICS).

References

1. P.M. Chaikin, T.C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, 1995).
   DOI 10.1017/CBO9780511813467
2. N. Bloembergen, Physica **15**, 386 (1949).
   DOI 10.1016/0031-8914(49)90114-7
3. L.P. Kadanoff, P.C. Martin, Ann. Phys. (NY) **24**, 419 (1963).
   DOI 10.1016/0003-4916(63)90078-2
4. A. Tucciarone, J.M. Hastings, L.M. Corliss, Phys. Rev. B **8**, 1103 (1973).
   DOI 10.1103/PhysRevB.8.1103
5. B.N. Narozhny, Phys. Rev. B **54**, 3311 (1996).
   DOI 10.1103/PhysRevB.54.3311
6. B.N. Narozhny, A.J. Millis, N. Andrei, Phys. Rev. B **58**, R2921 (1998).
   DOI 10.1103/PhysRevB.58.R2921
7. A. Rosch, N. Andrei, Phys. Rev. Lett. **85**, 1092 (2000).
   DOI 10.1103/PhysRevLett.85.1092
8. J. Sirker, R.G. Pereira, I. Affleck, Phys. Rev. Lett. **103**, 216602 (2009).
   DOI 10.1103/PhysRevLett.103.216602
9. B.L. Altshuler, A.G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, ed. by A.L. Efros, M. Pollak (North-Holland, Amsterdam, 1985).
   DOI 10.1016/B978-0-444-86916-6.50007-7
10. J.M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).
    DOI 10.1017/CBO9781139644075
11. C. Beenakker, H. van Houten, in *Semiconductor Heterostructures and Nanostructures*, *Solid State Physics*, vol. 44, ed. by H. Ehrenreich, D. Turnbull (Academic Press, 1991), pp. 1–228.
    DOI 10.1016/S0081-1947(08)60091-0
12. G. Zala, B.N. Narozhny, I.L. Aleiner, Phys. Rev. B **64**, 214204 (2001).
    DOI 10.1103/PhysRevB.64.214204
13. L.D. Landau, E.M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1987).
    DOI 10.1016/C2013-0-03799-1
14. H. Lamb, *Hydrodynamics* (Dover, New York, 1945)
15. D. Vollhardt, P. Wölfle, *The Superfluid Phases of Helium 3* (Taylor and Francis, London, 1990).
    DOI 10.1201/b12808
16. E.M. Lifshitz, L.P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).
    DOI 10.1016/C2009-0-25523-1
17. R.N. Gurzhi, Soviet Physics Uspekhi **11**(2), 255 (1968). [Usp. Fiz. Nauk **94**, 689 (1968)].
    DOI 10.1070/PU1968v011n02ABEH003815
18. Y. Machida, N. Matsumoto, T. Isono, K. Behnia, Science **367**(6475), 309 (2020).
    DOI 10.1126/science.aaz8043
19. B.I. Halperin, P.C. Hohenberg, Phys. Rev. **188**, 898 (1969).
    DOI 10.1103/PhysRev.188.898
20. R.N. Gurzhi, Zh. Eksp. Teor. Fiz. **44**, 771 (1963). [Sov. Phys. JETP **17**, 521 (1963)]
21. R.N. Gurzhi, Zh. Eksp. Teor. Fiz. **47**, 1415 (1965). [Sov. Phys. JETP **20**, 953 (1965)]
22. M. Dyakonov, M. Shur, Phys. Rev. Lett. **71**, 2465 (1993).
    DOI 10.1103/PhysRevLett.71.2465
23. B.N. Narozhny, I.V. Gornyi, A.D. Mirlin, J. Schmalian, Annalen der Physik **529**(11), 1700043 (2017).
    DOI 10.1002/andp.201700043

24. A. Lucas, K.C. Fong, J. Phys: Condens. Matter **30**(5), 053001 (2018). DOI 10.1088/1361-648x/aaa274
25. M. Polini, A.K. Geim, Phys. Today **73**(6), 28 (2020). DOI 10.1063/PT.3.4497
26. A.A. Abrikosov, I.M. Khalatnikov, Rep. Prog. Phys. **22**(1), 329 (1959). DOI 10.1088/0034-4885/22/1/310
27. D.A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G.H. Auton, E. Khestanova, K.S. Novoselov, I.V. Grigorieva, L.A. Ponomarenko, A.K. Geim, M. Polini, Science **351**, 1055 (2016). DOI 10.1126/science.aad0201
28. J. Crossno, J.K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T.A. Ohki, K.C. Fong, Science **351**(6277), 1058 (2016). DOI 10.1126/science.aad0343
29. P.J.W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, A.P. Mackenzie, Science **351**(6277), 1061 (2016). DOI 10.1126/science.aac8385
30. B.A. Braem, F.M.D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J.V. Koski, M. Berl, W. Dietsche, W. Wegscheider, M. Polini, T. Ihn, K. Ensslin, Phys. Rev. B **98**, 241304(R) (2018). DOI 10.1103/PhysRevB.98.241304
31. A. Jaoui, B. Fauqué, C.W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D.L. Maslov, C. Felser, K. Behnia, npj Quant. Mater. **3**, 64 (2018). DOI 10.1038/s41535-018-0136-x
32. G.M. Gusev, A.S. Jaroshevich, A.D. Levin, Z.D. Kvon, A.K. Bakarov, Sci. Rep. **10**(1), 7860 (2020). DOI 10.1038/s41598-020-64807-6
33. G. Varnavides, A.S. Jermyn, P. Anikeeva, C. Felser, P. Narang, Nat. Commun. **11**(1), 4710 (2020). DOI 10.1038/s41467-020-18553-y
34. U. Vool, A. Hamo, G. Varnavides, Y. Wang, T.X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C.A.C. Garcia, A.T. Pierce, J. Gooth, P. Anikeeva, C. Felser, P. Narang, A. Yacoby, Nat. Phys. (2021). DOI 10.1038/s41567-021-01341-w
35. A. Jaoui, B. Fauqué, K. Behnia, Nat. Commun. **12**(1), 195 (2021). DOI 10.1038/s41467-020-20420-9
36. A. Gupta, J.J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G.C. Gardner, M.J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021). DOI 10.1103/PhysRevLett.126.076803
37. S.S. Pershoguba, A.F. Young, L.I. Glazman, Phys. Rev. B **102**, 125404 (2020). DOI 10.1103/PhysRevB.102.125404
38. D.Y.H. Ho, I. Yudhistira, N. Chakraborty, S. Adam, Phys. Rev. B **97**, 121404(R) (2018). DOI 10.1103/PhysRevB.97.121404
39. R. Krishna Kumar, D.A. Bandurin, F.M.D. Pellegrino, Y. Cao, A. Principi, H. Guo, G.H. Auton, M. Ben Shalom, L.A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I.V. Grigorieva, L.S. Levitov, M. Polini, A.K. Geim, Nat. Phys. **13**(12), 1182 (2017). DOI 10.1038/nphys4240
40. F. Ghahari, H.Y. Xie, T. Taniguchi, K. Watanabe, M.S. Foster, P. Kim, Phys. Rev. Lett. **116**, 136802 (2016). DOI 10.1103/PhysRevLett.116.136802
41. D.A. Bandurin, A.V. Shytov, L.S. Levitov, R. Krishna Kumar, A.I. Berdyugin, M. Ben Shalom, I.V. Grigorieva, A.K. Geim, G. Falkovich, Nat. Commun. **9**(1), 4533 (2018). DOI 10.1038/s41467-018-07004-4
42. A.I. Berdyugin, S.G. Xu, F.M.D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M.B. Shalom, T. Taniguchi, K. Watanabe, I.V. Grigorieva, M. Polini, A.K. Geim, D.A. Bandurin, Science **364**(6436), 162 (2019). DOI 10.1126/science.aau0685

43. P. Gallagher, C.S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, F. Wang, Science **364**(6436), 158 (2019). DOI 10.1126/science.aat8687
44. A. Jenkins, S. Baumann, H. Zhou, S.A. Meynell, D. Yang, T.T. K. Watanabe, A. Lucas, A.F. Young, A.C. Bleszynski Jayich (2020). DOI 10.48550/arXiv.2002.05065
45. Z.J. Krebs, W.A. Behn, S. Li, K.J. Smith, K. Watanabe, T. Taniguchi, A. Levchenko, V.W. Brar (2021). DOI 10.48550/arXiv.2106.07212
46. A. Finkler, Y. Segev, Y. Myasoedov, M.L. Rappaport, L. Ne'eman, D. Vasyukov, E. Zeldov, M.E. Huber, J. Martin, A. Yacoby, Nano Lett. **10**(3), 1046 (2010). DOI 10.1021/nl100009r
47. D. Halbertal, J. Cuppens, M. Ben Shalom, L. Embon, N. Shadmi, Y. Anahory, H.R. Naren, J. Sarkar, A. Uri, Y. Ronen, Y. Myasoedov, L.S. Levitov, E. Joselevich, A.K. Geim, E. Zeldov, Nature **539**(7629), 407 (2016). DOI 10.1038/nature19843
48. A. Marguerite, J. Birkbeck, A. Aharon-Steinberg, D. Halbertal, K. Bagani, I. Marcus, Y. Myasoedov, A.K. Geim, D.J. Perello, E. Zeldov, Nature **575**(7784), 628 (2019). DOI 10.1038/s41586-019-1704-3
49. A. Uri, Y. Kim, K. Bagani, C.K. Lewandowski, S. Grover, N. Auerbach, E.O. Lachman, Y. Myasoedov, T. Taniguchi, K. Watanabe, J. Smet, E. Zeldov, Nat. Phys. **16**(2), 164 (2020). DOI 10.1038/s41567-019-0713-3
50. M.J.H. Ku, T.X. Zhou, Q. Li, Y.J. Shin, J.K. Shi, C. Burch, L.E. Anderson, A.T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M.M. Fogler, P. Kim, A. Yacoby, R.L. Walsworth, Nature **583**, 537 (2020). DOI 10.1038/s41586-020-2507-2
51. J.A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D.J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Stern, T. Scaffidi, A.K. Geim, S. Ilani, Nature **576**, 75 (2019). DOI 10.1038/s41586-019-1788-9
52. L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A.K. Geim, S. Ilani, J.A. Sulpizio, Nat. Nanotechnol. **14**, 480 (2019). DOI 10.1038/s41565-019-0398-x
53. A. Aharon-Steinberg, A. Marguerite, D.J. Perello, K. Bagani, T. Holder, Y. Myasoedov, L.S. Levitov, A.K. Geim, E. Zeldov, Nature **593**, 528 (2021). DOI 10.1038/s41586-021-03501-7
54. C. Kumar, J. Birkbeck, J.A. Sulpizio, D.J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A.K. Geim, S. Ilani (2021). DOI 10.48550/arXiv.2111.06412
55. G.M. Gusev, A.D. Levin, E.V. Levinson, A.K. Bakarov, Phys. Rev. B **98**, 161303(R) (2018). DOI 10.1103/PhysRevB.98.161303
56. O.E. Raichev, G.M. Gusev, A.D. Levin, A.K. Bakarov, Phys. Rev. B **101**, 235314 (2020). DOI 10.1103/PhysRevB.101.235314
57. G.M. Gusev, A.S. Jaroshevich, A.D. Levin, Z.D. Kvon, A.K. Bakarov, Phys. Rev. B **103**, 075303 (2021). DOI 10.1103/PhysRevB.103.075303
58. M.J.M. de Jong, L.W. Molenkamp, Phys. Rev. B **51**, 13389 (1995). DOI 10.1103/PhysRevB.51.13389
59. A. Lucas, R.A. Davison, S. Sachdev, Proc. Natl. Acad. Sci. **113**(34), 9463 (2016). DOI 10.1073/pnas.1608881113
60. E.V. Gorbar, V.A. Miransky, I.A. Shovkovy, P.O. Sukhachov, Phys. Rev. B **97**, 205119 (2018). DOI 10.1103/PhysRevB.97.205119
61. J.S. Bell, R. Jackiw, Nuovo Cimento A **60**, 47 (1969). DOI 10.1007/BF02823296

62. S.L. Adler, Phys. Rev. **177**, 2426 (1969).
DOI 10.1103/PhysRev.177.2426
63. R. Jackiw, Int. J. Mod. Phys. A **25**(04), 659 (2010).
DOI 10.1142/S0217751X10048391
64. C. Felser, B. Yan, Nat. Mater. **15**(11), 1149 (2016).
DOI 10.1038/nmat4741
65. M. Hirschberger, S. Kushwaha, Z. Wang, Q. Gibson, S. Liang, C.A. Belvin, B.A. Bernevig, R.J. Cava, N.P. Ong, Nat. Mater. **15**(11), 1161 (2016).
DOI 10.1038/nmat4684
66. D.T. Son, B.Z. Spivak, Phys. Rev. B **88**, 104412 (2013).
DOI 10.1103/PhysRevB.88.104412
67. F. Arnold, C. Shekhar, S.C. Wu, Y. Sun, R.D. dos Reis, N. Kumar, M. Naumann, M.O. Ajeesh, M. Schmidt, A.G. Grushin, J.H. Bardarson, M. Baenitz, D. Sokolov, H. Borrmann, M. Nicklas, C. Felser, E. Hassinger, B. Yan, Nat. Commun. **7**(1), 11615 (2016).
DOI 10.1038/ncomms11615
68. M. Knudsen, Annalen der Physik **333**(1), 75 (1909). [Ann. Phys. (Leipzig) **28**, 75 (1909)].
DOI 10.1002/andp.19093330106
69. L.J. van der Pauw, Philips Tech. Rev. **20**, 223 (1958)
70. D.A. Abanin, S.V. Morozov, L.A. Ponomarenko, R.V. Gorbachev, A.S. Mayorov, M.I. Katsnelson, K. Watanabe, T. Taniguchi, K.S. Novoselov, L.S. Levitov, A.K. Geim, Science **332**, 328 (2011).
DOI 10.1126/science.1199595
71. P. Drude, Annalen der Physik **306**(1), 566 (1900). [Ann. Phys. (Leipzig) **1**, 566 (1900)].
DOI 10.1002/andp.19003060312
72. W.J. Skocpol, P.M. Mankiewich, R.E. Howard, L.D. Jackel, D.M. Tennant, A.D. Stone, Phys. Rev. Lett. **58**, 2347 (1987).
DOI 10.1103/PhysRevLett.58.2347
73. H. van Houten, C.W.J. Beenakker, J.G. Williamson, M.E.I. Broekaart, P.H.M. van Loosdrecht, B.J. van Wees, J.E. Mooij, C.T. Foxon, J.J. Harris, Phys. Rev. B **39**, 8556 (1989).
DOI 10.1103/PhysRevB.39.8556
74. A.K. Geim, P.C. Main, P.H. Beton, P. Steda, L. Eaves, C.D.W. Wilkinson, S.P. Beaumont, Phys. Rev. Lett. **67**, 3014 (1991).
DOI 10.1103/PhysRevLett.67.3014
75. K.L. Shepard, M.L. Roukes, B.P. van der Gaag, Phys. Rev. Lett. **68**, 2660 (1992).
DOI 10.1103/PhysRevLett.68.2660
76. Y. Hirayama, A.D. Wieck, T. Bever, K. von Klitzing, K. Ploog, Phys. Rev. B **46**, 4035 (1992).
DOI 10.1103/PhysRevB.46.4035
77. G. Mihajlović, J.E. Pearson, M.A. Garcia, S.D. Bader, A. Hoffmann, Phys. Rev. Lett. **103**, 166601 (2009).
DOI 10.1103/PhysRevLett.103.166601
78. R.V. Gorbachev, J.C.W. Song, G.L. Yu, A.V. Kretinin, F. Withers, Y. Cao, A. Mishchenko, I.V. Grigorieva, K.S. Novoselov, L.S. Levitov, A.K. Geim, Science **346**(6208), 448 (2014).
DOI 10.1126/science.1254966
79. P.L. McEuen, A. Szafer, C.A. Richter, B.W. Alphenaar, J.K. Jain, A.D. Stone, R.G. Wheeler, R.N. Sacks, Phys. Rev. Lett. **64**, 2062 (1990).
DOI 10.1103/PhysRevLett.64.2062
80. J.K. Wang, V.J. Goldman, Phys. Rev. B **45**, 13479 (1992).
DOI 10.1103/PhysRevB.45.13479
81. A. Roth, C. Brune, H. Buhmann, L.W. Molenkamp, J. Maciejko, X.L. Qi, S.C. Zhang, Science **325**(5938), 294 (2009).
DOI 10.1126/science.1174736
82. X.P. Zhang, C. Huang, M.A. Cazalilla, 2D Materials **4**(2), 024007 (2017).
DOI 10.1088/2053-1583/aa5e9b

83. K. Komatsu, Y. Morita, E. Watanabe, D. Tsuya, K. Watanabe, T. Taniguchi, S. Moriyama, Science Advances **4**(5), eaaq0194 (2018). DOI 10.1126/sciadv.aaq0194
84. F. Chiappini, S. Wiedmann, M. Titov, A.K. Geim, R.V. Gorbachev, E. Khestanova, A. Mishchenko, K.S. Novoselov, J.C. Maan, U. Zeitler, Phys. Rev. B **94**, 085302 (2016). DOI 10.1103/PhysRevB.94.085302
85. S. Danz, M. Titov, B.N. Narozhny, Phys. Rev. B **102**, 081114(R) (2020). DOI 10.1103/PhysRevB.102.081114
86. M. Titov, R.V. Gorbachev, B.N. Narozhny, T. Tudorovskiy, M. Schütt, P.M. Ostrovsky, I.V. Gornyi, A.D. Mirlin, M.I. Katsnelson, K.S. Novoselov, A.K. Geim, L.A. Ponomarenko, Phys. Rev. Lett. **111**, 166601 (2013). DOI 10.1103/PhysRevLett.111.166601
87. P.S. Alekseev, A.P. Dmitriev, I.V. Gornyi, V.Y. Kachorovskii, B.N. Narozhny, M. Schütt, M. Titov, Phys. Rev. Lett. **114**, 156601 (2015). DOI 10.1103/PhysRevLett.114.156601
88. G.Y. Vasileva, D. Smirnov, Y.L. Ivanov, Y.B. Vasilyev, P.S. Alekseev, A.P. Dmitriev, I.V. Gornyi, V.Y. Kachorovskii, M. Titov, B.N. Narozhny, R.J. Haug, Phys. Rev. B **93**, 195430 (2016). DOI 10.1103/PhysRevB.93.195430
89. M.S. Foster, I.L. Aleiner, Phys. Rev. B **79**, 085415 (2009). DOI 10.1103/PhysRevB.79.085415
90. C.L. Navier, in *Mémoires de l'Académie des sciences de l'Institut de France - Année 1823* (Gauthier-Villars, Paris, 1827), pp. 389–440
91. G.G. Stokes, Transactions of the Cambridge Philosophical Society **8**, 287 (1845)
92. G.G. Stokes, Transactions of the Cambridge Philosophical Society **9**, 8 (1851)
93. S. Danz, B.N. Narozhny, 2D Materials **7**(3), 035001 (2020). DOI 10.1088/2053-1583/ab7bfa
94. I. Torre, A. Tomadin, A.K. Geim, M. Polini, Phys. Rev. B **92**, 165433 (2015). DOI 10.1103/PhysRevB.92.165433
95. F.M.D. Pellegrino, I. Torre, A.K. Geim, M. Polini, Phys. Rev. B **94**, 155414 (2016). DOI 10.1103/PhysRevB.94.155414
96. L.S. Levitov, G. Falkovich, Nat. Phys. **12**(7), 672 (2016). DOI 10.1038/nphys3667
97. R. Moessner, P. Surówka, P. Witkowski, Phys. Rev. B **97**, 161112 (2018). DOI 10.1103/PhysRevB.97.161112
98. R. Moessner, N. Morales-Durán, P. Surówka, P. Witkowski, Phys. Rev. B **100**, 155115 (2019). DOI 10.1103/PhysRevB.100.155115
99. J. Mayzel, V. Steinberg, A. Varshney, Nat. Commun. **10**, 937 (2019). DOI 10.1038/s41467-019-08916-5
100. A.V. Shytov, J.F. Kong, G. Falkovich, L.S. Levitov, Phys. Rev. Lett. **121**, 176805 (2018). DOI 10.1103/PhysRevLett.121.176805
101. H. Guo, E. Ilseven, G. Falkovich, L.S. Levitov, Proc. Natl. Acad. Sci. **114**(12), 3068 (2017). DOI 10.1073/pnas.1612181114
102. B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, C.T. Foxon, Phys. Rev. Lett. **60**, 848 (1988). DOI 10.1103/PhysRevLett.60.848
103. D.A. Wharam, T.J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J.E.F. Frost, D.G. Hasko, D.C. Peacock, D.A. Ritchie, G.A.C. Jones, Journal of Physics C **21**(8), L209 (1988). DOI 10.1088/0022-3719/21/8/002
104. Y.V. Sharvin, Zh. Eksp. Teor. Fiz. **48**, 984 (1965). [Sov. Phys. JETP **21**, 655 (1965)]
105. G. Zhang, V. Kachorovskii, K. Tikhonov, I. Gornyi, Phys. Rev. B **104**, 075417 (2021). DOI 10.1103/PhysRevB.104.075417
106. A. Lucas, Phys. Rev. B **95**, 115425 (2017). DOI 10.1103/PhysRevB.95.115425

107. H. Guo, E. Ilseven, G. Falkovich, L. Levitov. Stokes Paradox, Back Reflections and Interaction-Enhanced Conduction (2017). DOI 10.48550/arXiv.1612.09239
108. C.W. Oseen, Ark. f. Mat., Astr. och Fysik (Stockholm) **6**, 29 (1910)
109. P.S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016). DOI 10.1103/PhysRevLett.117.166601
110. T. Scaffidi, N. Nandi, B. Schmidt, A.P. Mackenzie, J.E. Moore, Phys. Rev. Lett. **118**, 226601 (2017). DOI 10.1103/PhysRevLett.118.226601
111. S. Tarucha, T. Saku, Y. Tokura, Y. Hirayama, Phys. Rev. B **47**, 4064 (1993). DOI 10.1103/PhysRevB.47.4064
112. S. Datta, *Electron Transport in Mesoscopic Systems* (Cambridge University Press, 1997). DOI 10.1017/CBO9780511805776
113. F.M.D. Pellegrino, I. Torre, M. Polini, Phys. Rev. B **96**, 195401 (2017). DOI 10.1103/PhysRevB.96.195401
114. P.S. Alekseev, A.P. Dmitriev, I.V. Gornyi, V.Y. Kachorovskii, B.N. Narozhny, M. Titov, Phys. Rev. B **97**, 085109 (2018). DOI 10.1103/PhysRevB.97.085109
115. P.S. Alekseev, A.P. Dmitriev, I.V. Gornyi, V.Y. Kachorovskii, B.N. Narozhny, M. Titov, Phys. Rev. B **98**, 125111 (2018). DOI 10.1103/PhysRevB.98.125111
116. J.L.M. Poiseuille, C. R. Acad. Sci. **11**, 961 (1840)
117. E.I. Kiselev, J. Schmalian, Phys. Rev. B **99**, 035430 (2019). DOI 10.1103/PhysRevB.99.035430
118. J.R. Maze, P.L. Stanwix, J.S. Hodges, S. Hong, J.M. Taylor, P. Cappellaro, L. Jiang, M.V. Gurudev Dutt, E. Togan, A.S. Zibrov, A. Yacoby, R.L. Walsworth, M.D. Lukin, Nature **455**, 644 (2008). DOI 10.1038/nature07279
119. B.N. Narozhny, M. Schütt, Phys. Rev. B **100**, 035125 (2019). DOI 10.1103/PhysRevB.100.035125
120. B.N. Narozhny, Ann. Phys. **411**, 167979 (2019). DOI 10.1016/j.aop.2019.167979
121. U. Briskot, M. Schütt, I.V. Gornyi, M. Titov, B.N. Narozhny, A.D. Mirlin, Phys. Rev. B **92**, 115426 (2015). DOI 10.1103/PhysRevB.92.115426
122. M. Schütt, P.M. Ostrovsky, I.V. Gornyi, A.D. Mirlin, Phys. Rev. B **83**, 155441 (2011). DOI 10.1103/PhysRevB.83.155441
123. M. Müller, S. Sachdev, Phys. Rev. B **78**, 115419 (2008). DOI 10.1103/PhysRevB.78.115419
124. B.N. Narozhny, I.V. Gornyi, M. Titov, M. Schütt, A.D. Mirlin, Phys. Rev. B **91**, 035414 (2015). DOI 10.1103/PhysRevB.91.035414
125. B.N. Narozhny, I.V. Gornyi, M. Titov, Phys. Rev. B **104**, 075443 (2021). DOI 10.1103/PhysRevB.104.075443
126. J.P. Eisenstein, in *Perspectives in Quantum Hall Effects*, ed. by S.D. Sarma, A. Pinczuk (Wiley, New York, 1997). DOI 10.1002/9783527617258.ch2
127. S.M. Girvin, A.H. MacDonald, in *Perspectives in Quantum Hall Effects*, ed. by S.D. Sarma, A. Pinczuk (Wiley, New York, 1997). DOI 10.1002/9783527617258.ch5
128. B.A. Bernevig, T.L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, 2013)
129. Y.T. Cui, B. Wen, E.Y. Ma, G. Diankov, Z. Han, F. Amet, T. Taniguchi, K. Watanabe, D. Goldhaber-Gordon, C.R. Dean, Z.X. Shen, Phys. Rev. Lett. **117**, 186601 (2016). DOI 10.1103/PhysRevLett.117.186601
130. J. Chae, S. Jung, S. Woo, H. Baek, J. Ha, Y.J. Song, Y.W. Son, N.B. Zhitenev, J.A. Stroscio, Y. Kuk, Nano Letters **12**(4), 1839 (2012). DOI 10.1021/nl2041222

131. J.M. Marmolejo-Tejada, J.H. García, M.D. Petrović, P.H. Chang, X.L. Sheng, A. Cresti, P. Plecháč, S. Roche, B.K. Nikolić, Journal of Physics: Materials **1**(1), 015006 (2018).
DOI 10.1088/2515-7639/aad585
132. M.T. Allen, O. Shtanko, I.C. Fulga, A.R. Akhmerov, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, L.S. Levitov, A. Yacoby, Nat. Phys. **12**, 128 (2016).
DOI 10.1038/nphys3534
133. S.M. Sze, K.K. Ng, *Physics of Semiconductor Devices* (Wiley, New York, 2006).
DOI 10.1002/0470068329
134. W. Schottky, Naturwissenschaften **26**, 843 (1938).
DOI 10.1007/BF01774216
135. M.J. Zhu, A.V. Kretinin, M.D. Thompson, D.A. Bandurin, S. Hu, G.L. Yu, J. Birkbeck, A. Mishchenko, I.J. Vera-Marun, K. Watanabe, T. Taniguchi, M. Polini, J.R. Prance, K.S. Novoselov, A.K. Geim, M. Ben Shalom, Nat. Commun. **8**, 14552 (2017).
DOI 10.1038/ncomms14552
136. D. Halbertal, M. Ben Shalom, A. Uri, K. Bagani, A.Y. Meltzer, I. Markus, Y. Myasoedov, J. Birkbeck, L.S. Levitov, A.K. Geim, E. Zeldov, Science **358**, 1303 (2017).
DOI 10.1126/science.aan0877
137. B.A. Braem, C. Gold, S. Hennel, M. Röösli, M. Berl, W. Dietsche, W. Wegscheider, K. Ensslin, T. Ihn, New Journal of Physics **20**(7), 073015 (2018).
DOI 10.1088/1367-2630/aad068
138. M. Shavit, A.V. Shytov, G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).
DOI 10.1103/PhysRevLett.123.026801
139. O. M. Corbino, Nuovo Cimento **1**, 397 (1911).
DOI 10.1007/BF02958241
140. A. Tomadin, G. Vignale, M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).
DOI 10.1103/PhysRevLett.113.235901
141. R. Franz, G. Wiedemann, Annalen der Physik **165**(8), 497 (1853).
DOI 10.1002/andp.18531650802
142. L. Lorenz, Annalen der Physik **223**(11), 429 (1872).
DOI 10.1002/andp.18722231107
143. J.G. Hust, A.B. Lankford, *Thermal conductivity of aluminum, copper, iron and tungsten for temperatures from 1 K to the melting point* (1984). National Bureau of Standards, Boulder, Colorado; NBSIR 84-3007
144. Y.M. Zuev, W. Chang, P. Kim, Phys. Rev. Lett. **102**, 096807 (2009).
DOI 10.1103/PhysRevLett.102.096807
145. J.G. Checkelsky, N.P. Ong, Phys. Rev. B **80**, 081413(R) (2009).
DOI 10.1103/PhysRevB.80.081413
146. A. Block, A. Principi, N.C.H. Hesp, A.W. Cummings, M. Liebel, K. Watanabe, T. Taniguchi, S. Roche, F.H.L. Koppens, N.F. van Hulst, K.J. Tielrooij, Nature Nanotechnology **16**, 1195 (2021).
DOI 10.1038/s41565-021-00957-6
147. S.A. Hartnoll, P.K. Kovtun, M. Müller, S. Sachdev, Phys. Rev. B **76**, 144502 (2007).
DOI 10.1103/PhysRevB.76.144502
148. P.R. Wallace, Phys. Rev. **71**, 622 (1947).
DOI 10.1103/PhysRev.71.622
149. G.W. Semenoff, Phys. Rev. Lett. **53**, 2449 (1984).
DOI 10.1103/PhysRevLett.53.2449
150. D.P. DiVincenzo, E.J. Mele, Phys. Rev. B **29**, 1685 (1984).
DOI 10.1103/PhysRevB.29.1685
151. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, A.A. Firsov, Nature **438**, 197 (2005).
DOI 10.1038/nature04233
152. M. Müller, L. Fritz, S. Sachdev, Phys. Rev. B **78**, 115406 (2008).
DOI 10.1103/PhysRevB.78.115406
153. L. Fritz, J. Schmalian, M. Müller, S. Sachdev, Phys. Rev. B **78**, 085416 (2008).
DOI 10.1103/PhysRevB.78.085416
154. E.G. Mishchenko, Phys. Rev. Lett. **98**, 216801 (2007).
DOI 10.1103/PhysRevLett.98.216801

155. E.G. Mishchenko, EPL (Europhysics Letters) **83**(1), 17005 (2008).
DOI 10.1209/0295-5075/83/17005
156. J. Gooth, F. Menges, N. Kumar, V. Süβ, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, B. Gotsmann, Nat. Commun. **9**, 4093 (2018).
DOI 10.1038/s41467-018-06688-y
157. N. Kumar, Y. Sun, M. Nicklas, S.J. Watzman, O. Young, I. Leermakers, J. Hornung, J. Klotz, J. Gooth, K. Manna, V. Süß, S.N. Guin, T. Förster, M. Schmidt, L. Muechler, B. Yan, P. Werner, W. Schnelle, U. Zeitler, J. Wosnitza, S. Parkin, C. Felser, C. Shekhar, Nat. Commun. **10**(1), 2475 (2019).
DOI 10.1038/s41467-019-10126-y
158. H.Y. Xie, M.S. Foster, Phys. Rev. B **93**, 195103 (2016).
DOI 10.1103/PhysRevB.93.195103
159. A. Principi, G. Vignale, Phys. Rev. Lett. **115**, 056603 (2015).
DOI 10.1103/PhysRevLett.115.056603
160. A. Lucas, S. Das Sarma, Phys. Rev. B **97**, 245128 (2018).
DOI 10.1103/PhysRevB.97.245128
161. M.R. van Delft, Y. Wang, C. Putzke, J. Oswald, G. Varnavides, C.A.C. Garcia, C. Guo, H. Schmid, V. Süss, H. Borrmann, J. Diaz, Y. Sun, C. Felser, B. Gotsmann, P. Narang, P.J.W. Moll, Nat. Commun. **12**, 4799 (2021).
DOI 10.1038/s41467-021-25037-0
162. E.H. Sondheimer, Phys. Rev. **80**, 401 (1950).
DOI 10.1103/PhysRev.80.401
163. E. Shuryak, Prog. Part. Nucl. Phys. **62**(1), 48 (2009).
DOI 10.1016/j.ppnp.2008.09.001
164. S.N. Shore, *Astrophysical Hydrodynamics: An Introduction* (Wiley, Weinheim, 2007).
DOI 10.1002/9783527619054
165. M.S. Steinberg, Phys. Rev. **109**, 1486 (1958).
DOI 10.1103/PhysRev.109.1486
166. M.I. Katsnelson, *Graphene* (Cambridge University Press, 2012).
DOI 10.1017/CBO9781139031080
167. M.J. Bhaseen, A.G. Green, S.L. Sondhi, Phys. Rev. B **79**, 094502 (2009).
DOI 10.1103/PhysRevB.79.094502
168. L. Euler, Mémoires de l'académie des sciences de Berlin **11**, 274 (1757)
169. T.V. Phan, J.C.W. Song, L.S. Levitov (2013).
DOI 10.48550/arXiv.1306.4972
170. L.S. Levitov, A.V. Shtyk, M.V. Feigelman, Phys. Rev. B **88**, 235403 (2013).
DOI 10.1103/PhysRevB.88.235403
171. Z. Sun, D.N. Basov, M.M. Fogler, Proc. Natl. Acad. Sci. **115**(13), 3285 (2018).
DOI 10.1073/pnas.1717010115
172. D. Svintsov, Phys. Rev. B **97**, 121405(R) (2018).
DOI 10.1103/PhysRevB.97.121405
173. A. Lucas, S. Das Sarma, Phys. Rev. B **97**, 115449 (2018).
DOI 10.1103/PhysRevB.97.115449
174. E.I. Kiselev, J. Schmalian, Phys. Rev. B **102**, 245434 (2020).
DOI 10.1103/PhysRevB.102.245434
175. D.V. Fateev, V.V. Popov, Semiconductors **54**, 941 (2020).
DOI 10.1134/S1063782620080084
176. B.N. Narozhny, I.V. Gornyi, M. Titov, Phys. Rev. B **103**, 115402 (2021).
DOI 10.1103/PhysRevB.103.115402
177. M. Auslender, M.I. Katsnelson, Phys. Rev. B **76**, 235425 (2007).
DOI 10.1103/PhysRevB.76.235425
178. A.A. Kozikov, A.K. Savchenko, B.N. Narozhny, A.V. Shytov, Phys. Rev. B **82**, 075424 (2010).
DOI 10.1103/PhysRevB.82.075424
179. D.E. Sheehy, J. Schmalian, Phys. Rev. Lett. **99**, 226803 (2007).
DOI 10.1103/PhysRevLett.99.226803
180. M. Müller, J. Schmalian, L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).
DOI 10.1103/PhysRevLett.103.025301

181. J. Gonzalez, F. Guinea, M.A.H. Vozmediano, Nuclear Physics B **424**(3), 595 (1994). DOI 10.1016/0550-3213(94)90410-3
182. J. González, F. Guinea, M.A.H. Vozmediano, Phys. Rev. B **59**, R2474 (1999). DOI 10.1103/PhysRevB.59.R2474
183. J.M. Link, B.N. Narozhny, E.I. Kiselev, J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018). DOI 10.1103/PhysRevLett.120.196801
184. I.L. Aleiner, O. Agam, Ann. Phys. **385**, 716 (2017). DOI 10.1016/j.aop.2017.08.017
185. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Mod. Phys. **81**, 109 (2009). DOI 10.1103/RevModPhys.81.109
186. G. Wagner, D.X. Nguyen, S.H. Simon, Phys. Rev. B **101**, 245438 (2020). DOI 10.1103/PhysRevB.101.245438
187. H.K. Tang, J.N. Leaw, J.N.B. Rodrigues, I.F. Herbut, P. Sengupta, F.F. Assaad, S. Adam, Science **361**(6402), 570 (2018). DOI 10.1126/science.aao2934
188. D.C. Elias, R.V. Gorbachev, A.S. Mayorov, S.V. Morozov, A.A. Zhukov, P. Blake, L.A. Ponomarenko, I.V. Grigorieva, K.S. Novoselov, F. Guinea, A.K. Geim, Nat. Phys. **7**, 701 (2011). DOI 10.1038/NPHYS2049
189. G. Alymov, V. Vyurkov, V. Ryzhii, A. Satou, D. Svintsov, Phys. Rev. B **97**, 205411 (2018). DOI 10.1103/PhysRevB.97.205411
190. A. Tomadin, D. Brida, G. Cerullo, A.C. Ferrari, M. Polini, Phys. Rev. B **88**, 035430 (2013). DOI 10.1103/PhysRevB.88.035430
191. A.B. Kashuba, Phys. Rev. B **78**, 085415 (2008). DOI 10.1103/PhysRevB.78.085415
192. M. Schütt, P.M. Ostrovsky, M. Titov, I.V. Gornyi, B.N. Narozhny, A.D. Mirlin, Phys. Rev. Lett. **110**, 026601 (2013). DOI 10.1103/PhysRevLett.110.026601
193. J.C.W. Song, M.Y. Reizer, L.S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012). DOI 10.1103/PhysRevLett.109.106602
194. M.W. Graham, S.F. Shi, D.C. Ralph, J. Park, P.L. McEuen, Nat. Phys. **9**, 103 (2013). DOI 10.1038/nphys2493
195. A.C. Betz, S.H. Jhang, E. Pallecchi, R. Ferreira, G. Fève, J.M. Berroir, B. Plaçais, Nat. Phys. **9**, 109 (2013). DOI 10.1038/nphys2494
196. K.S. Tikhonov, I.V. Gornyi, V.Y. Kachorovskii, A.D. Mirlin, Phys. Rev. B **97**, 085415 (2018). DOI 10.1103/PhysRevB.97.085415
197. J.F. Kong, L. Levitov, D. Halbertal, E. Zeldov, Phys. Rev. B **97**, 245416 (2018). DOI 10.1103/PhysRevB.97.245416
198. M. Zarenia, A. Principi, G. Vignale, Phys. Rev. B **102**, 214304 (2020). DOI 10.1103/PhysRevB.102.214304
199. P.S. Alekseev, A.P. Dmitriev, I.V. Gornyi, V.Y. Kachorovskii, B.N. Narozhny, M. Schütt, M. Titov, Phys. Rev. B **95**, 165410 (2017). DOI 10.1103/PhysRevB.95.165410
200. R.V. Gorbachev, A.K. Geim, M.I. Katsnelson, K.S. Novoselov, T. Tudorovskiy, I.V. Grigorieva, A.H. MacDonald, S.V. Morozov, K. Watanabe, T. Taniguchi, L.A. Ponomarenko, Nat. Phys. **8**(12), 896 (2012). DOI 10.1038/nphys2441
201. B.N. Narozhny, I.V. Gornyi, Frontiers in Physics **9**, 108 (2021). DOI 10.3389/fphy.2021.640649
202. R. Bistritzer, A.H. MacDonald, Phys. Rev. Lett. **102**, 206410 (2009). DOI 10.1103/PhysRevLett.102.206410
203. W.K. Tse, S. Das Sarma, Phys. Rev. B **79**, 235406 (2009). DOI 10.1103/PhysRevB.79.235406

204. H.Y. Xie, A. Levchenko, Phys. Rev. B **99**, 045434 (2019).
DOI 10.1103/PhysRevB.99.045434
205. O. Kashuba, B. Trauzettel, L.W. Molenkamp, Phys. Rev. B **97**, 205129 (2018).
DOI 10.1103/PhysRevB.97.205129
206. B.N. Narozhny, Phys. Rev. B **100**, 115434 (2019).
DOI 10.1103/PhysRevB.100.115434
207. G. Falkovich, L.S. Levitov, Phys. Rev. Lett. **119**, 066601 (2017).
DOI 10.1103/PhysRevLett.119.066601
208. I.M. Khalatnikov, Zh. Eksp. Teor. Fiz. **29**, 253 (1956). [Soviet Physics JETP-USSR **2**, 169 (1956)]
209. J. Sykes, G.A. Brooker, Ann. Phys. **56**(1), 1 (1970).
DOI 10.1016/0003-4916(70)90002-3
210. V.A. Zakharov, I.S. Burmistrov, Phys. Rev. B **103**, 235305 (2021).
DOI 10.1103/PhysRevB.103.235305
211. B. Bradlyn, M. Goldstein, N. Read, Phys. Rev. B **86**, 245309 (2012).
DOI 10.1103/PhysRevB.86.245309
212. P. Rao, B. Bradlyn, Phys. Rev. X **10**, 021005 (2020).
DOI 10.1103/PhysRevX.10.021005
213. L.V. Delacrétaz, A. Gromov, Phys. Rev. Lett. **119**, 226602 (2017).
DOI 10.1103/PhysRevLett.119.226602
214. T. Holder, R. Queiroz, A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).
DOI 10.1103/PhysRevLett.123.106801
215. J.M. Link, D.E. Sheehy, B.N. Narozhny, J. Schmalian, Phys. Rev. B **98**, 195103 (2018).
DOI 10.1103/PhysRevB.98.195103
216. A. Principi, G. Vignale, M. Carrega, M. Polini, Phys. Rev. B **93**, 125410 (2016).
DOI 10.1103/PhysRevB.93.125410
217. I.S. Burmistrov, M. Goldstein, M. Kot, V.D. Kurilovich, P.D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).
DOI 10.1103/PhysRevLett.123.026804
218. Y. Liao, V. Galitski, Phys. Rev. B **101**, 195106 (2020).
DOI 10.1103/PhysRevB.101.195106
219. B.N. Narozhny, A. Levchenko, Rev. Mod. Phys. **88**, 025003 (2016).
DOI 10.1103/RevModPhys.88.025003
220. B.N. Narozhny, M. Titov, I.V. Gornyi, P.M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).
DOI 10.1103/PhysRevB.85.195421
221. P.S. Alekseev, Phys. Rev. B **98**, 165440 (2018).
DOI 10.1103/PhysRevB.98.165440
222. B. Laikhtman, Phys. Rev. B **45**, 1259 (1992).
DOI 10.1103/PhysRevB.45.1259
223. R.N. Gurzhi, A.N. Kalinenko, A.I. Kopeliovich, Phys. Rev. Lett. **74**, 3872 (1995).
DOI 10.1103/PhysRevLett.74.3872
224. P. Ledwith, H. Guo, A. Shytov, L. Levitov, Phys. Rev. Lett. **123**, 116601 (2019).
DOI 10.1103/PhysRevLett.123.116601
225. P.S. Alekseev, A.P. Dmitriev, Phys. Rev. B **102**, 241409 (2020).
DOI 10.1103/PhysRevB.102.241409
226. H. Isobe, B.J. Yang, A. Chubukov, J. Schmalian, N. Nagaosa, Phys. Rev. Lett. **116**, 076803 (2016).
DOI 10.1103/PhysRevLett.116.076803
227. A. Kobayashi, Y. Suzumura, F. Piéchon, G. Montambaux, Phys. Rev. B **84**, 075450 (2011).
DOI 10.1103/PhysRevB.84.075450
228. V. Pardo, W.E. Pickett, Phys. Rev. Lett. **102**, 166803 (2009).
DOI 10.1103/PhysRevLett.102.166803
229. S. Banerjee, R.R.P. Singh, V. Pardo, W.E. Pickett, Phys. Rev. Lett. **103**, 016402 (2009).
DOI 10.1103/PhysRevLett.103.016402
230. C. Fang, L. Fu, Phys. Rev. B **91**, 161105 (2015).
DOI 10.1103/PhysRevB.91.161105

231. S.M. Huang, S.Y. Xu, I. Belopolski, C.C. Lee, G. Chang, T.R. Chang, B. Wang, N. Alidoust, G. Bian, M. Neupane, D. Sanchez, H. Zheng, H.T. Jeng, A. Bansil, T. Neupert, H. Lin, M.Z. Hasan, Proc. Natl. Acade. Sci. **113**(5), 1180 (2016). DOI 10.1073/pnas.1514581113
232. P.K. Kovtun, D.T. Son, A.O. Starinets, Phys. Rev. Lett. **94**, 111601 (2005). DOI 10.1103/PhysRevLett.94.111601
233. F. Peña Benitez, K. Saha, P. Surówka, Phys. Rev. B **99**, 045141 (2019). DOI 10.1103/PhysRevB.99.045141
234. C. Hoyos, R. Lier, F. Peña Benitez, P. Surówka, Phys. Rev. B **102**, 081303 (2020). DOI 10.1103/PhysRevB.102.081303
235. J.M. Luttinger, Phys. Rev. **102**, 1030 (1956). DOI 10.1103/PhysRev.102.1030
236. A.A. Abrikosov, Zh. Eksp. Teor. Fiz. **66**, 1443 (1974). [Sov. Phys. JETP **39**, 709 (1974)]
237. J.M. Link, I.F. Herbut, Phys. Rev. B **101**, 125128 (2020). DOI 10.1103/PhysRevB.101.125128
238. J.C. Maxwell, Phil. Trans. R. Soc. **170**, 231 (1879). DOI 10.1098/rstl.1879.0067
239. L.A. Falkovsky, Advances in Physics **32**(5), 753 (1983). DOI 10.1080/00018738300101601
240. T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J.A. Sulpizio, L. Ella, S. Ilani, A. Stern, Phys. Rev. B **100**, 245305 (2019). DOI 10.1103/PhysRevB.100.245305
241. O.E. Raichev, Phys. Rev. B **105**, L041301 (2022). DOI 10.1103/PhysRevB.105.L041301
242. A.i.e.i.f.C. Keser, D.Q. Wang, O. Klochan, D.Y.H. Ho, O.A. Tkachenko, V.A. Tkachenko, D. Culcer, S. Adam, I. Farrer, D.A. Ritchie, O.P. Sushkov, A.R. Hamilton, Phys. Rev. X **11**, 031030 (2021). DOI 10.1103/PhysRevX.11.031030
243. M.P. Marder, *Condensed Matter Physics* (Wiley, 2010). DOI 10.1002/9780470949955
244. R. Landauer, IBM J. Res. Dev. **1**(3), 223 (1957). DOI 10.1147/rd.13.0223
245. A. Lucas, Phys. Rev. B **93**, 245153 (2016). DOI 10.1103/PhysRevB.93.245153
246. M. Semenyakin, G. Falkovich, Phys. Rev. B **97**, 085127 (2018). DOI 10.1103/PhysRevB.97.085127
247. G. Giuliani, G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005). DOI 10.1017/CBO9780511619915
248. A. Hill, S.A. Mikhailov, K. Ziegler, EPL (Europhysics Letters) **87**(2), 27005 (2009). DOI 10.1209/0295-5075/87/27005
249. A. Principi, R. Asgari, M. Polini, Solid State Communications **151**(21), 1627 (2011). DOI 10.1016/j.ssc.2011.07.015
250. Z. Fei, A.S. Rodin, G.O. Andreev, W. Bao, A.S. McLeod, M. Wagner, L.M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M.M. Fogler, A.H.C. Neto, C.N. Lau, F. Keilmann, D.N. Basov, Nature **487**(7405), 82 (2012). DOI 10.1038/nature11253
251. J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, A. Zurutuza Elorza, N. Camara, F.J.G. de Abajo, R. Hillenbrand, F.H.L. Koppens, Nature **487**, 77 (2012). DOI 10.1038/nature11254
252. G.X. Ni, L. Wang, M.D. Goldflam, M. Wagner, Z. Fei, A.S. McLeod, M.K. Liu, F. Keilmann, B. Özyilmaz, A.H. Castro Neto, J. Hone, F.M. M., B.D. N., Nat. Photon. **10**, 244 (2016). DOI 10.1038/nphoton.2016.45
253. M.B. Lundeberg, Y. Gao, R. Asgari, C. Tan, B. Van Duppen, M. Autore, P. Alonso-González, A. Woessner, K. Watanabe, T. Taniguchi, R. Hillenbrand, J. Hone, M. Polini,

F.H.L. Koppens, Science **357**(6347), 187 (2017).
DOI 10.1126/science.aan2735
254. D. Alcaraz Iranzo, S. Nanot, E.J.C. Dias, I. Epstein, C. Peng, D.K. Efetov, M.B. Lundeberg, R. Parret, J. Osmond, J.Y. Hong, J. Kong, D.R. Englund, N.M.R. Peres, F.H.L. Koppens, Science **360**(6386), 291 (2018).
DOI 10.1126/science.aar8438
255. P. Novelli, I. Torre, F.H.L. Koppens, F. Taddei, M. Polini, Phys. Rev. B **102**, 125403 (2020).
DOI 10.1103/PhysRevB.102.125403
256. T. Giovannini, L. Bonatti, M. Polini, C. Cappelli, The Journal of Physical Chemistry Letters **11**(18), 7595 (2020).
DOI 10.1021/acs.jpclett.0c02051
257. N.C.H. Hesp, I. Torre, D. Rodan-Legrain, P. Novelli, Y. Cao, S. Carr, S. Fang, P. Stepanov, D. Barcons-Ruiz, H. Herzig-Sheinfux, K. Watanabe, T. Taniguchi, D.K. Efetov, E. Kaxiras, P. Jarillo-Herrero, M. Polini, F.H.L. Koppens. Observation of interband collective excitations in twisted bilayer graphene (2021).
DOI 10.1038/s41567-021-01327-8
258. A.T. Costa, P.A.D. Gonçalves, D.N. Basov, F.H.L. Koppens, N.A. Mortensen, N.M.R. Peres, Proc. Natl. Acad. Sci. **118**(4), e2012847118 (2021).
DOI 10.1073/pnas.2012847118
259. A. Klein, D.L. Maslov, L.P. Pitaevskii, A.V. Chubukov, Phys. Rev. Research **1**, 033134 (2019).
DOI 10.1103/PhysRevResearch.1.033134
260. B.A. Ferreira, B. Amorim, A.J. Chaves, N.M.R. Peres, Phys. Rev. A **101**, 033817 (2020).
DOI 10.1103/PhysRevA.101.033817
261. A. Klein, D.L. Maslov, A.V. Chubukov, npj Quant. Mater. **5**, 55 (2020).
DOI 10.1038/s41535-020-0250-4
262. Z.M. Raines, V.I. Fal'ko, L.I. Glazman, Phys. Rev. B **103**, 075422 (2021).
DOI 10.1103/PhysRevB.103.075422
263. I.L. Aleiner, B.I. Shklovskii, Phys. Rev. B **49**, 13721 (1994).
DOI 10.1103/PhysRevB.49.13721
264. V. Andreeva, D.A. Bandurin, M. Luskin, D. Margetis, Phys. Rev. B **102**, 205411 (2020).
DOI 10.1103/PhysRevB.102.205411
265. R. Cohen, M. Goldstein, Phys. Rev. B **98**, 235103 (2018).
DOI 10.1103/PhysRevB.98.235103
266. A. Levchenko, H.Y. Xie, A.V. Andreev, Phys. Rev. B **95**, 121301(R) (2017).
DOI 10.1103/PhysRevB.95.121301
267. S. Li, A. Levchenko, A.V. Andreev, Phys. Rev. B **102**, 075305 (2020).
DOI 10.1103/PhysRevB.102.075305
268. A. Hui, S. Lederer, V. Oganesyan, E.A. Kim, Phys. Rev. B **101**, 121107 (2020).
DOI 10.1103/PhysRevB.101.121107
269. S. Li, M. Khodas, A. Levchenko, Phys. Rev. B **104**, 155305 (2021).
DOI 10.1103/PhysRevB.104.155305
270. J.L.M. Poiseuille, Annales de chimie et de physique (Series 3) **21**, 76 (1847)
271. I. Matthaiakakis, D. Rodríguez Fernández, C. Tutschku, E.M. Hankiewicz, J. Erdmenger, R. Meyer, Phys. Rev. B **101**, 045423 (2020).
DOI 10.1103/PhysRevB.101.045423
272. A.A. Patel, R.A. Davison, A. Levchenko, Phys. Rev. B **96**, 205417 (2017).
DOI 10.1103/PhysRevB.96.205417
273. I. Mandal, A. Lucas, Phys. Rev. B **101**, 045122 (2020).
DOI 10.1103/PhysRevB.101.045122
274. E.I. Rashba, Z.S. Gribnikov, V.Y. Kravchenko, Usp. Fiz. Nauk **119**(5), 3 (1976). [Sov. Phys. Usp. 19, 361 (1976)].
DOI 10.3367/UFNr.0119.197605a.0003
275. S. Li, A. Levchenko, A.V. Andreev, Phys. Rev. B **105**, 125302 (2022).
DOI 10.1103/PhysRevB.105.125302

276. A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, S. Ilani (2021). DOI 10.48550/arXiv.2110.15369
277. O.E. Raichev (2022). DOI 10.48550/arXiv.2202.06623
278. M. Mendoza, H.J. Herrmann, S. Succi, Phys. Rev. Lett. **106**, 156601 (2011). DOI 10.1103/PhysRevLett.106.156601
279. I. Meric, M.Y. Han, A.F. Young, B. Ozylmaz, P. Kim, K.L. Shepard, Nat. Nanotechnol. **3**, 654 (2008). DOI 10.1038/nnano.2008.268
280. P.A. Cherenkov, C. R. Acad. Sci. USSR **2**, 451 (1934). [Dokl. Akad. Nauk SSSR **2**, 451 (1934)]. DOI 10.3367/UFNr.0093.196710n.0385
281. S.I. Vavilov, C. R. Acad. Sci. USSR **2**, 457 (1934). [Dokl. Akad. Nauk SSSR **2**, 457 (1934)]. DOI 10.3367/UFNr.0093.196710m.0383
282. I.E. Tamm, I.M. Frank, C. R. Acad. Sci. USSR **14**, 107 (1937). [Dokl. Akad. Nauk SSSR **14**, 107 (1937)]. DOI 10.3367/UFNr.0093.196710o.0388
283. D. Svintsov, Phys. Rev. B **100**, 195428 (2019). DOI 10.1103/PhysRevB.100.195428
284. D. Svintsov, V. Vyurkov, V. Ryzhii, T. Otsuji, Phys. Rev. B **88**, 245444 (2013). DOI 10.1103/PhysRevB.88.245444
285. T. Zdyrski, J. McGreevy, Phys. Rev. B **99**, 235435 (2019). DOI 10.1103/PhysRevB.99.235435
286. V.E. Dorgan, M.H. Bae, E. Pop, Applied Physics Letters **97**(8), 082112 (2010). DOI 10.1063/1.3483130
287. D. Di Sante, J. Erdmenger, M. Greiter, I. Matthaiakakis, R. Meyer, D. Rodriguez Fernández, R. Thomale, E. van Loon, T. Wehling, Nat. Commun. **11**, 3997 (2020). DOI 10.1038/s41467-020-17663-x
288. R.C.V. Coelho, M. Mendoza, M.M. Doria, H.J. Herrmann, Phys. Rev. B **96**, 184307 (2017). DOI 10.1103/PhysRevB.96.184307
289. O. Furtmaier, M. Mendoza, I. Karlin, S. Succi, H.J. Herrmann, Phys. Rev. B **91**, 085401 (2015). DOI 10.1103/PhysRevB.91.085401
290. Sir W. Thomson, F.R.S., The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **42**(281), 362 (1871). DOI 10.1080/14786447108640585
291. H. Helmholz, Monatsberichte der Königlichen Preussische Akademie der Wissenschaften zu Berlin **23**, 215 (1868)
292. H. Bénard, Les tourbillons cellulaires dans une nappe liquide propageant de la chaleur par convection en régime permanent. Ph.D. thesis, Collège de France (1901)
293. Lord Rayleigh, F.R.S., The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **32**, 529 (1916). DOI 10.1080/14786441608635602
294. L. Bazzanini, A. Gabbana, D. Simeoni, S. Succi, R. Tripiccione, Journal of Computational Science **51**, 101320 (2021). DOI https://doi.org/10.1016/j.jocs.2021.101320
295. C.B. Mendl, A. Lucas, Applied Physics Letters **112**(12), 124101 (2018). DOI 10.1063/1.5022187
296. J. Crabb, X. Cantos-Roman, J.M. Jornet, G.R. Aizin, Phys. Rev. B **104**, 155440 (2021). DOI 10.1103/PhysRevB.104.155440
297. K. Li, Y. Hao, X. Jin, W. Lu, Journal of Physics D: Applied Physics **51**(3), 035104 (2017). DOI 10.1088/1361-6463/aa9cd0
298. C.B. Mendl, M. Polini, A. Lucas, Applied Physics Letters **118**(1), 013105 (2021). DOI 10.1063/5.0030869

299. J.H. Farrel, N. Grisouard, T. Scaffidi (2022). DOI 10.48550/arXiv.2112.07683
300. A. Hui, V. Oganesyan, E.A. Kim, Phys. Rev. B **103**, 235152 (2021). DOI 10.1103/PhysRevB.103.235152
301. D. Bernoulli, *Hydrodynamica, sive de viribus et motibus fluidorum commentarii* (Strasbourg, 1738). DOI 10.3931/e-rara-3911
302. C. Eckart, Phys. Rev. **73**, 68 (1948). DOI 10.1103/PhysRev.73.68
303. Lord Rayleigh, F.R.S., Philosophical Transactions **175**, 1 (1884). DOI 10.1098/rstl.1884.0002
304. J. Maldacena, Adv. Theor. Math. Phys. **2**(2), 231 (1998). DOI 10.4310/ATMP.1998.v2.n2.a1
305. A. Donos, S.A. Hartnoll, Nat. Phys. **9**, 649 (2013). DOI 10.1038/nphys2701
306. S.A. Hartnoll, A. Lucas, S. Sachdev, *Holographic Quantum Matter* (MIT Press, 2018). DOI 10.48550/arXiv.1612.07324
307. T. Schäfer, Annu. Rev. Nucl. Part. Sci. **64**(1), 125 (2014). DOI 10.1146/annurev-nucl-102313-025439
308. J.E. Thomas, Nucl. Phys. A **830**(1-4), 665c (2009). DOI 10.1016/j.nuclphysa.2009.09.055
309. S.A. Hartnoll, Nat. Phys. **11**(1), 54 (2015). DOI 10.1038/nphys3174
310. G.A. Inkof, J.M.C. Küppers, J.M. Link, B. Goutéraux, J. Schmalian, Journal of High Energy Physics **2020**(11), 88 (2020). DOI 10.1007/JHEP11(2020)088
311. J.A.N. Bruin, H. Sakai, R.S. Perry, A.P. Mackenzie, Science **339**(6121), 804 (2013). DOI 10.1126/science.1227612
312. J. Zaanen, Nat. Phys. **9**, 609 (2013). DOI 10.1038/nphys2717
313. R.A. Davison, K. Schalm, J. Zaanen, Phys. Rev. B **89**, 245116 (2014). DOI 10.1103/PhysRevB.89.245116
314. C. Proust, L. Taillefer, Annu. Rev. Cond. Matt. Phys. **10**, 409 (2019). DOI 10.1146/annurev-conmatphys-031218-013210
315. A. Legros, S. Benhabib, W. Tabis, F. Laliberté, M. Dion, M. Lizaire, B. Vignolle, D. Vignolles, H. Raffy, Z.Z. Li, P. Auban-Senzier, N. Doiron-Leyraud, P. Fournier, D. Colson, L. Taillefer, C. Proust, Nat. Phys. **15**(2), 142 (2019). DOI 10.1038/s41567-018-0334-2
316. J. Pagline, R.L. Greene, Nat. Phys. **6**, 645 (2010). DOI 10.1038/nphys1759
317. D.C. Johnston, Adv. Phys. **59**(6), 803 (2010). DOI 10.1080/00018732.2010.513480
318. P. Dai, Rev. Mod. Phys. **87**, 855 (2015). DOI 10.1103/RevModPhys.87.855
319. Q. Si, R. Yu, E. Abrahams, Nat. Rev. Mater. **1**(4), 16017 (2016). DOI 10.1038/natrevmats.2016.17
320. Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, P. Jarillo-Herrero, Nature **556**, 45 (2018). DOI 10.1038/nature23160
321. Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigorda, K. Watanabe, T. Taniguchi, T. Senthil, P. Jarillo-Herrero, Phys. Rev. Lett. **124**(7), 076801 (2020). DOI 10.1103/PhysRevLett.124.076801
322. G.T. Horowitz, V.E. Hubeny, Phys. Rev. D **62**, 024027 (2000). DOI 10.1103/PhysRevD.62.024027
323. D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (CRC Press, 2018). DOI 10.1201/9780429493683
324. S.c.v. Grozdanov, K. Schalm, V. Scopelliti, Phys. Rev. Lett. **120**, 231601 (2018). DOI 10.1103/PhysRevLett.120.231601

325. A. Bagrov, B. Craps, F. Galli, V. Keränen, E. Keski-Vakkuri, J. Zaanen, Phys. Rev. D **97**, 086005 (2018).
DOI 10.1103/PhysRevD.97.086005
326. S. Bhattacharyya, S. Minwalla, J. High Energy Phys. **2009**(09), 034 (2009).
DOI 10.1088/1126-6708/2009/09/034
327. J.M. Deutsch, Phys. Rev. A **43**, 2046 (1991).
DOI 10.1103/PhysRevA.43.2046
328. M. Srednicki, Phys. Rev. E **50**, 888 (1994).
DOI 10.1103/PhysRevE.50.888
329. J. Zaanen, SciPost Phys. **6**(5), 061 (2019).
DOI 10.21468/SciPostPhys.6.5.061
330. A. Lucas, Phys. Rev. Lett. **122**, 216601 (2019).
DOI 10.1103/PhysRevLett.122.216601
331. D. van der Marel, H.J.A. Molegraaf, J. Zaanen, Z. Nussinov, F. Carbone, A. Damascelli, H. Eisaki, M. Greven, P.H. Kes, M. Li, Nature **425**, 271 (2003).
DOI 10.1038/nature01978
332. R.A. Cooper, Y. Wang, B. Vignolle, O.J. Lipscombe, S.M. Hayden, Y. Tanabe, T. Adachi, Y. Koike, M. Nohara, H. Takagi, C. Proust, N.E. Hussey, Science **323**(5914), 603 (2009).
DOI 10.1126/science.1165015
333. L. Delacrétaz, B. Goutéraux, S.A. Hartnoll, A. Karlsson, SciPost Phys. **3**(3), 025 (2017).
DOI 10.21468/SciPostPhys.3.3.025
334. E.W. Huang, C.B. Mendl, S. Liu, S. Johnston, H.C. Jiang, B. Moritz, T.P. Devereaux, Science **358**(6367), 1161 (2017).
DOI 10.1126/science.aak9546
335. T. Andrade, A. Krikun, K. Schalm, J. Zaanen, Nat. Phys. **14**(10), 1049 (2018).
DOI 10.1038/s41567-018-0217-6
336. A. Amoretti, D. Areán, B. Goutéraux, D. Musso, Phys. Rev. Lett. **120**, 171603 (2018).
DOI 10.1103/PhysRevLett.120.171603
337. D.V. Khveshchenko, Lithuan. J. Phys. **61**, 42 (2021).
DOI 10.3952/physics.v61i1.4406
338. J. Erdmenger, I. Matthaiakakis, R. Meyer, D.R. Fernández, Phys. Rev. B **98**, 195143 (2018).
DOI 10.1103/PhysRevB.98.195143
339. R. Peierls, Annalen der Physik **404**(2), 154 (1932). [Ann. Phys. (5) **12**, 154 (1932)].
DOI 10.1002/andp.19324040203
340. L.E. Gurevich, Zh. Eksp. Teor. Fiz. **16**, 193 (1946). [J. Phys. (USSR) **9**, 857 (1945)]
341. L.E. Gurevich, Zh. Eksp. Teor. Fiz. **16**, 416 (1946). [J. Phys. (USSR) **10**, 67 (1946)]
342. Y.G. Gurevich, O.L. Mashkevich, Phys. Rep. **181**, 327 (1989).
DOI 10.1016/0370-1573(89)90011-2
343. C. Fu, T. Scaffidi, J. Waissman, Y. Sun, R. Saha, S.J. Watzman, A.K. Srivastava, G. Li, W. Schnelle, P. Werner, M.E. Kamminga, S. Sachdev, S.S.P. Parkin, S.A. Hartnoll, C. Felser, J. Gooth (2018).
DOI 10.48550/arXiv.1802.09468
344. X. Huang, A. Lucas, Phys. Rev. B **103**, 155128 (2021).
DOI 10.1103/PhysRevB.103.155128
345. A. Levchenko, J. Schmalian, Ann. Phys. **419**, 168218 (2020).
DOI 10.1016/j.aop.2020.168218
346. N. Nandi, T. Scaffidi, P. Kushwaha, S. Khim, M.E. Barber, V. Sunko, F. Mazzola, P.D.C. King, H. Rosner, P.J.W. Moll, M. König, J.E. Moore, S. Hartnoll, A.P. Mackenzie, npj Quant. Mater. **3**(1), 66 (2018).
DOI 10.1038/s41535-018-0138-8
347. C.Q. Cook, A. Lucas, Phys. Rev. B **99**, 235148 (2019).
DOI 10.1103/PhysRevB.99.235148
348. J.M. Link, P.P. Orth, D.E. Sheehy, J. Schmalian, Phys. Rev. B **93**, 235447 (2016).
DOI 10.1103/PhysRevB.93.235447

349. N. Levy, S.A. Burke, K.L. Meaker, M. Panlasigui, A. Zettl, F. Guinea, A.H.C. Neto, M.F. Crommie, Science **329**(5991), 544 (2010). DOI 10.1126/science.1191700
350. K. Mnasri, B. Jeevanesan, J. Schmalian, Phys. Rev. B **92**, 134423 (2015). DOI 10.1103/PhysRevB.92.134423
351. N. Prasai, B.A. Trump, G.G. Marcus, A. Akopyan, S.X. Huang, T.M. McQueen, J.L. Cohn, Phys. Rev. B **95**, 224407 (2017). DOI 10.1103/PhysRevB.95.224407
352. C. Ulloa, A. Tomadin, J. Shan, M. Polini, B.J. van Wees, R.A. Duine, Phys. Rev. Lett. **123**, 117203 (2019). DOI 10.1103/PhysRevLett.123.117203
353. J.F. Rodriguez-Nieva, D. Podolsky, E. Demler (2018). DOI 10.48550/arXiv.1810.12333
354. C. Zu, F. Machado, B. Ye, S. Choi, B. Kobrin, T. Mittiga, S. Hsieh, P. Bhattacharyya, M. Markham, D. Twitchen, A. Jarmola, D. Budker, C.R. Laumann, J.E. Moore, N.Y. Yao, Nature **597**, 45 (2021). DOI 10.1038/s41586-021-03763-1
355. E. Ilievski, J. De Nardis, B. Wouters, J.S. Caux, F.H.L. Essler, T. Prosen, Phys. Rev. Lett. **115**, 157201 (2015). DOI 10.1103/PhysRevLett.115.157201
356. T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I.E. Mazets, T. Gasenzer, J. Schmiedmayer, Science **348**(6231), 207 (2015). DOI 10.1126/science.1257026
357. O.A. Castro-Alvaredo, B. Doyon, T. Yoshimura, Phys. Rev. X **6**, 041065 (2016). DOI 10.1103/PhysRevX.6.041065
358. B. Bertini, M. Collura, J. De Nardis, M. Fagotti, Phys. Rev. Lett. **117**, 207201 (2016). DOI 10.1103/PhysRevLett.117.207201
359. B. Bertini, F. Heidrich-Meisner, C. Karrasch, T. Prosen, R. Steinigeweg, M. Žnidarič, Rev. Mod. Phys. **93**, 025003 (2021). DOI 10.1103/RevModPhys.93.025003
360. J.S. Caux, B. Doyon, J. Dubail, R. Konik, T. Yoshimura, SciPost Phys. **6**, 70 (2019). DOI 10.21468/SciPostPhys.6.6.070
361. M. Schemmer, I. Bouchoule, B. Doyon, J. Dubail, Phys. Rev. Lett. **122**, 090601 (2019). DOI 10.1103/PhysRevLett.122.090601
362. J. Lopez-Piqueres, B. Ware, S. Gopalakrishnan, R. Vasseur, Phys. Rev. B **103**, L060302 (2021). DOI 10.1103/PhysRevB.103.L060302
363. Y. Tserkovnyak, J. Appl. Phys. **124**(19), 190901 (2018). DOI 10.1063/1.5054123
364. John Kerr, LL.D., The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **3**(19), 321 (1877). DOI 10.1080/14786447708639245
365. T. Van Mechelen, W. Sun, Z. Jacob, Nat. Commun. **12**, 4729 (2021). DOI 10.1038/s41467-021-25097-2
366. Y. Tserkovnyak, J. Zou, Phys. Rev. Res. **1**, 033071 (2019). DOI 10.1103/PhysRevResearch.1.033071
367. V. Galitski, M. Kargarian, S. Syzranov, Phys. Rev. Lett. **121**, 176603 (2018). DOI 10.1103/PhysRevLett.121.176603
368. K. Hattori, Y. Hirono, H.U. Yee, Y. Yin, Phys. Rev. D **100**, 065023 (2019). DOI 10.1103/PhysRevD.100.065023
369. C. Copetti, K. Landsteiner, Phys. Rev. B **99**, 195146 (2019). DOI 10.1103/PhysRevB.99.195146
370. R. Toshio, K. Takasan, N. Kawakami, Phys. Rev. Research **2**, 032021 (2020). DOI 10.1103/PhysRevResearch.2.032021
371. E.H. Hasdeo, J. Ekström, E.G. Idrisov, T.L. Schmidt, Phys. Rev. B **103**, 125106 (2021). DOI 10.1103/PhysRevB.103.125106
372. M. Moore, P. Surówka, V. Juričić, B. Roy, Phys. Rev. B **101**, 161111 (2020). DOI 10.1103/PhysRevB.101.161111

373. N.N. Fimin, V.M. Chechetkin, Comput. Math. Math. Phys. **58**, 449 (2018).
DOI 10.1134/S0965542518030053
374. A.G. Aksenov, A.V. Babakov, V.M. Chechetkin, Comput. Math. Math. Phys. **58**, 1287 (2018).
DOI 10.1134/S096554251808002X
375. I.L. Aleiner, B.L. Altshuler, M.E. Gershenson, Waves in Random Media **9**(2), 201 (1999).
DOI 10.1088/0959-7174/9/2/308
376. B.L. Altshuler, P.A. Lee, R.A. Webb (eds.), *Mesoscopic Phenomena in Solids* (North-Holland, New York, 1991)
377. B.N. Narozhny, I.L. Aleiner, Phys. Rev. Lett. **84**, 5383 (2000).
DOI 10.1103/PhysRevLett.84.5383
378. B.N. Narozhny, I.L. Aleiner, A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).
DOI 10.1103/PhysRevLett.86.3610
379. A.K. Geim, I.V. Grigorieva, Nature **499**, 419 (2013).
DOI 10.1038/nature12385
380. E.G. Mishchenko, B.I. Halperin, Phys. Rev. B **68**, 045317 (2003).
DOI 10.1103/PhysRevB.68.045317
381. T.S. Nunner, N.A. Sinitsyn, M.F. Borunda, V.K. Dugaev, A.A. Kovalev, A. Abanov, C. Timm, T. Jungwirth, J.i. Inoue, A.H. MacDonald, J. Sinova, Phys. Rev. B **76**, 235312 (2007).
DOI 10.1103/PhysRevB.76.235312
382. N.A. Sinitsyn, Journal of Physics: Condensed Matter **20**, 023201 (2008).
DOI 10.1088/0953-8984/20/02/023201
383. I. Žutić, J. Fabian, S. Das Sarma, Rev. Mod. Phys. **76**, 323 (2004).
DOI 10.1103/RevModPhys.76.323
384. J. Sinova, S.O. Valenzuela, J. Wunderlich, C.H. Back, T. Jungwirth, Rev. Mod. Phys. **87**, 1213 (2015).
DOI 10.1103/RevModPhys.87.1213

Anti-Poiseuille flow in neutral graphene

B.N. Narozhny,[1, 2] I.V. Gornyi,[3, 1, 4] and M. Titov[5]
[1]*Institut for Theoretical Condensed Matter Physics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[2]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[3]*Institut for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[4]*Ioffe Institute, 194021 St. Petersburg, Russia*
[5]*Radboud University Nijmegen, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*
(Dated: August 26, 2021)

Hydrodynamic flow of charge carriers in graphene is an energy flow unlike the usual mass flow in conventional fluids. In neutral graphene, the energy flow is decoupled from the electric current, making it difficult to observe the hydrodynamic effects and measure the viscosity of the electronic fluid by means of electric current measurements. In particular, we show that the hallmark Poiseuille flow in a narrow channel cannot be driven by the electric field irrespective of boundary conditions at the channel edges. Nevertheless one can observe nonuniform current densities similarly to the case of the well-known ballistic-diffusive crossover. The standard diffusive behavior with the uniform current density across the channel is achieved under the assumptions of specular scattering on the channel boundaries. This flow can also be made nonuniform by applying weak magnetic fields. In this case, the curvature of the current density profile is determined by the quasiparticle recombination processes dominated by the disorder-assisted electron-phonon scattering – the so-called supercollisions.

Electronic hydrodynamics has attracted substantial experimental and theoretical attention in recent years [1–3]. Hydrodynamic flows in two-dimensional (2D) materials can now be observed directly using several imaging techniques [4–14]. Two of these experiments [10, 11] were focusing on the Poiseuille flow, the simplest manifestation of viscous hydrodynamics in conventional fluids [15].

The Poiseuille flow [15–17] is a pressure-induced flow in a pipe or between parallel plates. The latter is equivalent to a 2D flow in a narrow channel (with the length L much greater than the width W). In the middle of the channel (away from both of its ends) the flow velocity is directed along the channel and depends only on the transverse coordinate. In that case, the hydrodynamic equations admit a simple solution with the parabolic velocity profile and the flow rate (discharge) that is proportional to the third power of the channel width (for a three-dimensional flow through a pipe – the fourth power of the radius, which is especially important in hematology [18]).

The possibility for an electronic system to exhibit the Poiseuille flow in a narrow wire was first pointed out by Gurzhi [19–21]. Recently, similar behavior has been a subject of intense theoretical [22–33] and experimental [10–13, 22, 34–44] research in the context of electronic transport in high-mobility 2D materials. In contrast to conventional fluids, the electronic flow is affected not only by viscous effects, but also by weak disorder scattering and is characterized by a typical length scale known as the Gurzhi length [26–29, 33]

$$\ell_G = \sqrt{\nu \tau_{\rm dis}}. \tag{1}$$

Here ν is the kinematic viscosity [3, 15, 45–47] and $\tau_{\rm dis}$ is the disorder mean free time. The resulting current profile is given by the catenary curve approaching the parabola in the limit $\ell_G \gg W$.

Nonuniform hydrodynamic flow in a narrow channel has to be contrasted with a conventional ballistic flow that in the case of realistic boundary conditions [10, 48] can also be nonuniform. Assuming rough edges, where electrons scatter off in all directions with equal probability ("diffusive scattering"), bulk impurity scattering competes with boundary effects leading to a ballistic-diffusive crossover. If the mean free path is much smaller than the channel width, $\ell_{\rm dis} \ll W$, then the electric current density is uniform, except for the small regions close to the edges. Reducing the channel width leads to the appearance of a curved current profile that is visually similar to the Poiseuille flow (with the maximum curvature corresponding to both length scales being of the same order of magnitude). In doped graphene this was observed in the recent imaging experiment [10].

Physics of neutral graphene [11, 49, 50] is more intricate. Here the electronic system is nondegenerate and both graphene bands contribute to transport on equal footing. Due to linearity of the Dirac spectrum, the Auger processes are kinematically suppressed and to the leading approximation the number of particles in each band is conserved independently [2, 3, 51, 52]. Another consequence of the peculiar kinematics of Dirac fermions in graphene is the so-called "collinear scattering singularity" [52–59] that gives rise to the "three-mode approximation" allowing one to solve the kinetic equation and derive the hydrodynamic theory [59–61]. The key feature of the resulting description is that the hydrodynamic flow in graphene is the *flow of energy* rather than mass in conventional fluids or charge in Ohmic conductors [2, 3, 60, 61]. Precisely at charge neutrality and in the absence of external magnetic field, the hydrodynamic energy flow is completely decoupled from the electric current. In an infinite system the latter exhibits usual Ohmic behavior with the dominant contribution to the mean free path coming from electron-electron interaction

[50, 54, 55, 60–63]. It is then reasonable to expect that in a narrow channel this current should exhibit the above ballistic-diffusive crossover with the only difference being the microscopic nature of the mean free path.

Hydrodynamic flows in neutral graphene were recently studied experimentally with the help of nanoscale magnetic imaging [11]. The authors reported measurements of inhomogeneous electric current density interpreting them in terms of the Poiseuille flow. Assuming that the curvature of the current density profile was determined by viscosity, the authors proceeded to extract the shear viscosity in graphene at and close to charge neutrality. The resulting values appeared to be in a surprisingly good agreement with the theoretical calculations of Ref. [47].

What exactly is the Poiseuille flow and can it be used as a hallmark of hydrodynamic behavior? The Poiseuille flow is a particular solution to the Navier-Stokes equation [15] in the case where a viscous, incompressible fluid is constrained by (straight and infinitely long) stationary boundaries. The problem is usually solved under the assumption of the so-called *no-slip boundary conditions*, i.e. the vanishing flow velocity at the boundaries. Then the Navier-Stokes equation becomes an ordinary second-order differential equation yielding the standard parabolic velocity profile. The solution can be extended to the case of more general Maxwell's boundary conditions [64] with a finite slip length [65]. The limit of the infinite slip length however (i.e., with no-stress boundary conditions) does not admit any solutions for the Poiseuille problem. In other words: *a pressure-induced viscous flow in a pipe cannot be homogeneous.* On the contrary, an inviscid fluid is described by the Euler equation [15], which is a nonlinear, first-order differential equation. As such, it does not require boundary conditions on the longitudinal (along the boundary) component of the velocity and allows for homogeneous solutions. Hence, the Poiseuille flow can be used as a hallmark of viscosity.

Adapting the above arguments to electronic transport is straightforward for single-band, Fermi-liquid-like systems, such as doped graphene. Here all physical quantities are determined by the Fermi energy, all macroscopic currents are physically equivalent and can be represented by a single vector quantity, the velocity $\boldsymbol{u}$. In the "hydrodynamic regime", i.e., if the electron-electron interaction is the dominant scattering mechanism in the problem, $\ell_{ee} \ll \ell_{\mathrm{dis}}, \ell_{e-ph}, W$ (in the self-evident notation), $\boldsymbol{u}$ obeys a Navier-Stokes-like equation [2, 3, 61] and may exhibit a Poiseuille-like behavior in a channel [10].

Electronic hydrodynamics in graphene – In a two-band system the situation is more involved. An out-of-equilibrium (current-carrying) state may be characterized either by the chemical potentials $\mu_\pm$ of each band, or by their linear combinations [51, 61]

$$\mu = (\mu_+ + \mu_-)/2, \qquad \mu_I = (\mu_+ - \mu_-)/2, \tag{2a}$$

conjugate to the charge and imbalance densities

$$n = n_+ - n_-, \qquad n_I = n_+ + n_-. \tag{2b}$$

In equilibrium $\mu_I = 0$. Although macroscopic currents are no longer equivalent [51, 59–61], one can still introduce the hydrodynamic velocity associating it with one (nearly) conserved current, namely the momentum flux. In the case of linear spectrum, the momentum flux is equivalent to the energy current. As a result, the electric ($\boldsymbol{j}$), quasiparticle (or "imbalance", $\boldsymbol{j}_I$), and energy ($\boldsymbol{j}_E$) currents in graphene can be defined as [2, 3, 61]

$$\boldsymbol{j} = n\boldsymbol{u} + \delta\boldsymbol{j}, \quad \boldsymbol{j}_I = n_I\boldsymbol{u} + \delta\boldsymbol{j}_I, \quad \boldsymbol{j}_E = \mathcal{W}\boldsymbol{u}, \tag{3}$$

where $\mathcal{W}$ is the enthalpy density and $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ are the dissipative corrections, see Eqs. (7) below and the Appendix. In the degenerate limit $\mu \gg T$ the dissipative corrections vanish [61, 63] justifying the applicability of the above single-band picture to doped graphene. At charge neutrality $n = 0$, the electric and energy currents in Eq. (3) appear to be decoupled [61].

The quasiparticle currents $\boldsymbol{j}$ and $\boldsymbol{j}_I$ satisfy the continuity equations [2, 3, 61, 66]

$$\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0, \tag{4a}$$

$$\partial_t n_I + \boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\frac{n_I - n_{I,0}}{\tau_R} = -\frac{12\ln 2}{\pi^2}\frac{n_{I,0}\mu_I}{T\tau_R}, \tag{4b}$$

where $n_{I,0} = \pi T^2/(3v_g^2)$ is the equilibrium value of the total quasiparticle density (i.e., at $\mu_I = 0$) and τ_R is the recombination time [66, 67]. The hydrodynamic velocity $\boldsymbol{u}$ satisfies the generalized Navier-Stokes equation [61]

$$\mathcal{W}(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} + v_g^2\boldsymbol{\nabla}P + \boldsymbol{u}\partial_t P + e(\boldsymbol{E}\cdot\boldsymbol{j})\boldsymbol{u} = \tag{4c}$$

$$= v_g^2\left[\eta\Delta\boldsymbol{u} + en\boldsymbol{E} + \frac{e}{c}\boldsymbol{j}\times\boldsymbol{B}\right] - \mathcal{W}\boldsymbol{u}/\tau_{\mathrm{dis}},$$

where P and η are the thermodynamic pressure and shear viscosity. The full hydrodynamic equations [51, 68] also includes the thermal transport equation [66]

$$T\left[\frac{\partial s}{\partial t} + \boldsymbol{\nabla}\cdot\left(s\boldsymbol{u} - \delta\boldsymbol{j}\frac{\mu}{T} - \delta\boldsymbol{j}_I\frac{\mu_I}{T}\right)\right] = \tag{4d}$$

$$= \delta\boldsymbol{j}\cdot\left[e\boldsymbol{E} + \frac{e}{c}\boldsymbol{u}\times\boldsymbol{B} - T\boldsymbol{\nabla}\frac{\mu}{T}\right] - T\delta\boldsymbol{j}_I\cdot\boldsymbol{\nabla}\frac{\mu_I}{T}$$

$$+\frac{\eta}{2}\left(\nabla_\alpha u_\beta + \nabla_\beta u_\alpha - \delta_{\alpha\beta}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right)^2$$

$$-\frac{n_E - n_{E,0}}{\tau_{RE}} + \mu_I\frac{n_I - n_{I,0}}{\tau_R} + \frac{\mathcal{W}\boldsymbol{u}^2}{v_g^2\tau_{\mathrm{dis}}},$$

which is typically used in hydrodynamics [15] instead of the continuity equation representing energy conservation. Here $n_{E,0}$ denotes the equilibrium value of the energy density similarly to $n_{I,0}$ (i.e., at $\mu_I = 0$) and τ_{RE} is the energy relaxation time (due to, e.g., supercollisions [66]). The last three terms in Eq. (4d) represent energy relaxation, entropy increase due to quasiparticle recombination, and local heating due to impurity scattering.

Consider now linear response transport in the channel geometry (see Refs. [10, 11] for experimental realization) at charge neutrality ($n = 0$) in the steady state. Linearizing the hydrodynamic equations (4), we obtain [66]

$$\boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{5a}$$

$$n_{I,0}\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -(12\ln 2/\pi^2)n_{I,0}\mu_I/(T\tau_R), \tag{5b}$$

$$\boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} + (e/c)\delta\boldsymbol{j}\times\boldsymbol{B} - 3P\boldsymbol{u}/(v_g^2\tau_{\mathrm{dis}}), \tag{5c}$$

$$3P\boldsymbol{\nabla}\cdot\boldsymbol{u} = -2\delta P/\tau_{RE}, \tag{5d}$$

where we have used the "equation of state" [61]

$$\mathcal{W} = 3P = 3n_E/2.$$

Here we follow the standard approach [15] where the thermodynamic quantities are replaced by the corresponding equilibrium functions of the hydrodynamic variables. Equations (5) should be solved for the unknowns $\boldsymbol{u}$, μ_I, and δP keeping the rest of the quantities, e.g., $n_{I,0}$, P, and T, constant (the dissipative corrections $\delta\boldsymbol{j}$, $\delta\boldsymbol{j}_I$ are specified below).

At charge neutrality, the electric field vanishes from the linearized Navier-Stokes equation (5c) and hence cannot drive a hydrodynamic flow.

Channel geometry: absence of the Poiseuille flow in neutral graphene – The channel geometry can be modeled by an "infinite" strip (i.e., with the length of the sample much greater than its width). Transport measurements are assumed to be performed in the two-terminal scheme [10, 11] with the leads placed at the far away ends of the channel. In the middle of the sample, the electric current is flowing along the channel and all physical quantities are independent of the longitudinal coordinate x (this is not true in small regions close to the leads at the ends of the channel). At $n = 0$, the electric current is given by the dissipative correction (y is the transverse coordinate)

$$\boldsymbol{j} = \delta\boldsymbol{j} = \delta j_x(y)\boldsymbol{e}_x, \tag{6a}$$

automatically satisfying the continuity equation (5a). The pressure is also a function of y

$$\delta P = \delta P(y) \quad\Rightarrow\quad \boldsymbol{\nabla}\delta P = \frac{\partial\delta P}{\partial y}\boldsymbol{e}_y, \tag{6b}$$

and similarly

$$\mu_I = \mu_I(y) \quad\Rightarrow\quad \boldsymbol{\nabla}\mu_I = \frac{\partial\mu_I}{\partial y}\boldsymbol{e}_y. \tag{6c}$$

Projecting the Navier-Stokes equation (5c) onto the longitudinal direction, we find

$$\eta\frac{\partial^2 u_x}{\partial y^2} = \frac{3Pu_x}{v_g^2\tau_{\mathrm{dis}}} \quad\Rightarrow\quad u_x = 0. \tag{6d}$$

This is a homogeneous equation that yields the trivial solution $u_x = 0$ for either the no-slip or no-stress boundary conditions. As a result,

$$\boldsymbol{u} = u_y(y)\boldsymbol{e}_y \quad\Rightarrow\quad \boldsymbol{\nabla}\cdot\boldsymbol{u} = \frac{\partial u_y}{\partial y}. \tag{6e}$$

Equations (6) represent the key difference between the usual hydrodynamic flow and electronic transport in neutral graphene. The standard Poiseuille flow is driven by the pressure gradient. In contrast, charge carriers in graphene may be driven by the electric field. At charge neutrality, the field term vanishes from the Navier-Stokes equation leading to the homogeneous equation (6d) for the longitudinal component of the velocity. In other words, *in neutral graphene the Poiseuille flow cannot be driven by the electric field*. Instead, one should apply a temperature gradient along the channel [in this case, the pressure gradient in Eq. (6b) will acquire an x-component contributing a driving term to Eq. (6d)], see also Ref. [69]. We emphasize that this result does not depend on microscopic details of carrier scattering off the channel edges.

What does this mean for the electric current? To clarify this question, we have to specify the dissipative corrections $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$. Their general form was derived in bulk graphene in Refs. [61, 63], see also Appendix. This derivation relied on the specific form of the nonequilibrium correction to the distribution function [see Eq. (A.2) in the Appendix] representing a natural generalization of the usual solution to the kinetic equation in metals [70] to the two-band Dirac system in graphene. In a narrow channel, solutions to the kinetic equation should be subjected to boundary conditions [48] reflecting the nature of the electron scattering off the channel edges. Specifically at charge neutrality, the typical wavelength of Dirac quasiparticles is determined by temperature and thus is much larger than the length scale of the edge roughness that may lead to diffusive boundary scattering [48]. As a result, specular boundary conditions can be expected to adequately describe neutral graphene samples.

In the limit of specular scattering, the distribution function Eq. (A.2) satisfies the boundary conditions and the form of the dissipative corrections remains the same as in the bulk system. At charge neutrality, the corrections are given by

$$\delta\boldsymbol{j} = \frac{1}{e^2\tilde{R}}\left[e\boldsymbol{E} + \omega_B\boldsymbol{e}_B\times\left(\frac{\alpha_1\delta_I\boldsymbol{\nabla}\mu_I}{\tau_{\mathrm{dis}}^{-1}+\delta_I^{-1}\tau_{22}^{-1}} - \frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right)\right], \tag{7a}$$

$$\delta\boldsymbol{j}_I = -\frac{\delta_I}{\tau_{\mathrm{dis}}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{1}{e^2\tilde{R}}\times \tag{7b}$$

$$\times\left[\alpha_1\omega_B\boldsymbol{e}_B\times\boldsymbol{E}+\frac{2T\ln 2}{\pi}e^2R_0\boldsymbol{\nabla}\mu_I+\alpha_1\omega_B^2\frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right],$$

$$\tilde{R} = R_0+\alpha_1^2\delta_I\tilde{R}_B. \tag{7c}$$

Here R_0 [see Eq. (A.14)] is the zero-field bulk resistivity in neutral graphene [56, 59, 61], $\tilde{R}_B \propto \omega_B^2\tau_{\mathrm{dis}}$ is defined in

Eq. (A.15), $\omega_B = eBv_g^2/(2cT\ln 2)$ is the generalized cyclotron frequency (at $\mu = 0$), $\alpha_1 \approx 2.08$ and $\delta_I \approx 0.28$ are detailed in Appendix (v_g is the band velocity in graphene, c is the speed of light, and e is the electron charge). The parameter τ_{22} describes the integrated collision integral, see Eqs. (A.10). Both $\tau_{\rm dis}$ and τ_{22} are functions of the chemical potential and temperature [61, 63, 71].

At $\boldsymbol{B} = 0$, the corrections (7) simplify. The electric current ($e\delta\boldsymbol{j} = \boldsymbol{E}/R_0$) is governed by Ohmic dissipative processes and is independent of the hydrodynamic velocity. Thus, we immediately arrive at the conclusion that in the absence of magnetic field the resulting current density in neutral graphene with specular boundaries is uniform [61, 63] (in contrast to conventional hydrodynamics that does not allow for a stationary pressure-induced flow in a channel without boundary friction [15]).

Nonuniform flows in magnetic field – Now we show that even in the case of specular scattering on the channel boundaries the electric current density can be made nonuniform by applying weak external magnetic field. In the presence of the field all three macroscopic currents are entangled [59] and one may expect a nontrivial solution. The electric current is still flowing along the channel, but is accompanied by the lateral flow of quasiparticles [67, 72]. Since the latter cannot leave the sample, this flow has to vanish at both edges and (nontrivial) homogeneous solutions are no longer allowed. In the two-fluid model of compensated semimetals [28, 72–74] the nontrivial inhomogeneous solution becomes possible due to quasiparticle recombination.

Quasiparticle recombination refers to any scattering process that violates the "approximate" conservation of the number of particles in each individual band including the kinematically suppressed Auger processes, three-particle collisions, scattering by optical phonons [68, 75], and the disorder-assisted electron-phonon coupling (or "supercollisions") [66, 76–80]. The resulting quasiparticle recombination is manifested by an additional term in the continuity equation (4b) for the total quasiparticle ("imbalance") density, first established in Ref. [51] in the context of thermoelectric phenomena. Recently, recombination effects were shown to lead to linear magnetoresistance in compensated semimetals [28, 72, 73, 81], giant magnetodrag [67, 82], and giant nonlocality [74, 83].

Supercollisions involve electron-phonon scattering in a close proximity to an impurity. This is a second-order process where an electron in the upper graphene band may scatter into an empty state in the lower band while emitting a phonon and losing its momentum to the impurity. In the reverse process, the phonon can be absorbed by an electron in the lower band scattering into the upper band (while the impurity compensates the momentum mismatch). Unlike the Auger or three-particle processes, supercollisions also lead to energy relaxation [66]. Taking into account recombination without energy relaxation leads to a problem: the continuity equations for energy and imbalance densities allow only homogeneous solutions, which are incompatible with the boundary conditions at the channel edges. Here we show that energy relaxation due to supercollisions provides the missing piece of the puzzle allowing one to solve the hydrodynamic equations in graphene at charge neutrality. The solution exhibits the inhomogeneous electric current profile in neutral graphene samples with specular reflective boundaries subjected to weak magnetic field. We find that the curvature of the current profile is determined by supercollisions (by means of energy relaxation and quasiparticle recombination) rather than viscosity. A case of rough edges and the corresponding ballistic-diffusive crossover will be discussed elsewhere.

Substituting Eqs. (6) into Eqs. (7) and (5) we find five equations for five unknowns. Excluding δP, μ_I, and δj_x, we are left with two equations for u_y and $\delta j_{I,y}$. For further analysis it is convenient to express them in terms of dimensionless quantities

$$q = \frac{n_{I,0}u_y}{q_0}, \quad p = \frac{\delta j_{I,y}}{q_0}, \quad q_0 = \frac{\omega_B\tau_{\rm dis}E}{e\tilde{R}}, \tag{8}$$

in the matrix form

$$\widehat{L}\begin{pmatrix} q'' \\ p'' \end{pmatrix} = \widehat{M}\begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} \alpha_3 \\ p_0 \end{pmatrix}. \tag{9}$$

The matrix $\widehat{L}$ comprises squares of the recombination-related length scales

$$\widehat{L} = \begin{pmatrix} \ell_{RG}^2 - \ell_{R1}^2 & -\ell_{R1}^2 \\ \ell_{R2}^2 & \ell_{R2}^2 \end{pmatrix}, \tag{10a}$$

$$\ell_{RG}^2 = \frac{1}{2}\ell_{RE}^2 + \frac{2\pi}{9\zeta(3)}\frac{\eta v_g^4\tau_{\rm dis}}{T^3}, \quad \ell_{RE}^2 = v_g^2\tau_{RE}\tau_{\rm dis}, \tag{10b}$$

$$\ell_{R1}^2 = \alpha_1\alpha_3\delta_I\frac{\tilde{R}_B}{2\tilde{R}}\ell_R^2, \quad \ell_R^2 = v_g^2\tau_R\tau_{\rm dis}, \tag{10c}$$

$$\ell_{R2}^2 = \delta_I\frac{R_0}{2\tilde{R}}\frac{\ell_R^2}{1+\tau_{\rm dis}/(\delta_I\tau_{22})}, \tag{10d}$$

while the remaining quantities are dimensionless

$$\alpha_3 = \frac{2\pi^2\ln 2}{27\zeta(3)} \approx 0.42, \quad p_0 = \frac{\alpha_1\delta_I}{1+\tau_{\rm dis}/(\delta_I\tau_{22})}, \tag{10e}$$

$$\widehat{M} = \begin{pmatrix} C_1 & 0 \\ C_2 & 1 \end{pmatrix}, \tag{10f}$$

$$C_1 = \frac{R_0 + \delta R(B)}{\tilde{R}}, \quad C_2 = \frac{12\alpha_1\delta_I\tilde{R}_B\ln^2 2}{\pi^2\tilde{R}}. \tag{10g}$$

The correction $\delta R(B) \propto \omega_B^2\tau_{\rm dis}$ is defined in Eq. (A.20).

Once Eqs. (9) are solved, we can find the electric current (6a) by substituting the solutions $q(y)$ and $p(y)$ into Eq. (7a) using Eqs. (8) and (5b). As a result, we find

$$\delta j_x(y) = \frac{E}{eR_0}\left[1 + \frac{\omega_B^2\tau_{\rm dis}}{e^2\tilde{R}T}\left(\frac{\pi\alpha_1}{2\ln 2}p + \frac{6\ln 2}{\pi}q\right)\right]. \tag{11}$$

We reiterate, that Eq. (11) describes *viscous electronic fluid* in neutral graphene (in contrast to the inviscid system of carriers considered in Ref. [59]).

Anti-Poiseuille flow – Equations similar to Eq. (9) have been solved in Refs. [28, 29, 59, 72, 73] focusing on the resulting magnetoresistance. In this paper, we are interested in the spatial profile of the quasiparticle currents. Requiring the "hard-wall" boundary conditions

$$u_y(\pm W/2) = \delta j_{I,y}(\pm W/2) = 0, \tag{12}$$

we find the solution to Eq. (9) in the form of the catenary curve

$$\begin{pmatrix} q \\ p \end{pmatrix} = \left[\frac{\cosh(\widehat{\mathcal{K}} y)}{\cosh(\widehat{\mathcal{K}} W/2)} - 1\right] \widehat{M}^{-1} \begin{pmatrix} \alpha_3 \\ p_0 \end{pmatrix}, \tag{13}$$

where

$$\widehat{\mathcal{K}}^2 = \widehat{L}^{-1} \widehat{M}.$$

Substituting the result (13) into Eq. (11) we find the electric current profile. The analytical expression for $\delta j_x(y)$ contains a y-independent contribution inherited from the first term in Eq. (11) and the second term in Eq. (13) as well as the catenary terms describing the y dependence of q and p from Eq. (13). Following Ref. [11], we normalize the current by its average value

$$\bar{j_x} = \frac{1}{W} \int\limits_{-W/2}^{W/2} dy\, \delta j_x(y), \tag{14}$$

which can be obtained by averaging the solution (13) and substituting the result into Eq. (11). Averaging of Eq. (13) can be performed in the matrix form yielding

$$\begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix} = \left[\frac{\tanh(\widehat{\mathcal{K}} W/2)}{\widehat{\mathcal{K}} W/2} - 1\right] \widehat{M}^{-1} \begin{pmatrix} \alpha_3 \\ p_0 \end{pmatrix}. \tag{15}$$

The resulting inhomogeneous current density is illustrated in Fig. 1. In some sense, the profile in Fig. 1 can be regarded as "*anti-Poiseuille*": unlike the true Poiseuille flow, this current density exhibits a *minimum* in the center of the channel and is finite at the edges (in fact, there it reaches its maximum). The numerical values of the current density were obtained by using a typical experimental value $\tau_{\text{dis}} \approx 0.8\,\text{THz}$ [50], and assuming the effective coupling constant $\alpha_g \approx 0.2$ following Refs. [50, 84], temperature $T = 250\,\text{K}$, magnetic field $B = 0.1\,\text{T}$, and channel width $W = 1\,\mu\text{m}$. The viscosity affects the current only through the length scale ℓ_{RG}, see Eq. (10b). This effect is rather weak: varying the kinematic viscosity in the range $\nu \approx 0.2 - 0.4\,\text{m}^2/\text{s}$ [47] does not significantly change the results. The recombination length $\ell_R \approx 2\,\mu\text{m}$ and the energy relaxation length $\ell_{RE} \approx 5\,\mu\text{m}$ were chosen phenomenologically, using the data of Ref. [67] as a guide (see also Ref. [66] for theoretical estimates).

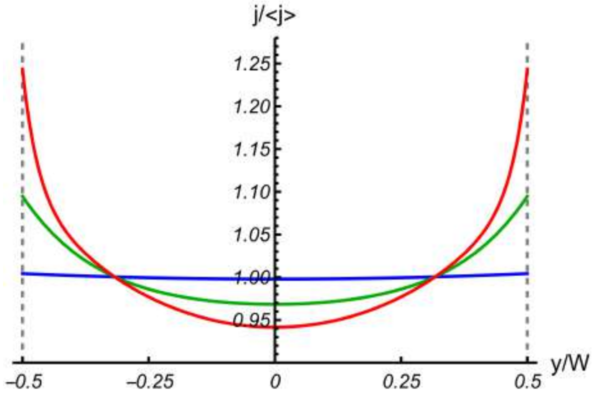


FIG. 1. Catenary curves of the current density in the narrow channel Eq. (11) normalized by the averaged current density Eq. (14). The numerical results were obtained for typical parameter values ($\tau_{\text{dis}} \approx 0.8\,\text{THz}$ [50], $\alpha_g \approx 0.2$ [50, 84], $\nu \approx 0.4\,\text{m}^2/\text{s}$ [11, 47], $B = 0.1\,\text{T}$, $T = 250\,\text{K}$) and correspond to three values of the channel width, $W = 0.1$, 1, 5 μm (blue, green, and red curves, respectively).

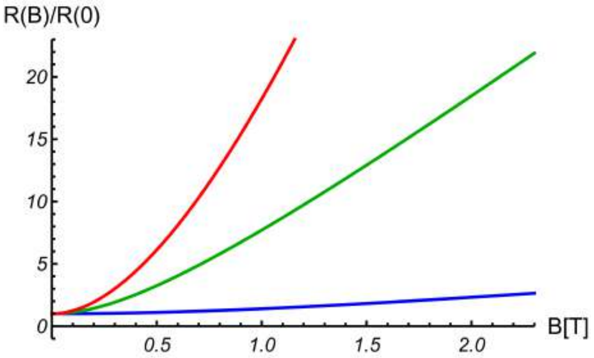


FIG. 2. Magnetoresistance in the narrow channel following from Eqs. (11) and (14) normalized by the zero field resistance R_0. The numerical results were obtained for typical parameter values ($\tau_{\text{dis}} \approx 0.8\,\text{THz}$ [50], $\alpha_g \approx 0.2$ [50, 84], $\nu \approx 0.4\,\text{m}^2/\text{s}$ [11, 47], $B = 0.1\,\text{T}$, $T = 250\,\text{K}$) and correspond to three values of the channel width, $W = 0.1$, 1, 5 μm (blue, green, and red curves, respectively).

Discussion – The results presented in this paper have to be contrasted with recent developments in the field. Most theoretical work on hydrodynamic behavior in neutral (or compensated) materials has been devoted to infinite (or bulk) systems [2, 3, 55, 56, 59, 61]. A bulk system is translationally invariant and hence the current density is uniform with the corresponding sheet resistance given by R_0. In confined geometries the resulting flow profiles are determined by the interplay of sample geometry, boundary conditions, and bulk interaction effects [85]. With respect to electron-electron interaction, three types of theories have been proposed: (i) macroscopic linear response theory of the inviscid electronic fluid [59], (ii) two-fluid hydrodynamics [28, 72–74], and (iii) viscous electronic hydrodynamics that is the subject of the present paper. The difference between the three theories can be summarized as follows. (i) Ref. [59] generalized the standard transport theory (basically the Ohm's law) to graphene close to charge neutrality, where electron-electron interaction contributes to resistivity directly due to lack of Galilean invariance. The resulting theory comprises three (algebraic) equations for three macroscopic currents and does not take into account any possible viscous effects. (ii) The two-fluid model of Ref. [28] assumes that the electron and hole subsystems (i.e. quasiparticles in two different bands) are independently equilibrated and form two separate fluids, while the electron-hole scattering leads to a (weak) friction between the two resembling the drag effect [86]. The theory is described by two sets of hydrodynamic equations, including two Navier-Stokes-like equations. In contrast, (iii) the present hydrodynamic theory [2, 3, 60, 61] assumes that the whole system of charge carriers is equilibrated and is described by a single local equilibrium distribution function leading to the generalized Navier-Stokes equation (4c).

The only theory (out of the above three) yielding the Poiseuille-like flow for the electric current in the channel geometry in the absence of magnetic field is the two-fluid model of Ref. [28], which assumes no-slip boundary conditions for each fluid. Neither the linear response theory of Ref. [59], nor the theory presented in this paper allow for this behavior. The fact that both approaches yield qualitatively similar results (e.g., the absence of the Poiseuille flow and linear magnetoresistance) is quite remarkable since these are two very different theories describing two different systems, one being a (nearly relativistic) viscous fluid and the other being a standard, inviscid (two-band) system of charge carriers. Even though in the latter approach viscosity as a stress-stress correlator [69, 87, 88] might not not necessarily vanish, none of the macroscopic currents satisfy a second-order differential equation of the Navier-Stokes type. It is then rather natural that this approach does not allow for a Poiseuille-like flow. In contrast, the present theory is fully hydrodynamic and hence does in principle yield Poiseuille-like solutions [89]. What we have shown here is that such flows cannot be driven by the electric field leaving the temperature gradient [89] as the only possibility to induce the Poiseuille flow in neutral graphene.

All of the above references agree that in the absence of magnetic field the electric current density is uniform not only in the bulk (infinite) systems, but also in the channel geometry. Based on the arguments presented in this paper, we believe that this intuitively expected conclusion follows from implicit assumptions of either specular boundary conditions or diffusive bulk transport (where one typically neglects narrow regions of inhomogeneity at the sample edges). Here we considered a narrow channel, which is no longer translationally invariant in the lateral direction. In the special case of specular scattering off the boundaries, we find basically the same results: the current density (11) is uniform with R_0 being the resistance. Note, that similarly to the bulk case, R_0 remains finite even in the limit of a completely clean system, $\tau_{\mathrm{dis}} \to \infty$. Once magnetic field is applied, the bulk system exhibits [56, 59] positive, parabolic magnetoresistance $\delta R(B)$, see Eq. (A.20). In contrast, the electronic flow constrained to the narrow channel exhibits linear magnetoresistance [59] in classically strong magnetic fields, see Fig. 2.

Linear magnetoresistance was also discussed in the context of the two-fluid hydrodynamics in Refs. [28, 72–74]. These papers considered a phenomenological model of compensated semimetals where elementary excitations of the conductance and valence bands, i.e. electrons and holes, independently formed hydrodynamic flows, which were only weakly coupled by a mutual friction term. In the language of scattering rates, this model assumed that intraband scattering (characterized by τ_{ee} and τ_{hh} in self-evident notation) was much more effective that interband scattering, such that $\tau_{eh} \gg \tau_{ee}$, τ_{hh}. The zero-field resistance of this model is provided by disorder and intraband scattering, such that even in a clean system ($\tau_{\mathrm{dis}} \to \infty$) the resistance is finite (and is determined by τ_{eh} in a way that is reminiscent of Coulomb drag [67, 86, 90, 91]).

We also stress the importance of boundary conditions on the distribution function. In particular, Ref. [59] considered linear magnetoresistance in a narrow channel, but avoided the issue of the boundary conditions altogether (moreover, energy relaxation was considered purely phenomenologically). Based on the present results, we conclude that the theory presented in Ref. [59] is valid for specular scattering off the channel boundaries. The two-fluid model of Refs. [28, 29, 72, 73] assumed hydrodynamic no-slip boundary conditions for each of the fluids, such that the resulting electric current would vanish at the boundaries. This approach is justified in a different parameter regime from that of the hydrodynamic theory of electronic transport in graphene [2, 3, 61] with a single hydrodynamic flow. Here the electric current comprises both the hydrodynamic and dissipative contributions [63], see Eq. (3). At charge neutrality, the current is decoupled from the hydrodynamic flow and hence the hydrodynamic boundary conditions [65]. Instead, one should consider the kinetics of scattering off the boundaries [48]. In the special case of specular scattering considered in this paper, the nonequilibrium distribution function retains the form (A.2). In the case of diffusive scattering the distribution function is more complicated; in both cases the boundary condition on the distribution function does not easily translate into a boundary condition for electric current: in particular, the electric current is not expected to vanish at the channel boundaries [10]. The alternative no-stress boundary condition [24, 65], that could have been chosen in the two-fluid model of Refs. [28, 29, 72, 73], would not yield the results shown in Fig. 1 as well: then the current density profile would have been flat at the channel boundaries.

Finally, our conclusions should be contrasted with the results of the recent imaging experiment of Ref. [11]. In particular, the vanishing current density at the channel boundaries reported in Ref. [11] are consistent with the hydrodynamic no-slip boundary condition that within our theory is incompatible with the charge flow in neutral graphene. Based on the arguments presented in this paper, as well as our preliminary results for the case of diffusive scattering of the channel boundaries, we expect that bulk recombination processes (most notably, supercollisions) are responsible for the small dip in the current density seen in Ref. [11] in the center of the channel. The overall shape of the current density profile reported in Ref. [11] is consistent with the charge flow under assumptions of the diffusive boundary conditions (to be discussed in a subsequent publication). However, at this time we are not aware of any theoretical argument that would predict precise vanishing of electric current at the channel boundaries (in particular, a recent study of hydrodynamic boundary conditions in graphene [65] reported a nonvanishing slip length). This point appears even more intriguing in view of the recent experiment demonstrating current-carrying edge states in graphene [14], possibly a manifestation of the edge charge accumulation. The latter physics (in particular, the role of such "edge reconstruction" in the hydrodynamic regime) is yet to be addressed in a consistent theoretical fashion. Combining the observations of Ref. [11] and Ref. [14] with the peculiarities of the hydrodynamic approach for neutral graphene remains an important open question.

To conclude, we have discussed electronic transport in graphene at charge neutrality exhibiting a behavior that is strikingly different from any single-component fluid including that in strongly doped graphene. For weak doping ($\mu \ll T$), the hydrodynamic contribution to the electric current Eq. (3) yields a small correction to the results presented in this paper (e.g, the hydrodynamic contribution to optical conductivity in weakly doped graphene was shown [63] to be proportional to μ^2/T^2). For $\mu \sim T$, both the hydrodynamic and dissipative ("kinetic") contributions are of the same order. Now there is no small parameter in the theory and the full system of linearized hydrodynamic equations can be represented by a 6×6 matrix [92]. In the strongly doped regime ($\mu \gg T$), the hydrodynamic contribution dominates and in addition the boundary scattering becomes diffusive. As a result, the electronic flow in a channel exhibits the Poiseuille profile in agreement with the experimental observations in Ref. [10]. Thus we expect the crossover from the anti-Poiseuille to Poiseuille flow to take place at $\mu \sim T$.

Summary – In this paper we have shown that electronic flow in neutral graphene is qualitatively different from that in a conventional viscous fluid. Our main results can be summarized as follows: (i) in response to external electric field, channel-shaped samples of neutral graphene do not exhibit Poiseuille-like flows, while the resulting electric current is independent of viscosity regardless of the choice of the boundary conditions; (ii) for specular boundaries, the electric current density is spatially homogeneous; but (iii) it can be made inhomogeneous by applying the external magnetic field. In the latter case the current profile is anti-Poiseuille, see Fig. 1.

ACKNOWLEDGMENTS

The authors are grateful to P. Alekseev, U. Briskot, I. Burmistrov, A. Dmitriev, V. Gall, V. Kachorovskii, E. Kiselev, A. Mirlin, J. Schmalian, M. Schütt, and A. Shnirman for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 Program (Project 873028 HYDROTRONICS), by the German Research Foundation DFG project NA 1114/5-1 (BNN), by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant No. I-1505-303.10/2019 (IVG) and by the Russian Science Foundation, Grant No. 20-12-00147 (IVG). BNN acknowledges the support by the MEPhI Academic Excellence Project, Contract No. 02.a03.21.0005.

Appendix: Dissipative corrections to macroscopic currents

Within the three-mode approximation [61], the hydrodynamic theory in graphene is formulated in terms of three macroscopic currents (3). In local equilibrium, all three currents are proportional to the hydrodynamic velocity $\boldsymbol{u}$. The effect of electron-electron interaction beyond local equilibrium is captured by the dissipative corrections that can be found following the standard perturbative approach [15]. In the context of electronic hydrodynamics in graphene, the dissipative corrections were derived in Refs. [60, 61, 63]. Here we present a slightly modified approach better suited for the problem at hand.

Let us highlight the main differences between the electronic hydrodynamics in graphene and the conventional hydrodynamics of Galilean-invariant fluids: (i) the band structure of graphene contains two bands touching at the Dirac points leading to the presence of two types of carriers characterized by two quasiparticle currents, $\boldsymbol{j}$ and $\boldsymbol{j}_I$; (ii) neither of the two currents represent the flow of momentum described by the energy current $\boldsymbol{j}_E$; (iii) charge carriers in graphene may scatter off lattice imperfections (impurities), lattice vibrations (phonons), and experience other scattering processes leading to violation of conservation laws including momentum conservation.

Due to the latter issue, the hydrodynamic approach to electronic transport in graphene (as well as any other solid) may be justified only in an intermediate temperature regime, where the electron-electron interaction is the dominant scattering process characterized by the largest relaxation rate or the smallest timescale [2, 3]

$$\tau_{ee} \ll \tau_{\text{dis}}, \tau_R, \text{ etc.}$$

Local equilibrium is formed at the shortest timescales of the order of τ_{ee}. As pointed out in Ref. [59], in graphene this local equilibrium is not equivalent to a steady state since the electron-electron interactions do not relax momentum and hence the hydrodynamic energy flow. To overcome this difficulty one has to take into account weak disorder scattering leading, e.g., to parabolic magnetoresistance [56, 59]. We emphasize that disorder scattering contributes to the hydrodynamic theory already at local equilibrium [61]. Technically this can be understood from the fact that the local equilibrium distribution function does not nullify the disorder collision integral. Similarly, local equilibrium in graphene is affected by electron-phonon scattering [51, 59, 61, 66–68, 75]. Since the lowest-order electron-phonon scattering is kinematically suppressed (within the same valley), the dominant process appears to be the disorder-assisted electron-phonon scattering (or supercollisions) [66, 76]. As compared to the direct impurity scattering, these processes are second-order. Nevertheless, we assume that the mean free time $\tau_{\rm dis}$ includes the (small) contribution of supercollisions as well. The more important effect of supercollisions are the weak decay terms in the continuity equations for the energy and imbalance densities, Eqs. (4d) and (4b) that are characterized by the timescales τ_{RE} and τ_R [66]. Again, these effects appear already at local equilibrium.

Within linear response, the local equilibrium state we have described so far is fully equivalent [61] to the standard transport theory yielding the Ohm's law, classical Hall effect, and – at charge neutrality – positive, parabolic magnetoresistance. As such, the hydrodynamic theory already includes the dissipative processes related to the weak disorder and electron-phonon coupling. This point represents the most important difference between electronic hydrodynamics and conventional fluids, where the ideal flow is always isentropic [15]. In the latter case, dissipative processes (viscosity and thermal conductivity) are attributed to the same interparticle collisions that are responsible for equilibration. By analogy, the effect of electron-electron interaction in electronic hydrodynamics beyond local equilibrium is also described in terms of the "dissipative corrections" to quasiparticle currents (as well as viscosity), the term that might cause confusion (since some dissipation is already taken into account). Moreover, electron-electron interaction does not lead to any further correction to the energy current (since it conserves momentum). It is therefore logical to consider two corrections $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ due to electron-electron interaction instead of three introduced in Ref. [61].

To describe the dissipative processes beyond local equilibrium one introduces a nonequilibrium correction to the local equilibrium distribution function $f^{(0)}_{\lambda\boldsymbol{k}}$ [93]

$$\delta f_{\lambda\boldsymbol{k}} = f_{\lambda\boldsymbol{k}} - f^{(0)}_{\lambda\boldsymbol{k}} = -T\frac{\partial f^{(0)}_{\lambda\boldsymbol{k}}}{\partial\epsilon_{\lambda\boldsymbol{k}}} h_{\lambda\boldsymbol{k}} = f^{(0)}_{\lambda\boldsymbol{k}}\left(1 - f^{(0)}_{\lambda\boldsymbol{k}}\right) h_{\lambda\boldsymbol{k}}, \tag{A.1}$$

where the single-particle states are labeled by the band index $\lambda = \pm$ and the momentum $\boldsymbol{k}$. Taking advantage of the so-called collinear scattering singularity in graphene [52–61], we adopt the "three-mode approximation" [59–61] and write the correction h in the form

$$h_{\lambda\boldsymbol{k}} = \frac{\boldsymbol{v}_{\lambda\boldsymbol{k}}}{v_g}\sum_1^3 \phi_i \boldsymbol{h}^{(i)} + \frac{v^\alpha_{\lambda\boldsymbol{k}} v^\beta_{\lambda\boldsymbol{k}}}{v_g^2}\sum_1^3 \phi_i h^{(i)}_{\alpha\beta} + \ldots, \tag{A.2a}$$

where ... stands for higher-order tensors and the "three modes" are expressed by means of ($\epsilon_{\lambda\boldsymbol{k}}$ denotes the quasiparticle spectrum)

$$\phi_1 = 1, \quad \phi_2 = \lambda, \quad \phi_3 = \epsilon_{\lambda\boldsymbol{k}}/T. \tag{A.2b}$$

The first term in h is responsible for dissipative corrections to the currents, the second term – for viscosity [61].

The coefficients $\boldsymbol{h}^{(i)}$ and $h^{(i)}_{\alpha\beta}$ in Eq. (A.2a) satisfy general constraints [93] reflecting the postulate that electron-electron collisions should not alter conserved thermodynamic quantities. To maintain conservation of the number of particles and energy one sets [60, 61]

$$\operatorname{Tr} h^{(i)}_{\alpha\beta} = 0. \tag{A.2c}$$

To maintain momentum conservation, we require that any correction to the energy current due to the nonequilibrium correction (A.1) should vanish leading to

$$\boldsymbol{h}^{(3)} = -\frac{2T}{3n_E}\left(n\boldsymbol{h}^{(1)} + n_I\boldsymbol{h}^{(2)}\right), \tag{A.2d}$$

following from the linear correspondence between the coefficients $\boldsymbol{h}^{(i)}$ and the corrections to the currents [60, 61]

$$\begin{pmatrix} \delta\boldsymbol{j} \\ \delta\boldsymbol{j}_I \\ \delta\boldsymbol{j}_E/T \end{pmatrix} = \frac{v_g T}{2}\widehat{M}_h \begin{pmatrix} \boldsymbol{h}^{(1)} \\ \boldsymbol{h}^{(2)} \\ \boldsymbol{h}^{(3)} \end{pmatrix}, \tag{A.3}$$

where

$$\widehat{M}_h = \begin{pmatrix} \frac{\partial n}{\partial\mu} & \frac{\partial n_I}{\partial\mu} & \frac{2n}{T} \\ \frac{\partial n_I}{\partial\mu} & \frac{\partial n}{\partial\mu} & \frac{2n_I}{T} \\ \frac{2n}{T} & \frac{2n_I}{T} & \frac{3n_E}{T^2} \end{pmatrix}. \tag{A.4}$$

Enforcing the constraint (A.2d) we find $\delta\boldsymbol{j}_E = 0$, while for the remaining two dissipative corrections we obtain

$$\delta\boldsymbol{j} = \frac{v_g T}{2}\left[\left(\frac{\partial n}{\partial\mu} - \frac{4n^2}{3n_E}\right)\boldsymbol{h}^{(1)} + \left(\frac{\partial n_I}{\partial\mu} - \frac{4nn_I}{3n_E}\right)\boldsymbol{h}^{(2)}\right], \tag{A.5a}$$

$$\delta\boldsymbol{j}_I = \frac{v_g T}{2}\left[\left(\frac{\partial n_I}{\partial\mu} - \frac{4nn_I}{3n_E}\right)\boldsymbol{h}^{(1)} + \left(\frac{\partial n}{\partial\mu} - \frac{4n_I^2}{3n_E}\right)\boldsymbol{h}^{(2)}\right], \tag{A.5b}$$

At charge neutrality these expressions simplify to

$$\delta\boldsymbol{j} = \frac{v_g T}{2}\frac{\partial n}{\partial\mu}\boldsymbol{h}^{(1)}, \tag{A.6a}$$

$$\delta\boldsymbol{j}_I = \frac{v_g T}{2}\frac{\partial n}{\partial\mu}\delta_I\boldsymbol{h}^{(2)}, \tag{A.6b}$$

where

$$\delta_I = 1 - \frac{\pi^4}{162\zeta(3)\ln 2} \approx 0.28, \tag{A.6c}$$

and $\zeta(z)$ is the Riemann's zeta function.

The approach described so far is fully justified in bulk (or infinite) systems where one may assume rotational invariance. In contrast, if the electronic system is confined to a narrow channel, then the specific form of the nonequilibrium distribution function (A.2) cannot be assumed on symmetry grounds. Instead, one should solve the kinetic equation in the presence of the boundaries imposing proper boundary conditions on the distribution function reflecting physical assumptions of the nature of electron scattering off the channel boundaries [48]. In the case of specular scattering, the distribution function satisfies

$$f(\pm W/2, \varphi) = f(\pm W/2, -\varphi), \tag{A.7}$$

where φ is the angle between the velocity $\boldsymbol{v}_{\lambda\boldsymbol{k}}$ and the boundary (i.e., the direction along the channel). One can easy convince oneself that the first term in Eq. (A.2a) satisfies this condition. Indeed, the vectors $\boldsymbol{h}^{(1,2)}$ are linear combinations of the currents $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$, see Eqs. (A.5). The electric current $\delta\boldsymbol{j}$ has only a component along the channel, see Eq. (6a), while the lateral component of the imbalance current vanishes at the boundary, see Eq. (12). Precisely at the boundary, the angular dependence of the first term in Eq. (A.2a) is therefore

$$h \propto \cos\varphi.$$

Similarly, the lateral component of the hydrodynamic velocity $\boldsymbol{u}$ vanishes at the boundary, see Eq. (12), such that the product $\boldsymbol{u}{\cdot}\boldsymbol{k}$ has the same angular dependence (recall that both velocity and momentum have the same direction). As a result, at the boundary the full distribution function depends on $\cos\varphi$ only, thus satisfying Eq. (A.7).

The nonequilibrium correction to the distribution function can be found using the standard iterative solution of the kinetic equation [93]. In the context of the three-mode approximation in graphene, we may solve the kinetic equation directly in terms of the dissipative corrections (A.5) by integrating the kinetic equation to obtain the macroscopic equations for the quasiparticle currents. The iterative procedure is implemented by using the local equilibrium distribution function in the left-hand side of the kinetic equation, while retaining the nonequilibrium correction in the right-hand side to the linear order [60, 61]. At charge neutrality, the resulting equations have the form [61]

$$-\frac{v_g^2}{2}\frac{\partial n}{\partial \mu} e\boldsymbol{E} + \omega_B \boldsymbol{e}_B \times \boldsymbol{\mathcal{K}} = \boldsymbol{\mathcal{I}}_1, \tag{A.8a}$$

$$\frac{v_g^2}{2}\boldsymbol{\nabla} n_I - \frac{v_g^2 n_I}{3n_E}\boldsymbol{\nabla} n_E + \frac{2ev_g^2 n_I}{3cn_E}\delta\boldsymbol{j}\times\boldsymbol{B} + \omega_B \boldsymbol{e}_B \times \boldsymbol{\mathcal{K}}_I = \boldsymbol{\mathcal{I}}_2, \tag{A.8b}$$

where the Lorentz terms are given by [61]

$$\boldsymbol{\mathcal{K}}(\mu=0) = (T\ln 2)\frac{\partial n}{\partial \mu}\boldsymbol{u} + \alpha_1 \delta\boldsymbol{j}_I, \tag{A.9a}$$

$$\alpha_1 = \frac{1-\alpha_3}{\delta_I} \approx 2.08, \quad \alpha_3 = \frac{4n_I T\ln 2}{3n_E} = \frac{2\pi^2 \ln 2}{27\zeta(3)}, \tag{A.9b}$$

$$\boldsymbol{\mathcal{K}}_I(\mu=0) = \delta\boldsymbol{j}. \tag{A.9c}$$

The integrated collision integrals due to electron-electron interaction $\boldsymbol{\mathcal{I}}_i$ were discussed in detail in Refs. [61, 63]. At charge neutrality

$$\boldsymbol{\mathcal{I}}_1(\mu=0) = -\left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\rm dis}}\right)\delta\boldsymbol{j}, \tag{A.10a}$$

$$\boldsymbol{\mathcal{I}}_2(\mu=0) = -\left(\frac{1}{\delta_I \tau_{22}} + \frac{1}{\tau_{\rm dis}}\right)\delta\boldsymbol{j}_I, \tag{A.10b}$$

where the corresponding timescales are determined only by temperature and to the leading order have the form

$$\tau^{-1}_{11(22)}(\mu=0) = \frac{\alpha_g^2 T t^{-1}_{11(22)}}{4\pi\ln 2}, \tag{A.10c}$$

$$t_{11}^{-1} \approx 33.13, \qquad t_{22}^{-1} \approx 5.45, \tag{A.10d}$$

while the integrated collision integral due to impurity scattering is characterized by the timescale [$\tau_{\rm tr}(\epsilon)$ is the transport scattering time]

$$\tau_{\rm dis}^{-1} = -\int d\epsilon \frac{\partial f^{(0)}}{\partial \epsilon}\tau_{\rm tr}^{-1}(\epsilon). \tag{A.10e}$$

In this paper we choose the imbalance chemical potential as a hydrodynamic variable using the relation (at charge neutrality [61])

$$\frac{1}{2}\boldsymbol{\nabla} n_I - \frac{n_I}{3n_E}\boldsymbol{\nabla} n_E = \tag{A.11}$$

$$= \frac{1}{2}\frac{\partial n}{\partial \mu}\left[1 - \frac{4n_I^2}{3n_E}\frac{1}{\partial n/\partial \mu}\right]\boldsymbol{\nabla}\mu_I = \frac{\delta_I}{2}\frac{\partial n}{\partial \mu}\boldsymbol{\nabla}\mu_I.$$

Resolving the equation for the imbalance current, we find

$$\delta\boldsymbol{j}_I = -\frac{\delta_I \frac{v_g^2}{2}\frac{\partial n}{\partial \mu}\boldsymbol{\nabla}\mu_I + \omega_B(1-\alpha_3)\boldsymbol{e}_B\times\delta\boldsymbol{j}}{\tau_{\rm dis}^{-1} + \delta_I^{-1}\tau_{22}^{-1}}. \tag{A.12}$$

Substituting this expression into Eq. (A.8a), we find the dissipative correction to the electric current

$$\delta\boldsymbol{j} = \frac{1}{e^2(R_0 + \alpha_1^2\delta_I \tilde{R}_B)}\Bigg[e\boldsymbol{E} + \tag{A.13}$$

$$+ \frac{\alpha_1\delta_I\omega_B}{\tau_{\rm dis}^{-1} + \delta_I^{-1}\tau_{22}^{-1}}\boldsymbol{e}_B\times\boldsymbol{\nabla}\mu_I - \omega_B\frac{2T\ln 2}{v_g^2}\boldsymbol{e}_B\times\boldsymbol{u}\Bigg],$$

where R_0 denotes the intrinsic resistivity [3, 55] at $\boldsymbol{B}=0$

$$R_0=\frac{\pi}{2\ln 2}\frac{1}{e^2T}\left(\frac{1}{\tau_{11}}+\frac{1}{\tau_{\rm dis}}\right), \tag{A.14}$$

and

$$\tilde{R}_B=\frac{\pi}{2e^2T\ln 2}\frac{\omega_B^2}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}. \tag{A.15}$$

Substituting this result into Eq. (A.12), we find the dissipative correction to the imbalance current

$$\delta\boldsymbol{j}_I=-\frac{\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{1}{e^2(R_0+\alpha_1^2\delta_I\tilde{R}_B)}\times \tag{A.16}$$

$$\times\left[\alpha_1\omega_B\boldsymbol{e}_B\times\boldsymbol{E}+\frac{2T\ln 2}{\pi}e^2R_0\boldsymbol{\nabla}\mu_I+\alpha_1\omega_B^2\frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right].$$

To recover the positive magnetoresistance [56, 59, 61] in bulk graphene, we recall that in an infinite system all currents and densities are uniform. In this case, the generalized Navier-Stokes equation (4c) reduces to

$$0=v_g^2\frac{e}{c}\delta\boldsymbol{j}\times\boldsymbol{B}-\frac{3n_E\boldsymbol{u}}{2\tau_{\rm dis}}, \tag{A.17}$$

which yields the hydrodynamics velocity

$$\boldsymbol{u}=-\omega_B\tau_{\rm dis}\frac{4T\ln 2}{3n_E}\boldsymbol{e}_B\times\delta\boldsymbol{j}. \tag{A.18}$$

Substituting this expression into Eq. (A.13), we find

$$\delta\boldsymbol{j}=\frac{\boldsymbol{E}}{eR_0+e\delta R(B)}, \tag{A.19}$$

where

$$\begin{aligned}\delta R(B)&=\alpha_1^2\delta_I\tilde{R}_B+\frac{8\ln^3 2}{9\zeta(3)}\frac{\pi}{2e^2T\ln 2}\omega_B^2\tau_{\rm dis}\\&=\frac{\omega_B^2\tau_{\rm dis}}{2e^2T\ln 2}\frac{\pi}{9\zeta(3)}\left[1+\frac{9\zeta(3)}{\pi}\frac{\alpha_1^2\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\right]\\&=\mathcal{C}\frac{v_g^4B^2\tau_{\rm dis}}{c^2T^3},\end{aligned} \tag{A.20}$$

with

$$\mathcal{C}\approx\frac{1.71+1.04\frac{\tau_{\rm dis}}{\tau_{22}}}{1+3.59\frac{\tau_{\rm dis}}{\tau_{22}}}\underset{\tau_{\rm dis}\to\infty}{\longrightarrow}\frac{\pi}{9\zeta(3)}\approx 0.29.$$

The positive, parabolic magnetoresistance (A.20) in bulk graphene was previously found in this form in Refs. [59, 61] and in Ref. [56] (where the limiting value of $\mathcal{C}$ was first obtained in the two-mode limit, $\tau_{\rm dis}/\tau_{22}\to\infty$).

[1] M. Polini and A. K. Geim, Physics Today **73**, 28 (2020).

[2] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[3] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).

[4] A. Finkler, Y. Segev, Y. Myasoedov, M. L. Rappaport, L. Ne'eman, D. Vasyukov, E. Zeldov, M. E. Huber, J. Martin, and A. Yacoby, Nano Lett. **10**, 1046 (2010).

[5] D. Halbertal, J. Cuppens, M. B. Shalom, L. Embon, N. Shadmi, Y. Anahory, H. R. Naren, J. Sarkar, A. Uri, Y. Ronen, et al., Nature **539**, 407 (2016).

[6] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, et al., Phys. Rev. B **98**, 241304(R) (2018).

[7] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, et al., Nat. Nanotechnol. **14**, 480 (2019).

[8] A. Marguerite, J. Birkbeck, A. Aharon-Steinberg, D. Halbertal, K. Bagani, I. Marcus, Y. Myasoedov, A. K. Geim, D. J. Perello, and E. Zeldov, Nature **575**, 628 (2019).

[9] A. Uri, Y. Kim, K. Bagani, C. K. Lewandowski, S. Grover, N. Auerbach, E. O. Lachman, Y. Myasoedov, T. Taniguchi, K. Watanabe, et al., Nature Physics **16**, 164 (2020).

[10] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[11] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).

[12] U. Vool, A. Hamo, G. Varnavides, Y. Wang, T. X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C. A. C. Garcia, A. T. Pierce, et al. (2020), arXiv:2009.04477.

[13] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, T. T. K. Watanabe, A. Lucas, A. F. Young, and A. C. B. Jayich (2020), arXiv:2002.05065.

[14] A. Aharon-Steinberg, A. Marguerite, D. J. Perello, K. Bagani, T. Holder, Y. Myasoedov, L. S. Levitov, A. K. Geim, and E. Zeldov, Nature **593**, 528 (2021).

[15] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1959).

[16] J. L. M. Poiseuille, C. R. Acad. Sci. **11**, 961 (1840).

[17] J. L. M. Poiseuille, Annales de chimie et de physique (Series 3) **21**, 76 (1847).

[18] K. Kaushansky, M. A. Lichtman, J. T. Prchal, M. Levi, L. J. Burns, and D. C. Linch, eds., *Williams Hematology* (McGraw Hill, New York, 2021).

[19] R. N. Gurzhi, Zh. Eksp. Teor. Fiz. **44**, 771 (1963), [Sov. Phys. JETP **17**, 521 (1963)].

[20] R. N. Gurzhi, Zh. Eksp. Teor. Fiz. **47**, 1415 (1964), [Sov. Phys. JETP **20**, 953 (1965)].

[21] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968).

[22] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[23] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys.

Rev. B **92**, 165433 (2015).

[24] L. S. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).

[25] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).

[26] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[27] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[28] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).

[29] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).

[30] T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, Phys. Rev. B **100**, 245305 (2019).

[31] I. Matthaiakakis, D. Rodríguez Fernández, C. Tutschku, E. M. Hankiewicz, J. Erdmenger, and R. Meyer, Phys. Rev. B **101**, 045423 (2020).

[32] S. S. Pershoguba, A. F. Young, and L. I. Glazman, Phys. Rev. B **102**, 125404 (2020).

[33] S. Danz and B. N. Narozhny, 2D Materials **7**, 035001 (2020).

[34] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[35] M. Kim, S. G. Xu, A. I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, et al., Nature Communications **11**, 2339 (2020).

[36] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).

[37] G. M. Gusev, A. D. Levin, E. V. Levinson, and A. K. Bakarov, Phys. Rev. B **98**, 161303(R) (2018).

[38] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Scientific Reports **10**, 7860 (2020).

[39] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Phys. Rev. B **103**, 075303 (2021).

[40] O. E. Raichev, G. M. Gusev, A. D. Levin, and A. K. Bakarov, Phys. Rev. B **101**, 235314 (2020).

[41] A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021).

[42] G. Varnavides, A. S. Jermyn, P. Anikeeva, C. Felser, and P. Narang, Nature Communications **11**, 4710 (2020).

[43] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[44] A. Jaoui, B. Fauqué, and K. Behnia, Nature Communications **12**, 195 (2021).

[45] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[46] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[47] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).

[48] C. Beenakker and H. van Houten, in *Semiconductor Heterostructures and Nanostructures*, edited by H. Ehrenreich and D. Turnbull (Academic Press, 1991), vol. 44 of *Solid State Physics*, pp. 1–228.

[49] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, et al., Science **351**, 1058 (2016).

[50] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[51] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[52] A. Tomadin, D. Brida, G. Cerullo, A. C. Ferrari, and M. Polini, Phys. Rev. B **88**, 035430 (2013).

[53] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[54] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[55] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).

[56] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).

[57] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).

[58] M. Schütt, P. M. Ostrovsky, M. Titov, I. V. Gornyi, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. Lett. **110**, 026601 (2013).

[59] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

[60] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).

[61] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[62] Z. Sun, D. N. Basov, and M. M. Fogler, Proceedings of the National Academy of Sciences **115**, 3285 (2018).

[63] B. N. Narozhny, Phys. Rev. B **100**, 115434 (2019).

[64] J. C. Maxwell, Phil. Trans. R. Soc. **170**, 231 (1879).

[65] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).

[66] B. N. Narozhny and I. V. Gornyi, Frontiers in Physics **9**, 108 (2021).

[67] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[68] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[69] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018).

[70] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).

[71] B. N. Narozhny, M. Titov, I. V. Gornyi, and P. M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).

[72] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[73] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).

[74] S. Danz, M. Titov, and B. N. Narozhny, Phys. Rev. B **102**, 081114(R) (2020).

[75] H.-Y. Xie and M. S. Foster, Phys. Rev. B **93**, 195103 (2016).

[76] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).

[77] M. W. Graham, S.-F. Shi, D. C. Ralph, J. Park, and P. L. McEuen, Nature Physics **9**, 103 (2013).

[78] A. C. Betz, S. H. Jhang, E. Pallecchi, R. Ferreira, G. Fève, J.-M. Berroir, and B. Plaçais, Nature Physics **9**, 109 (2013).

[79] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **97**, 085415 (2018).

[80] J. F. Kong, L. Levitov, D. Halbertal, and E. Zeldov, Phys. Rev. B **97**, 245416 (2018).

[81] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).

[82] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, et al., Nat. Phys. **8**, 896 (2012).

[83] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, A. S. Mayorov, M. I. Katsnelson, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levitov, et al., Science **332**, 328 (2011).

[84] A. A. Kozikov, A. K. Savchenko, B. N. Narozhny, and A. V. Shytov, Phys. Rev. B **82**, 075424 (2010).

[85] O. Kashuba, B. Trauzettel, and L. W. Molenkamp, Phys. Rev. B **97**, 205129 (2018).

[86] B. N. Narozhny and A. Levchenko, Rev. Mod. Phys. **88**, 025003 (2016).

[87] B. Bradlyn, M. Goldstein, and N. Read, Phys. Rev. B **86**, 245309 (2012).

[88] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).

[89] J. M. Link, D. E. Sheehy, B. N. Narozhny, and J. Schmalian, Phys. Rev. B **98**, 195103 (2018).

[90] B. N. Narozhny and I. L. Aleiner, Phys. Rev. Lett. **84**, 5383 (2000).

[91] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).

[92] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **103**, 115402 (2021).

[93] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).

arXiv:2206.07414v2 [cond-mat.mes-hall] 29 Aug 2022

Electronic viscosity and energy relaxation in neutral graphene

Vanessa Gall,[1, 2] Boris N. Narozhny,[2, 3] and Igor V. Gornyi[1, 2, 4]
[1] *Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2] *Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[3] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[4] *Ioffe Institute, 194021 St. Petersburg, Russia*

(Dated: August 30, 2022)

We explore hydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. In the absence of magnetic field, the bulk Ohmic charge flow and the hydrodynamic energy flow are decoupled. However, the energy flow does affect the overall resistance of the system through viscous dissipation and energy relaxation that has to be compensated by the work done by the current source. Solving the hydrodynamic equations, we find that local temperature and electric potential are discontinuous at the interfaces with the leads as well as the device resistance and argue that this makes Corbino geometry a feasible choice for an experimental observation of the Dirac fluid.

Quantum dynamics of charge carriers is one of the most important research directions in condensed matter physics. In many materials transport properties can be successfully described under the assumption of weak electron-electron interaction allowing for free-electron theories [1]. An extension of this approach to strongly-correlated systems remains a major unsolved problem. The advent of "ultra-clean" materials poses new challenges, especially if the electronic system is nondegenerate. At high temperatures such systems may exhibit signatures of a collective motion of charge carriers resembling the hydrodynamic flow of a viscous fluid [2–14].

Electronic viscosity has been discussed theoretically for a long time [15–20], but became the subject of dedicated experiments [2, 9] only recently, after ultra-clean materials became available. Up until now, most experimental efforts were focusing on graphene [2–11] where the hydrodynamic regime is apparently easier to achieve [21, 22]. Viscous effects manifest themselves in nonuniform flows. In the common "linear" geometry (channels, wires, Hall bars, etc.) this occurs in "narrow" samples where the typical length scale associated with viscosity is of the same order as the channel width [23–27]. In contrast, in the "circular" Corbino geometry, see Fig. 1, the electric current is nonuniform even in the simplest Drude picture (in the absence of magnetic field, $\boldsymbol{j} \propto \boldsymbol{e_r}/|\boldsymbol{r}|$, where $\boldsymbol{e_r} = \boldsymbol{r}/|\boldsymbol{r}|$) making it an excellent platform to measure electronic viscosity [28–31]. In the last year, electronic hydrodynamics in the Corbino geometry has been studied both experimentally [32] and theoretically [33–36].

In this paper we address the "Dirac fluid" [3, 9] (the hydrodynamic flow of charge carriers in neutral graphene) in the Corbino geometry. Unlike doped graphene where degenerate, Fermi-liquid-like electrons may be described by the Navier-Stokes equation with a weak damping term due to disorder [16, 21, 23], the two-band physics of neutral graphene leads to unconventional hydrodynamics [22, 37]. In the hydrodynamic approach any macroscopic current can be expressed as a product of the corresponding density and hydrodynamic velocity $\boldsymbol{u}$ (up to dissipative corrections), e.g., the electric and energy current densities are $\boldsymbol{j} = n\boldsymbol{u}$ and $\boldsymbol{j}_E = n_E\boldsymbol{u}$, respectively. In the degenerate regime the charge and energy densities are proportional to each other (to the leading approximation in thermal equilibrium $n_E = 2\mu n/3$, where μ is the chemical potential) and the two currents are equivalent [38]. In contrast, the equilibrium charge density vanishes at charge neutrality, $n(\mu = 0) = 0$, while the energy density remains finite. The two currents "decouple": the energy current remains "hydrodynamic", the charge current is completely determined by the dissipative correction $\delta\boldsymbol{j}$.

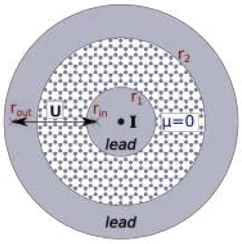


FIG. 1. Corbino geometry: the annulus-shaped sample of neutral graphene ($\mu = 0$) is placed between the the two leads: the inner circle of the radius r_1 and the outer shell with the inner radius r_2. A current I is injected through at the center point and a voltage U is measured between electrodes placed at the inner and outer radius $r_{\rm in}$ and $r_{\rm out}$.

Electronic transport at charge neutrality has been a subject of intensive research [9, 24–27, 38–46] leading to general consensus on the basic result: in the absence of magnetic field, $\boldsymbol{B} = 0$, resistivity of neutral graphene is determined by the electron-electron interaction

$$R_0 = \frac{\pi}{2e^2 T \ln 2}\left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\rm dis}}\right) \underset{\tau_{\rm dis}\to\infty}{\longrightarrow} \frac{1}{\sigma_Q}. \qquad (1)$$

Here $\tau_{11} \propto \alpha_g^{-2} T^{-1}$ describes the appropriate electron-electron collision integral and σ_Q is the "intrinsic" or "quantum" conductivity of graphene. Disorder scattering is characterized by the mean free time $\tau_{\rm dis}$, which is large under the assumptions of the hydrodynamic regime,

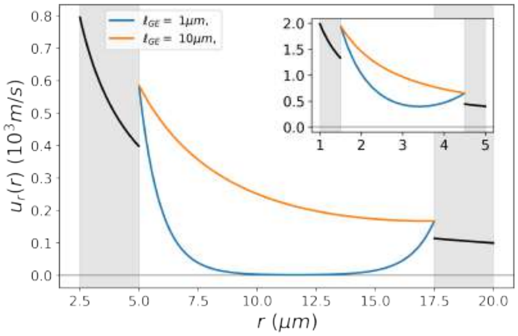


FIG. 2. Radial component of the hydrodynamic velocity u_r. Black lines show the drift velocity in the leads, $u_r^{\rm in(out)} \propto 1/r$. Colored curves correspond to the solution Eq. (4) for the two indicated values of ℓ_{GE}. The results are plotted for the two cases of a large (main panel) and small (inset) device.

$\tau_{\rm dis} \gg \tau_{11}$ and yields a negligible contribution to Eq. (1). Equation (1) describes the uniform bulk current and is independent of viscosity (i.e., in a channel [21, 24, 44, 46]). In contrast, in the Corbino geometry the current flow is necessarily inhomogeneous and hence viscous dissipation must be taken into account.

We envision the following experiment: a graphene sample (at charge neutrality) in the shape of an annulus is placed between the inner (a disk of radius r_1) and outer (a ring with the inner radius r_2) metallic contacts (leads). For simplicity, we assume both leads to be of the same material, e.g., highly doped graphene with the same doping level. The electric current I is injected into the center of the inner lead preserving the rotational invariance (e.g., through a thin vertical wire attached to the center point) and spreads towards the outer lead, which for concreteness we assume to be grounded. The overall voltage drop U is measured between two points in the two leads (at the radii $r_{\rm in} < r_1$ and $r_{\rm out} > r_2$) yielding the device resistance, $R = U/I$. In most traditional measurements, the leads' resistance is minimal, while the contact resistance is important only in ballistic systems, see e.g., Ref. [10]. Hence, one may interpret the measured voltage drop in terms of resistivity of the sample material. Here we focus on the device resistance and show that in the hydrodynamic regime there is an additional contribution due to electronic viscosity and energy relaxation.

Charge flow through the Corbino disk can be described as follows. The injected current spreads through the inner lead according to the Ohm's law and continuity equation. In the stationary case, the latter determines the radial component of the current density, $j_r^{\rm in} = I/(2\pi e r)$. This defines the drift velocity $\boldsymbol{u}^{\rm in} = \boldsymbol{j}^{\rm in}/n^{\rm in}$ ($n^{\rm in}$ is the carrier density in the inner lead) and the energy current $\boldsymbol{j}_E^{\rm in} = n_E^{\rm in}\boldsymbol{u}^{\rm in}$. Reaching the interface, both currents continue to flow into the graphene sample. Here (at $n = 0$ and $\boldsymbol{B} = 0$) the energy current $\boldsymbol{j}_E = n_E\boldsymbol{u}$ is decoupled from the electric current $\boldsymbol{j} = \delta\boldsymbol{j}$. Charge conservation requires the radial component of the electric current to be continuous at the interface, $\delta\boldsymbol{j}(r_1) = \boldsymbol{j}^{\rm in}(r_1)$. Due to the continuity equation, the current density in graphene has the same functional form, $\delta j_r = I/(2\pi e r)$. Does this mean that the device resistance trivially follows if one knows the resistivity of graphene? The answer is "no", since the electrochemical potential is discontinuous at the interface! There are two mechanisms for the "jump" of the potential: (i) the usual Schottky contact resistance [42, 47], and (ii) dissipation due to viscosity [31] and energy relaxation [48]. Since the lost energy must come from the current source, both contribute to R.

The energy flow in neutral graphene is described by the set of hydrodynamic equations developed in Refs. [37, 44, 48] and most recently solved in Ref. [46] in the channel geometry. Within linear response, the equations are

$$\boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{2a}$$

$$n_I\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -(12\ln 2/\pi^2)n_I\mu_I/(T\tau_R), \tag{2b}$$

$$\boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} - 3P\boldsymbol{u}/(v_g^2\tau_{\rm dis}), \tag{2c}$$

$$3P\boldsymbol{\nabla}\cdot\boldsymbol{u} = -2\delta P/\tau_{RE}. \tag{2d}$$

Here Eq. (2a) is the continuity equation; Eq. (2b) is the "imbalance" continuity equation [37, 42] (μ_I is the imbalance chemical potential, $n_I = \pi T^2/(3v_g^2)$ is the equilibrium imbalance density, v_g is the band velocity in graphene, and τ_R is the recombination time); Eq. (2c) is the linearized Navier-Stokes equation [37, 46, 49, 50]; and Eq. (2d) is the linearized "thermal transport" equation (τ_{RE} is the energy relaxation time [48]). Equilibrium thermodynamic quantities (the pressure $P = 3\zeta(3)T^3/(\pi v_g^2)$, enthalpy density $\mathcal{W}$, and energy density are related by the "equation of state", $\mathcal{W} = 3P = 3n_E/2$. The dissipative corrections to the macroscopic currents are given by

$$\delta\boldsymbol{j} = \boldsymbol{E}/(eR_0), \tag{3a}$$

$$\delta\boldsymbol{j}_I = -\frac{2\gamma\ln 2}{\pi}T\tau_{\rm dis}\boldsymbol{\nabla}\mu_I, \quad \gamma = \frac{\delta_I}{1+\tau_{\rm dis}/(\delta_I\tau_{22})}, \tag{3b}$$

where $\tau_{22} \propto \alpha_g^{-2}T^{-1}$ describes a component of the collision integral that is qualitatively similar, but quantitatively distinct from τ_{11} and $\delta_I \approx 0.28$. The equations (2) and (3) should be solved for $\boldsymbol{u}$, $\delta\boldsymbol{j}$, $\delta\boldsymbol{j}_I$, $\boldsymbol{E}$, μ_I, and δP.

Excluding δP from Eqs. (2c) and (2d) we find a second-order differential equation for $\boldsymbol{u}$

$$\eta'\Delta\boldsymbol{u} = 3P\boldsymbol{u}/(v_g^2\tau_{\rm dis}), \quad \eta' = \eta + 3P\tau_{RE}/2. \tag{4a}$$

In the Corbino disk, the general solution for the radial component of the velocity has the form

$$u_r = a_1 I_1\left(\frac{r}{\ell_{GE}}\right) + a_2 K_1\left(\frac{r}{\ell_{GE}}\right), \quad \ell_{GE}^2 = \frac{v_g^2\eta'\tau_{\rm dis}}{3P}, \tag{4b}$$

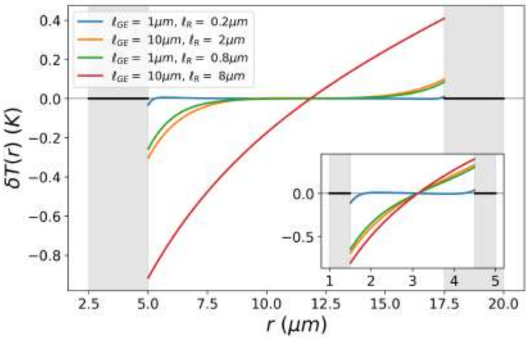


FIG. 3. Temperature distribution in the device. Colored curves correspond to the solution of the hydrodynamic equations for the indicated values of ℓ_{GE} and ℓ_R. The results are plotted for the two cases of a large (main panel) and small (inset) device. In the leads $\delta T = 0$, shown by black lines.

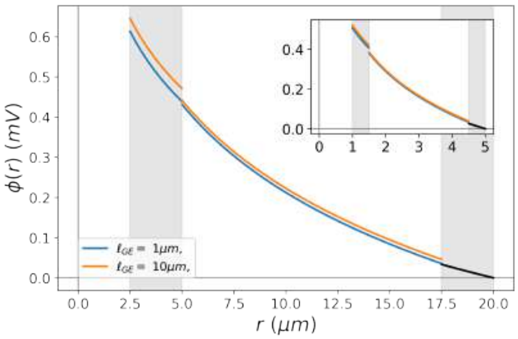


FIG. 4. Electrochemical potential (voltage drop) throughout the device. The black line shows the Ohmic behavior in the outer lead relative to the ground. The jumps at the interfaces are due to dissipative effects (viscosity and energy relaxation) in the bulk of the sample.

where $I_1(z)$ and $K_1(z)$ are the Bessel functions. The coefficients a_1 and a_2 can be found using the continuity of the entropy current at the two interfaces (within linear response). The resulting behavior in shown in Fig. 2 (here we choose to show our results in graphical form since the analytic expressions are somewhat cumbersome [51]; quantitative calculations were performed for $T = 100\,\mathrm{K}$ and experimentally relevant values of the parameters taken from Refs. [8–10, 48]).

In the hydrodynamic regime, the electron-electron scattering time is the shortest scale in the problem, hence the spatial variation of $\boldsymbol{u}$ is determined by energy relaxation. If $\ell_{GE} \ll r_{\rm out} - r_{\rm in}$, then the energy current injected from the leads decays in a (relatively small) boundary region while in the bulk of the sample $\boldsymbol{u} \to 0$. On the other hand, if ℓ_{GE} is of the same order as (or larger than) the system size, then u_r does not vanish and approaches the standard Corbino profile, $u_r \propto 1/r$. At each interface, u_r exhibits a jump due to the mismatch of the entropy densities in the sample and leads.

The nonequilibrium quantities δP and μ_I can now be found straightforwardly. The former follows directly from Eq. (2d) using the solution (4), while the differential equation for the latter can be found by substituting Eq. (3b) into Eq. (2b) and using the solution (4). The boundary conditions for δP and μ_I follow from the continuity equations for the charge and imbalance. The two quantities can be combined to determine the nonequilibrium temperature variation, δT, shown in Fig. 3. For a large sample ($\ell_{GE}, \ell_R \ll r_{\rm out} - r_{\rm in}$, $\ell_R^2 = \gamma v_g^2 \tau_{\rm dis} \tau_R / 2$), δT exhibits fast decay and vanishes in the bulk of the sample. For larger values of ℓ_{GE}, ℓ_R energy relaxation is less effective and the system exhibits an inhomogeneous temperature profile.

The obtained solutions completely describe the hydrodynamic energy flow in neutral graphene. Our remaining task is to find the behavior of the electrochemical potential at the two interfaces enabling us to determine R.

The standard description of interfaces between metals or semiconductors [47] can be carried over to neutral graphene [42] in terms of the contact resistance. Typically, this is a manifestation of the difference of work functions of the two materials across the interface. In graphene, the contact resistance was recently measured in Ref. [10], see also Refs. [32, 52, 53]. In the standard diffusive (or Ohmic) case, the contact resistance leads to a voltage drop that is small compared to the voltage drop in the bulk of the sample and can be ignored. In contrast, in the ballistic case there is almost no voltage drop in the bulk, such that most energy is dissipated at the contacts. Both scenarios neglect electron-electron interactions.

In the diffusive case interactions lead to corrections to the bulk resistivity [54, 55] and the contact resistance can still be ignored. In the ballistic case electron-electron interaction may give rise to a "Knudsen-Poiseuille" crossover [16] and drive the electronic system to the hydrodynamic regime. While the Ohmic resistivity of the electronic fluid may remain small, the hydrodynamic flow possesses another channel for dissipation through viscosity [31]. At charge neutrality, this effect is subtle, since the electric current is decoupled from the hydrodynamic energy flow, see Eq. (3a). At the same time, both are induced by the current source providing the energy dissipated not only by Ohmic effects, but also by viscosity [31] and energy relaxation processes [48] that should be taken into account in the form of an additional voltage drop. Since the voltage drop in the bulk of the sample is completely determined by Eq. (3a), the additional contribution takes the form of a jump in ϕ at the interface corresponding to an excess electric field induced in the thin Knudsen layer around the interface [31].

The magnitude of the jump in ϕ can be established by considering the flow of energy through the interface. Following the standard route [31, 56], we consider the time derivative of the kinetic energy, $\mathcal{A} = \dot{\mathcal{E}}$, where $\mathcal{E}$ is obtained by integrating the energy density $n_E(\boldsymbol{u}) - n_E(0)$

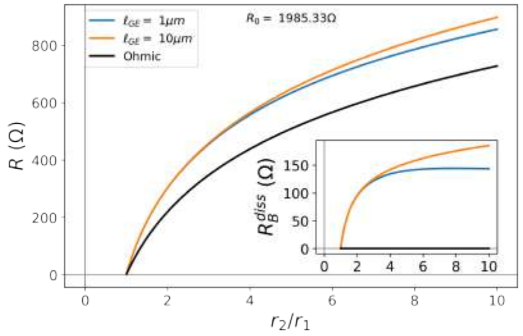


FIG. 5. Total resistance of the Corbino device for different values of ℓ_{GE} (here $r_1 = 0.5\,\mu$m). Inset: additional contribution to the resistance due to viscous dissipation.

over the volume. Working within linear response, we expand the latter to the leading order in the hydrodynamic velocity. Finding time derivatives from the equations of motion and using the continuity equation and partial integration, we then separate the "bulk" and "boundary" contributions, $\mathcal{A} = \mathcal{A}_{\rm bulk} + \mathcal{A}_{\rm edge}$. We interpret the former as the bulk dissipation, while $\mathcal{A}_{\rm edge}$ includes the energy brought in (carried away) through the boundary by the incoming (outgoing) flow. In the stationary state $\dot{\mathcal{E}} = 0$, dissipation is balanced by the work done by the source. Assuming that no energy is accumulated at the interface, we find the corresponding boundary condition.

The specific form of the equations of motion depends on the choice of the material. Assuming the leads' material is highly doped graphene, the equation of motion is the usual Ohm's law with the diffusion term [57] coming from the gradient of the stress-energy tensor [38], here we include a viscous contribution due to disorder [58] and find [31] (omitting the continuous entropy flux)

$$\mathcal{A}^{\rm lead}_{\rm edge} = \int dS_\beta \left(u^L_\alpha \sigma'_{L;\alpha\beta} - u^L_\beta \delta P_L - e\boldsymbol{j}^L_\beta \phi \right), \qquad (5a)$$

where $\boldsymbol{j}^L = n_L \boldsymbol{u}^L$ is the current density, $\boldsymbol{u}^L$ is the drift velocity, δP_L is the nonequilibrium pressure, and σ'_L is the viscous stress tensor in the lead. The first two terms are the usual dissipative contributions to the energy flow across the boundary [56], the last term is the Joule heat.

In neutral graphene, we obtain similar results from the Navier-Stokes equation, except that the Joule heat is now determined by $\delta\boldsymbol{j}$

$$\mathcal{A}^{\rm sample}_{\rm edge} = \int dS_\beta \left(u_\alpha \sigma'_{\alpha\beta} - u_\beta \delta P - e\delta j_\beta \phi \right). \qquad (5b)$$

Equating the two contributions (5) and using the above solutions for the velocity and pressure, we find the jumps of the potential ϕ at the two interfaces. This allows us to determine ϕ everywhere in the device, see Fig. 4, as well as the device resistance.

The total resistance of the Corbino device is shown in Fig. 5. Neglecting hydrodynamic effects, we find the usual logarithmic dependence of R on the system size. Viscosity and energy relaxation provide an additional dissipation channel and hence increase R. Energy relaxation contributes to this increase since it dominates the hydrodynamic energy flow, see Eq. (4). At the same time, the boundary condition for the electric potential, Eqs. (5), is determined by viscosity.

In this paper we have solved the hydrodynamic equations in neutral graphene. We have shown, that despite the known decoupling of the Ohmic charge flow and hydrodynamic energy flow, in Corbino geometry the latter does affect the observable behavior leading to jumps in temperature (shown in Fig. 3) and the electric potential, see Fig. 4. The potential jump is distinct from the usual contact resistance insofar it is a function of the system size. Both effects are observable using the modern imaging techniques (the local temperature variation can be measured using the approach of Refs. [59–61], while measurements of the local potential are at the heart of the technique proposed in Refs. [10, 62]). Hydrodynamics also affects the more conventional transport measurements through the size-dependent contribution to the device resistance, see Fig. 5.

Our results highlight several particular features of the Dirac fluid in neutral graphene. Firstly, the "linear response" currents (3) are independent of the temperature gradient due to exact particle-hole symmetry [42]. Secondly, in contrast to the case of doped graphene [31] *the Dirac fluid is compressible* even within linear response (due to energy relaxation, see Eq. (2d). Finally, the hydrodynamic flow in neutral graphene is the energy flow. Hence, energy relaxation effectively dominates over viscous effects, see Eqs. (4), complicating experimental determination of η.

External magnetic field is also known to couple the charge and energy flows in neutral graphene [37]. We expect that our theory will yield interesting results on Corbino magnetoresistance [52]. Another extension of our theory is the study of thermoelectric phenomena, which is more interesting if one moves away from the neutrality point [34] (where the thermopower must vanish due to the exact particle-hole symmetry). Our results on both issues will be reported elsewhere.

The authors are grateful to P. Hakonen, V. Kachorovskii, A. Levchenko, A. Mirlin, J. Schmalian, A. Shnirman, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 Program (Project 873028 HYDROTRONICS), by the German Research Foundation DFG project NA 1114/5-1 (BNN), and by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant No. I-1505-303.10/2019 (IVG).

Supplemental material

Starting with the general form of the hydrodynamic equations in graphene, we obtain the analytical results presented graphically in the main text. In Sec. we summarize the hydrodynamic equations for graphene. In Sec. we specify these equations within linear response in polar coordinates at charge neutrality and $B = 0$. In Sec. we formulate a description of the leads followed by the relevant boundary conditions at the lead-graphene interfaces in Sec. . Next, in Sec. we present the full analytical solution for the hydrodynamic equations in the Corbino geometry with the above boundary conditions. In Sec. we discuss the dissipation in the system and corroborate the argument used in the main text to obtain the device resistance. Finally, we conclude with a brief analysis in Sec. .

Electronic hydrodynamics in graphene

Following Ref. [63] we combine the chemical potentials of the two bands in graphene $\mu_\pm$ into

$$\mu = (\mu_+ + \mu_-)/2, \quad \mu_I = (\mu_+ - \mu_-)/2 \tag{6}$$

and introduce their conjugate charge and imbalance densities

$$n = n_+ - n_-, \quad n_I = n_+ + n_-. \tag{7}$$

Taking into account dissipative corrections due to electron-electron collisions we then obtain the electric ($\vec{j}$) and imbalance ($\vec{j}_I$) currents as

$$\vec{j} = n\vec{u} + \delta\vec{j}, \quad \vec{j}_I = n_I\vec{u} + \delta\vec{j}_I, \tag{8}$$

where $\vec{u}$ is the drift velocity. The energy current $\vec{j}_E = n_E\vec{u}$ is proportional to the momentum density and is not relaxed by electron-electron collisions. The currents $\vec{j}$ and $\vec{j}_I$ satisfy the continuity equations

$$\partial_t n + \vec{\nabla}\cdot\vec{j} = 0, \tag{9a}$$

which describes the exact conservation of charge and

$$\partial_t n_I + \vec{\nabla}\cdot\vec{j}_I = -\frac{n_I - n_{I,0}}{\tau_R} = -\frac{12\ln 2}{\pi^2}\frac{n_{I,0}\mu_I}{T\tau_R}, \tag{9b}$$

where $n_{I,0} = \pi T^2/(3v_g^2)$ is the equilibrium value of the total quasiparticle density (at $\mu_I = 0$) and τ_R is the recombination time.

A similar equation can be formulated for the energy density

$$\partial_t n_E + \vec{\nabla}\cdot\vec{j}_E = e\vec{j}\vec{E} - \frac{n_E - n_{E,0}}{\tau_{\mathrm{RE}}}, \tag{9c}$$

where τ_{RE} is the energy relaxation time. Typically this is replaced by the thermal transport equation

$$\begin{aligned} T\left[\frac{\partial s}{\partial t} + \vec{\nabla}_{\vec{r}}\cdot\left(s\vec{u} - \delta\vec{j}\frac{\mu}{T} - \delta\vec{j}_I\frac{\mu_I}{T}\right)\right] &= \delta\vec{j}\cdot\left[e\vec{E} + \frac{e}{c}\vec{u}\times\vec{B} - T\vec{\nabla}\frac{\mu}{T}\right] - T\delta\vec{j}_I\cdot\vec{\nabla}\frac{\mu_I}{T} \\ &+ \frac{\eta}{2}\left(\nabla_\alpha u_\beta + \nabla_\beta u_\alpha - \delta_{\alpha\beta}\vec{\nabla}\cdot\vec{u}\right)^2 - \frac{n_E - n_{E,0}}{\tau_{RE}} + \mu_I\frac{n_I - n_{I,0}}{\tau_R} + \frac{\mathcal{W}\vec{u}^2}{v_g^2\tau_{\mathrm{dis}}}. \end{aligned} \tag{9d}$$

Within linear response the two equations coincide. Finally, the generalized Navier-Stokes equation is given by

$$\mathcal{W}(\partial_t + \vec{u}\cdot\vec{\nabla})\vec{u} + v_g^2\vec{\nabla}P + \vec{u}\partial_t P + e(\vec{E}\cdot\vec{j})\vec{u} = v_g^2\left[\eta\Delta\vec{u} - \eta_H\Delta\vec{u}\times\vec{e}_B + en\vec{E} + \frac{e}{c}\vec{j}\times\vec{B}\right] - \frac{\vec{j}_E}{\tau_{\mathrm{dis}}}. \tag{9e}$$

Here η and η_H are the shear and Hall viscosity coefficients, respectively.

The expressions for the dissipative corrections can be found in the Appendix of Ref. [63].

Charge neutral Corbino disk at $B = 0$

Taking into account the rotational symmetry of the Corbino disk, we express the hydrodynamic theory in polar coordinates (r, ϑ). All quantities can only depend on the radial component r. Within linear response and at $B = 0$, the hydrodynamic equations (9) can be transformed to

$$\frac{1}{r}\frac{\partial(r\delta j_r)}{\partial r} = 0, \tag{10a}$$

$$n_{I,0}\frac{1}{r}\frac{\partial(ru_r)}{\partial r} + \frac{1}{r}\frac{\partial(r\delta j_{Ir})}{\partial r} = -\frac{12\ln 2}{\pi^2}\frac{n_{I,0}\mu_I(r)}{T\tau_R}, \tag{10b}$$

$$u_\vartheta = 0, \tag{10c}$$

$$\frac{\partial \delta P}{\partial r} = \eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_r)}{\partial r}\right) - \frac{3Pu_r}{v_g^2\tau_{\rm dis}}, \tag{10d}$$

$$3P\frac{1}{r}\frac{\partial(ru_r)}{\partial r} = -\frac{2\delta P(r)}{\tau_{RE}}. \tag{10e}$$

The electric field $\vec{E}$ does not appear in Eqs. (10) due to charge neutrality. It does however determine the dissipative correction $\delta\vec{j}$ which at charge neutrality is the whole current. In the absence of the magnetic field, all currents are radial. In polar coordinates, the dissipative corrections take the form

$$\delta j_r = \frac{E_r(r)}{eR_0}, \tag{11a}$$

$$\delta j_\vartheta = 0, \tag{11b}$$

$$\delta j_{Ir} = -\frac{\delta_I}{\tau_{\rm dis}^{-1} + \delta_I^{-1}\tau_{22}^{-1}}\frac{2T\ln 2}{\pi}\frac{\partial \mu_I}{\partial r}, \tag{11c}$$

$$\delta j_{I\vartheta} = 0. \tag{11d}$$

Equations (10) and (11) have to be solved together taking into account the corresponding boundary conditions see below.

For the purposes of establishing the boundary conditions we also need to specify the stress tensor. At $B = 0$ (and within linear response, meaning neglecting terms that are higher than the leading order in velocity or its derivatives), the stress tensor is

$$\Pi_E^{\alpha\beta} = P\delta^{\alpha\beta} - \sigma^{\alpha\beta}. \tag{12}$$

Since the Hall viscosity vanishes at charge neutrality (as well as at $B = 0$), the viscous stress tensor in Cartesian coordinates is given by

$$\sigma^{\alpha\beta} = \eta\left(\nabla^\alpha u^\beta + \nabla^\beta u^\alpha - \delta^{\alpha\beta}\vec{\nabla}\cdot\vec{u}\right), \tag{13}$$

which in polar coordinates becomes

$$\sigma_{rr} = -\sigma_{\vartheta\vartheta} = \eta\left(\partial_r - \frac{1}{r}\right)u_r, \quad \sigma_{r\vartheta} = \sigma_{\vartheta r} = \eta\left(\partial_r - \frac{1}{r}\right)u_\vartheta. \tag{14}$$

Description of leads

The leads, which are attached at the inner and outer radius of the Corbino disk, are assumed to be a normal metal in the degenerate regime ($\mu_L \gg T$), where transport is dominated by disorder scattering characterized by the relaxation time τ_L. In this case we may restrict ourselves to a single band, such that there is a single macroscopic current satisfying the continuity equation

$$\partial_t n_L + \vec{\nabla}\vec{j} = 0. \tag{15}$$

Within linear response, one can obtain the macroscopic equation of motion (or generalized Ohm's law) integrating the kinetic equation [64]. This way one finds

$$m\partial_t\vec{j} + \vec{\nabla}\check{\Pi}_E - en_L\vec{E} - \frac{e}{c}\vec{j}\times\vec{B} = -\frac{m}{\tau_L}\vec{j}, \tag{16}$$

where the stress tensor may me expressed in terms of thermodynamic pressure and disorder-induced viscosity

$$\Pi_E^{\alpha\beta} = P\delta^{\alpha\beta} - \sigma^{\alpha\beta}, \qquad \eta_L = \frac{\mu^3\tau_L}{4\pi v_g^2\hbar^2}. \tag{17}$$

To be concrete, we assume that the leads' material is doped graphene. In that case we may introduce the "effective mass" $m = \mu_L/v_g^2$ and the drift velocity $\vec{u}_L$, such that $\vec{j} = n_L\vec{u}_L$. Expressing the carrier density in terms of pressure, we find

$$m\vec{j} = \frac{3P_L}{v_g^2}\vec{u}_L, \tag{18}$$

where to lowest order in temperature we find $P_L = \mu^3/(3\pi v_g^2\hbar^2)$. In the stationary state and at $B = 0$, the equations of motion become

$$\vec{\nabla}\vec{u}_L = 0, \tag{19}$$

$$\vec{\nabla}\check{\Pi}_E + en_L\vec{\nabla}\phi = -\frac{3P_L}{v_g^2\tau_L}\vec{u}_L. \tag{20}$$

Experimentally, the density n_L and the chemical potential μ are fixed by the gate voltage. Moreover, under the common assumption of fast equilibration in the leads, we may assume a uniform temperature T as well. The general variation of P_L is found to be

$$\delta P_L = \left(\frac{2\pi\mu T\delta T}{3v_g^2} + \frac{\pi T^2\delta\mu}{3v_g^2} + \frac{\mu^2\delta\mu}{\pi v_g^2}\right) \tag{21}$$

and thus vanishes under the condition we consider. Since the leads are highly doped, we find $n_L = n_+ = n_I$, such that the imbalance chemical potential μ_I vanishes.

Boundary conditions

The differential equations (10) and (11) should be supplemented by a suitable set of boundary conditions. The only boundaries present in the Corbino are boundaries between the sample and the leads. Since charge conservation is exact and also holds in the leads, we find

$$j_r(r_1 - \epsilon) = \delta j_r(r_1 + \epsilon), \quad \delta j_r(r_2 - \epsilon) = j_r(r_2 + \epsilon). \tag{22}$$

Fixing the total current I in a radially symmetric system completely determines the current density

$$I = e\int \mathrm{d}\vec{A}\cdot\vec{j} = 2\pi e r j_r. \tag{23}$$

In contrast, the total quasiparticle number (imbalance) and entropy are not conserved due to recombination and energy relaxation processes. However, assuming that the corresponding relaxation rates are not singular at the interface, the continuity equations (9b) and (9d) yield the following boundary conditions for the radial components of the current densities.

The resulting boundary conditions at the two interfaces can be summarized as follows

$$j_r(r_1-\epsilon) = n_L u_r(r_1-\epsilon) = \delta j_r(r_1+\epsilon), \tag{24}$$

$$j_{I,r}(r_1-\epsilon) = n_L u_{L,r}(r_1-\epsilon) = n_{I,0} u_r(r_1+\epsilon) + \delta j_{I,r}(r_1+\epsilon) = \delta j_r(r_1+\epsilon), \tag{25}$$

$$s_L u_{L,r}(r_1-\epsilon) = s_B u_r(r_1+\epsilon) \tag{26}$$

$$j_r(r_2+\epsilon) = n_L u_{L,r}(r_2+\epsilon) = \delta j_r(r_2-\epsilon), \tag{27}$$

$$j_{I,r}(r_2+\epsilon) = n_L u_{L,r}(r_2+\epsilon) = n_{I,0} u_r(r_2-\epsilon) + \delta j_{I,r}(r_2-\epsilon) = \delta j_r(r_2-\epsilon), \tag{28}$$

$$s_L u_{L,r}(r_2+\epsilon) = s_B u_r(r_2-\epsilon). \tag{29}$$

Full solution

Solving the equations of motion in the leads, we find

$$u_{L,r} = \frac{I}{2\pi e n_L r}, \quad u_{L,\vartheta} = 0, \tag{30}$$

$$\sigma_{rr} = \frac{-I\eta_L}{\pi e n_L r^2}, \quad \sigma_{r\vartheta} = 0, \tag{31}$$

$$E_r = \frac{2P_L}{e n_L v_g^2 \tau_L} \frac{I}{2\pi e n_L r}, \tag{32}$$

$$\phi(r) = -\frac{I}{2\pi} \frac{2P_L}{e^2 n_L^2 v_g^2 \tau_L} \log\left(\frac{r}{r_0}\right). \tag{33}$$

Here the drift velocity follows from the continuity equation and the relation to the current which in turn is given by Eq. (23). After that, the assumption $\delta P = 0$ leads to the simple $1/r$ behavior for the electrical field E_r as well. Consequently, the charge density (from the Poisson equation) is indeed constant. On the other hand, the constant r_0 in the potential is not fixed by the boundary conditions we have imposed so far. Finally, neither the electric field nor the current depend on the disorder dominated viscosity η_L. However, the viscous stress tensor itself is not zero, which will be used below later.

The above expressions can be re-written in terms of the temperature T and the chemical potential μ_L in the leads. Under our assumptions, the leads' material is graphene, where the entropy density is defined as

$$Ts = 3P - \mu n - \mu_I n_I. \tag{34}$$

For $\mu \gg T$ in the leads we then find

$$P_L = \frac{\pi T^2 \mu}{3 v_g^2} + \frac{\mu^3}{3\pi v_g^2} = P_L^T + P_L^{T=0}, \tag{35}$$

$$n_L = \frac{\pi T^2}{3 v_g^2} + \frac{\mu^2}{\pi v_g^2} \tag{36}$$

$$s_L T = 3P_L - n_L \mu = \frac{\pi T^2 \mu}{v_g^2} + \frac{\mu^3}{\pi v_g^2} - \frac{\pi T^2 \mu}{3 v_g^2} - \frac{\mu^3}{\pi v_g^2} = \frac{2}{3}\frac{\pi T^2 \mu}{v_g^2} = 2P_L^T, \tag{37}$$

so we need to keep finite temperature corrections in the leads as well.

In our sample, the situation is more involved since in neutral graphene the electric current is not related to the hydrodynamic velocity. As a manifestation of this fact, the differential equations (10) and (11) decouple into two disjunct sets. The first one consists of equations (10a) and(11a) with the solution

$$\delta j_r = \frac{I}{2\pi e r}, \tag{38}$$

$$E_r = \frac{I R_0}{2\pi r}, \quad \phi = -\frac{I R_0}{2\pi} \log\left(\frac{r}{r_0}\right). \tag{39}$$

The constant r_0 (not necessarily the same as in the corresponding solution for the leads) is not fixed by the boundary conditions we have imposed so far.

The second set of equations consists of (10b), (10d), (10e) and (11c). Expressing δP through u_r, we find

$$0 = \partial_r \left(\frac{1}{r} \frac{\partial (r u_r)}{\partial r} \right) - \frac{u_r}{\ell_{\mathrm{GE}}^2} \tag{40}$$

$$\frac{1}{\ell_{\mathrm{GE}}^2} = \left(\eta + \frac{3P\tau_{RE}}{2} \right)^{-1} \frac{3P}{v_g^2 \tau_{\mathrm{dis}}}, \tag{41}$$

where the Gurzhi length is renormalized by energy relaxation through the combination $\eta' = \eta + 3P\tau_{RE}/2$. The other two equations can be combined to form

$$\partial_r \left(\frac{1}{r} \frac{\partial (r u_r)}{\partial r} \right) - M \partial_r \left(\frac{1}{r} \frac{\partial (r \frac{\partial \mu_I}{\partial r})}{\partial r} \right) = -\frac{M}{\ell_R^2} \frac{\partial \mu_I(r)}{\partial r} \tag{42}$$

$$M = \frac{2T \ln 2}{n_{I,0}\pi} \frac{\delta_I}{\tau_{\mathrm{dis}}^{-1} + \delta_I^{-1} \tau_{22}^{-1}}, \quad \ell_R^2 = \frac{\delta_I}{\tau_{\mathrm{dis}}^{-1} + \delta_I^{-1} \tau_{22}^{-1}} \frac{\pi T^2 \tau_R}{6 n_{I,0}}. \tag{43}$$

The two coupled Bessel differential equations for u_r and $\partial_r \mu_I$ can be expressed using the differential operator $\mathbb{D} = \partial_r (1/r) \partial_r r$. This way we can write the system of equations in the matrix form

$$\mathbb{D} \begin{pmatrix} 1 & 0 \\ 1 & -M \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} = \begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ 0 & -\frac{M}{\ell_R^2} \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} \Leftrightarrow \mathbb{D} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} = \begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ \frac{1}{M \ell_{\mathrm{GE}}^2} & \frac{1}{\ell_R^2} \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix}. \tag{44}$$

This can be formally solved by diagonalizing the matrix

$$\begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ \frac{1}{M \ell_{\mathrm{GE}}^2} & \frac{1}{\ell_R^2} \end{pmatrix} = \hat{U}^{-1} \hat{D} \hat{U}, \tag{45}$$

where $\hat{D}$ is a diagonal matrix with the eigenvalues d_1 and d_2 (in units of inverse length squared) and then transforming back to the u_r and $\partial_r \mu_I$ basis. Then this coupled Bessel differential equation has the general solution

$$u_r = M \left(1 - \frac{\ell_{\mathrm{GE}}^2}{\ell_R^2} \right) \left[f_1 I_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + f_2 K_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) \right] \tag{46}$$

$$\frac{\partial \mu_I}{\partial r} = f_1 I_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + f_2 K_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + g_1 I_1 \left(\frac{r}{\ell_R} \right) + g_2 K_1 \left(\frac{r}{\ell_R} \right), \tag{47}$$

where the coefficients f_1, f_2, g_1 and g_2 should be determined from the boundary conditions. These involve the entropy density

$$T s_B = 3P = 3 \frac{3T^3 \zeta(3)}{\pi v_g^2}. \tag{48}$$

From the conservation of entropy current Eqs. (26) and (29) we find f_1 and f_2 so that

$$u_r = \frac{I s_L}{2\pi e n_L s_B} \left\{ \frac{I_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right] - K_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}{r_1 r_2 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 r_2 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)} \right\}. \tag{49}$$

This leads to the stress tensor elements

$$\sigma_{rr} = \frac{\eta I s_L}{2\pi e \ell_{\mathrm{GE}} n_L s_B} \frac{I_2\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right] + K_2\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}{r_1 r_2 \left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}, \tag{50}$$

$$\sigma_{r\vartheta} = 0 \tag{51}$$

and

$$
\begin{aligned}
\delta P = & -\frac{3P\tau_{\mathrm{RE}}}{2}\frac{1}{r}\frac{\partial(ru_r)}{\partial r} = \\
= & -\frac{3P\tau_{\mathrm{RE}}}{2}\frac{Is_L}{2\pi e\ell_{\mathrm{GE}}n_Ls_B}\left[\frac{K_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)-r_2I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]+I_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)-r_2K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]}{r_1r_2K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-r_1r_2I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right],
\end{aligned}
\tag{52}
$$

Using the conservation of the imbalance current Eqs. (25) and (28) we find the imbalance chemical potential

$$
\begin{aligned}
\mu_I(r) = & \frac{Is_L\ell_R}{2\pi eMn_Lr_1r_2s_B(\ell_{\mathrm{GE}}^2-\ell_R^2)}\left[\frac{K_0\left(\frac{r}{\ell_R}\right)\left[r_1I_1\left(\frac{r_1}{\ell_R}\right)-r_2I_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2+(\ell_R^2-\ell_{\mathrm{GE}}^2)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right)I_1\left(\frac{r_2}{\ell_R}\right)-I_1\left(\frac{r_1}{\ell_R}\right)K_1\left(\frac{r_2}{\ell_R}\right)}\right. \\
& +\frac{I_0\left(\frac{r}{\ell_R}\right)\left[r_1K_1\left(\frac{r_1}{\ell_R}\right)-r_2K_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2+(\ell_R^2-\ell_{\mathrm{GE}}^2)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right)I_1\left(\frac{r_2}{\ell_R}\right)-I_1\left(\frac{r_1}{\ell_R}\right)K_1\left(\frac{r_2}{\ell_R}\right)} \\
& \left.+\frac{\ell_{\mathrm{GE}}\ell_RK_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-r_1I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}+\frac{\ell_{\mathrm{GE}}\ell_RI_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-r_1K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right]
\end{aligned}
\tag{53}
$$

and the dissipative correction to the imbalance current

$$
\begin{aligned}
\delta j_{Ir}(r) = & \frac{In_{I,0}s_L}{2\pi en_Lr_1r_2s_B\left(\ell_{\mathrm{GE}}^2-\ell_R^2\right)}\left[\frac{K_1\left(\frac{r}{\ell_R}\right)\left[r_1I_1\left(\frac{r_1}{\ell_R}\right)-r_2I_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2+(\ell_R^2-\ell_{\mathrm{GE}}^2)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right)I_1\left(\frac{r_2}{\ell_R}\right)-I_1\left(\frac{r_1}{\ell_R}\right)K_1\left(\frac{r_2}{\ell_R}\right)}\right. \\
& -\frac{I_1\left(\frac{r}{\ell_R}\right)\left[r_1K_1\left(\frac{r_1}{\ell_R}\right)-r_2K_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2+(\ell_R^2-\ell_{\mathrm{GE}}^2)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right)I_1\left(\frac{r_2}{\ell_R}\right)-I_1\left(\frac{r_1}{\ell_R}\right)K_1\left(\frac{r_2}{\ell_R}\right)} \\
& \left.+\frac{\ell_R^2K_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)-r_2I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}+\frac{\ell_R^2I_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-r_1K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)-I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right].
\end{aligned}
\tag{54}
$$

From δP and μ_I we find δT according to

$$
\delta T = \frac{\pi v_g^2}{9T^2\zeta(3)}\delta P - \frac{\pi^2}{27\zeta(3)}\mu_I. \tag{55}
$$

Our hydrodynamic system is not characterized by a local thermal conductivity κ. In other words, the heat current

$$
\vec{j}_Q(r) = 3P\vec{u} - \mu\vec{j} - \mu_I\vec{j}_I \tag{56}
$$

is related to the temperature gradient $\nabla\delta T(r)$ at the same point r non locally. The non-local (integral) relation between $\vec{j}_Q(r)$ and $\nabla\delta T(r')$ characterized by a non-local kernel $\kappa(r,r')$ follows from the fact that the equation for $\vec{u}(r)$ is now a second-order differential equation with a non-local Green's function. Expressing $\delta P(r)$ there in terms of $\delta T(r)$ and $\mu_I(r)$, we have a non-local relation between $\vec{u}(r)$, $\delta T(r')$ and $\nabla\mu_I(r')$. Substituting this $\vec{u}(r)$ into the definition of $\vec{j}_Q(r)$, we obtain a non-local thermal conductivity. As a result one can only introduce the thermal conductance for the device, relating the temperature difference between the contacts with the total heat current through the system. This will be done in a subsequent publication.

Dissipation and total resistance

The above solution is not sufficient to determine the drop in electrochemical potential between the points r_{in} and r_{out} (in the inner and outer lead, respectively) since it contains the undefined constant r_0 that has to be determined from a boundary condition for the electric potential. Although microscopically the potential has to be continuous, several effects might contribute to an apparent discontinuity on the hydrodynamic scale. The most obvious contribution is the contact resistance that is a manifestation of the different work functions in the two materials across the interface as well as the mismatch in their chemical potentials [65]. A more subtle effect due to electron-electron interaction giving rise to viscosity and hence an additional dissipation channel [66]. At charge neutrality, this effect is subtle, since the electric current is decoupled from the hydrodynamic energy flow. However, both flows are induced by the same current source providing the energy dissipated by means of both the Ohmic and viscous effects [66] as well as energy relaxation [48]. The latter processes should be taken into account in the form of an additional voltage drop at the interface.

Under the assumption that energy is not being accumulated at the interface, we generalize the idea proposed in Ref. [66] and consider viscous dissipation in the sample. Since the electric field in bulk of the sample is completely determined by the Ohmic resistance R_0, additional dissipation due to viscosity and energy relaxation corresponds to a jump in the electric potential (on the hydrodynamic scale) at the interface. Microscopically, the voltage jump is due to an excess electric field in the Knudsen layer around the sample-lead boundary.

Consider the kinetic energy associated with the hydrodynamic flow that can be found from the energy density

$$\mathcal{E} = \int \mathrm{d}V\,(n_E - n_E(\vec{u}=0)) \approx \int \mathrm{d}V \frac{6P}{v_g^2}\vec{u}^2. \tag{57}$$

Working within linear response, here we only keep terms up to the second order in $\vec{u}$ (and thus the drive I). Dissipation is then describe by

$$\mathcal{A} = \dot{\mathcal{E}} = 2\frac{6P}{v_g^2}\int \mathrm{dV}\vec{u}\partial_t\vec{u} = 0, \tag{58}$$

vanishing in the steady state. This expression can now be simplified using the generalized Navier-Stokes equation.

In the leads (still assuming graphene at finite carrier density) we find

$$\begin{aligned}\frac{3P_L}{v_g^2}\vec{u}_L\partial_t\vec{u}_L &= \vec{u}_L\left(-\frac{3P_L}{v_g^2}\frac{\vec{u}_L}{\tau_L} - \vec{\nabla}\check{\Pi}_E + n_L e\vec{E}\right)\\ &= -\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} - \vec{\nabla}\delta P\vec{u}_L + \vec{u}_L\vec{\nabla}\check{\sigma} - e\vec{j}\vec{\nabla}\phi.\\ &= -\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} - \frac{\partial u_{L,i}}{\partial x_j}\sigma_{ij} + \vec{\nabla}\left(\vec{u}_L\check{\sigma} - e\vec{j}\phi - \vec{u}_L\delta P\right).\end{aligned} \tag{59}$$

The term $en_L\vec{u}_L\vec{E} = e\vec{j}\vec{E}$ is the Joule heating. Using the divergence theorem we can divide this into a boundary and a bulk term

$$0 = \mathcal{A} = \mathcal{A}_{\mathrm{boundary}} - \mathcal{A}_{\mathrm{bulk}}, \tag{60}$$

$$\mathcal{A}_{\mathrm{boundary}} = 4\int \mathrm{d}\vec{A}\left(\vec{u}_L\check{\sigma} - \vec{u}_L\delta P - e\vec{j}\phi\right), \tag{61}$$

$$\mathcal{A}_{\mathrm{bulk}} = 4\int \mathrm{d}V\left(\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} + \frac{\partial u_{L,i}}{\partial x_j}\sigma_{ij}\right). \tag{62}$$

The boundary term includes the energy transmitted through the interface.

Since the current density is conserved at the interface, we can immediately write down the corresponding equation in the neutral graphene sample, where the Joule heating is given by $e\delta\vec{j}\vec{E}$. Then we find

$$0 = \mathcal{A} = \mathcal{A}_{\mathrm{boundary}} - \mathcal{A}_{\mathrm{bulk}}, \tag{63}$$

$$\mathcal{A}_{\mathrm{boundary}} = 4\int \mathrm{d}\vec{A}\left(\vec{u}\sigma - \vec{u}\delta P - e\delta\vec{j}\phi\right), \tag{64}$$

$$\mathcal{A}_{\mathrm{bulk}} = 4\int \mathrm{d}V\left(\frac{3P}{v_g^2}\frac{\vec{u}^2}{\tau_{\mathrm{dis}}} + \frac{\partial u_i}{\partial x_j}\sigma_{ij} - \delta P(\vec{\nabla}\cdot\vec{u})\right). \tag{65}$$

As stated above, under realistic experimental conditions the non-equilibrium part of the pressure at $u = 0$ on the lead side vanishes

$$\delta P = 0. \tag{66}$$

At the same time, in neutral graphene sample we find

$$\delta P = \left(\frac{9T^2\delta T\zeta(3)}{\pi v_g^2} + \frac{\pi\mu_I T^2}{3v_g^2}\right) = \frac{T^2}{v_g^2}\left(\frac{9\delta T\zeta(3)}{\pi} + \frac{\pi\mu_I}{3}\right). \tag{67}$$

Using the hydrodynamic equations, one may replace δP by $[-3P_B\tau_{\mathrm{RE}}/(2r)]\partial(ru_r)/\partial r$, thus determining δP for finite τ_{RE} without any additional boundary conditions. The same goes for μ_I. Thus we may use the viscous part of the dissipation to find the difference in the electrochemical potential across the interface.

In addition, we may include the contact resistance described by

$$I^2 R_c = \vec{I}^T \check{R}\vec{I}, \tag{68}$$

where $\vec{I}$ includes charge and entropy current and $\check{R}$ includes the thermoelectric coefficients of the interface.

In the absence of magnetic field, both u_ϑ and $\sigma_{r\vartheta}$ vanish. In the leads $\delta P = 0$ and hence we find the condition

$$\begin{aligned} &4\pi\left[r\left(u_r\sigma_{rr}\right)\right]_{r_1-\epsilon} - 2I\phi(r_1-\epsilon) = 4\pi\left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_1+\varepsilon} - 2I\phi(r_1+\epsilon) - 2\vec{I}^T\check{R}\vec{I} \\ &\Leftrightarrow \phi(r_1-\epsilon) - \phi(r_1+\epsilon) = \frac{2\pi}{I}\left\{\left[r\left(u_r\sigma_{rr}\right)\right]_{r_1-\epsilon} - \left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_1+\varepsilon}\right\} + IR_c \end{aligned} \tag{69}$$

at the first interface and similarly for the second interface

$$\begin{aligned} &4\pi\left[r\left(u_r\sigma_{rr} - \delta P u_r\right)\right]_{r_2-\epsilon} - 2I\phi(r_2-\epsilon) = 4\pi\left[r\left(u_r\sigma_{rr}\right)\right]_{r_2+\varepsilon} - 2I\phi(r_2+\epsilon) - 2\vec{I}^T\check{R}\vec{I} \\ &\Leftrightarrow \phi(r_2-\epsilon) - \phi(r_2+\epsilon) = -\frac{2\pi}{I}\left\{\left[r\left(u_r\sigma_{rr}\right)\right]_{r_2+\epsilon} - \left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_2-\varepsilon}\right\} + IR_c. \end{aligned} \tag{70}$$

Combining the above general solution with these conditions, we find at r_1

$$ru_r(\sigma_{rr} - \delta P) = \frac{I^2 s_L^2\eta\ell_{\mathrm{GE}}}{4\pi^2 e^2 n_L^2 s_B^2 \ell_G^2}\frac{\ell_{\mathrm{GE}} - r_2 K_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_2 I_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}{r_1 r_2\left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]} - \frac{\eta I^2 s_L^2}{2\pi^2 e^2 n_L^2 r_1^2 s_B^2}$$

and at r_2

$$ru_r(\sigma_{rr} - \delta P) = -\frac{I^2 s_L^2\eta\ell_{\mathrm{GE}}}{4\pi^2 e^2 \ell_G^2 n_L^2 s_B^2}\frac{\ell_{\mathrm{GE}} - r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}{r_1 r_2\left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]} - \frac{\eta I^2 s_L^2}{2\pi^2 e^2 n_L^2 r_2^2 s_B^2}.$$

As a result, we find the total resistance R of the system in the form

$$IR = \phi(r_{\mathrm{in}}) - \phi(r_{\mathrm{out}}) = I(R_L + R_B + 2R_C + R_L^{\mathrm{diss}} + R_B^{\mathrm{diss}}), \tag{71}$$

$$R_L = \frac{3P_L}{2\pi e^2 n_L^2 v_g^2\tau_L}\log\left(\frac{r_1 r_{\mathrm{out}}}{r_{\mathrm{in}} r_2}\right), \quad R_B = \frac{R_0}{2\pi}\log\left(\frac{r_2}{r_1}\right), \tag{72}$$

$$R_C = \frac{\vec{I}^T\check{R}\vec{I}}{I^2}, \tag{73}$$

$$R_L^{\mathrm{diss}} = \frac{\eta_L}{\pi e^2 n_L^2}\left(\frac{1}{r_2^2} - \frac{1}{r_1^2}\right) \tag{74}$$

$$\begin{aligned} R_B^{\mathrm{diss}} = \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\Bigg\{&\frac{1}{r_1^2} - \frac{1}{r_2^2} + \frac{\ell_{\mathrm{GE}}}{2\ell_G^2} \\ \times\ &\frac{r_2\left[K_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) + I_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right] + r_1\left[I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) + K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right] - 2\ell_{\mathrm{GE}}}{r_1 r_2\left(K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right)}\Bigg\}. \end{aligned} \tag{75}$$

Analysis of results

The behavior of the obtained resistance depends on the hierarchy of length scales r_1, r_2, $r_2 - r_1$, ℓ_{GE} and ℓ_R. In this Section, we specify the quantitative values of the parameters used to produce the plots shown in the main text. For clarity, here we restore the constants $\hbar$ and k_B.

We perform our quantitative analysis assuming the carrier density in the leads to be $n_L = 5 \times 10^{14}$ m^{-2}. The equilibrium temperature in the device (including both leads and the sample) is fixed to $T = 100$ K. The current, that is supplied by the source is $I = 1$ μA and we assume that the effective interaction constant is screened to $\alpha = 0.2$. We further use $\tau_{\text{dis}} = 1.25 \times 10^{-12}$ s and $\tau_L = 0.189 \times 10^{-12}$ s [65], since the density is higher in the leads. This determines all other parameters, except for τ_{RE} and τ_R (or alternatively ℓ_{GE} and ℓ_R). Since these quantities are difficult to extract from the available experimental data, we show results for several different regimes.

The time scales related to electron-electron interaction are given by [63]

$$\tau_{ii} = \hbar \frac{4\pi t_{ii} \log 2}{\alpha^2 k_B T}, \qquad t_{11} = \frac{1}{33.13}, \quad t_{22} = \frac{1}{5.45}. \tag{76}$$

For the above parameter values, we find $\tau_{11} = 0.5 \times 10^{-12}$ s and $\tau_{22} = 3 \times 10^{-12}$ s. The viscosity can be estimated as

$$\eta = \frac{0.446 k_B^2 T^2}{\alpha^2 v_g^2 \hbar} \tag{77}$$

and amounts to $\nu = v_g^2 \eta/(3P) = 0.25$ m^2/s. In addition

$$R_0 = \frac{\pi}{2 \log 2} \frac{\hbar^2}{e^2 k_B T} \left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\text{dis}}} \right) = 1985.33\Omega \tag{78}$$

When describing the hydrodynamic velocity u_r and the pressure δP one can consider three different limits. If $\ell_{\text{GE}} \ll r_1, r_2$, which is achieved for very small τ_{dis}, one finds

$$u_r \approx \frac{I s_L \left(\sqrt{r r_2} \sinh\left(\frac{r - r_2}{\ell_{\text{GE}}}\right) - \sqrt{r r_1} \sinh\left(\frac{r - r_1}{\ell_{\text{GE}}}\right) \right)}{2\pi e n_L r s_B \sqrt{r_1 r_2} \sinh\left(\frac{r_1 - r_2}{\ell_{\text{GE}}}\right)} \tag{79}$$

which means, that the velocity vanishes exponentially close to the interface and is very small in the bulk of the sample. In the opposite limit $\ell_{\text{GE}} \gg r_1, r_2$ u_r shows a behavior similar to the drift velocity in the leads with logarithmic corrections

$$u_r \approx \frac{I s_L}{2\pi e n_L r s_B} + \frac{I s_L r_1^2 r_2^2 \log\left(\frac{r_1}{r_2}\right)}{4\pi e \ell_{\text{GE}}^2 n_L r s_B (r_1^2 - r_2^2)} + \frac{I r s_L \left[r_1^2 \log\left(\frac{r}{r_1}\right) - r_2^2 \log\left(\frac{r}{r_2}\right) \right]}{4\pi e \ell_{\text{GE}}^2 n_L s_B \left(r_1^2 - r_2^2\right)}. \tag{80}$$

Finally, if $r_2 - r_1 \ll r_1, r_2, \ell_{\text{GE}}$ we find the same $1/r$ behavior as in the leads

$$u_r \approx \frac{I s_L}{2\pi e n_L r s_B} \tag{81}$$

The resulting velocity u_r is shown in Fig. 2 of the main text. In the leads, the drift velocity shows a simple $1/r$ behavior, while one finds a jump due to the mismatch of entropy directly at the interface. Inside the sample, the situation depends on the relative size of ℓ_{GE}. If $\ell_{\text{GE}} \ll r_1, r_2$ we indeed observe, that the velocity decreases rapidly close to the interface and exactly vanishes in the bulk of the sample. This behavior is generally only observable in rather large samples, since the quantity τ_{dis} cannot be arbitrarily small while still staying in the hydrodynamic regime. In all other cases, u_r resembles a $1/r$ behavior, that is slightly modified by logarithmic corrections.

The plot of δT are shown in Fig. 3 of the main text. In the limit of $\ell_{\text{GE}} \ll r_1, r_2$ the non-equilibrium part of the temperature δT vanishes in the bulk of the sample. In this limit energy relaxation processes transfer any heating, that may develop in the sample to the substrate and out of the device. There is only a small finite effect very close to the interface. Since this is an effect of τ_R it is in principle independent of ℓ_{GE} and τ_{RE}, however we need $\ell_R < \ell_{\text{GE}}$ to remain in the hydrodynamic regime. In all other scenarios, there is a finite temperature profile, which may amount to 0.5% of the equilibrium temperature.

Finally we take a look at the total resistance R of the system. In general one might place the measuring points r_{in} and r_{out} very close to the interface, in which case the bulk resistance of the leads R_L would not contribute to the total resistance R. We will further disregard the influence of the phenomenological contact resistance R_C, which only depends on the used materials and their relative chemical potential. Then one can consider again three limiting cases of the hydrodynamic, dissipative contribution to the resistance R_B^{diss}. The first limit is $\ell_{\text{GE}} \ll r_1, r_2$ in which case we find

$$\begin{aligned} R_B^{\text{diss}} \approx{}& \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left(\frac{1}{r_1^2}-\frac{1}{r_2^2}\right) \\ &- \frac{s_L^2(A+\eta)\left((r_1+r_2)\cosh\left(\frac{r_1-r_2}{\ell_{\text{GE}}}\right)-2\sqrt{r_1 r_2}\right)\operatorname{csch}\left(\frac{r_1-r_2}{\ell_{\text{GE}}}\right)}{2\pi e^2 \ell_{\text{GE}} n_L^2 r_1 r_2 s_B^2} && (82) \\ \approx{}& \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left[\frac{1}{r_1^2}-\frac{1}{r_2^2}-\frac{\sqrt{\ell_G^2+\frac{v_g^2\tau_{\text{dis}}\tau_{\text{RE}}}{2}}}{2\ell_G^2}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right] && (83) \end{aligned}$$

where the second approximation requires $r_1 - r_2 \gg \ell_{\text{GE}}$. The result of Ref. [66] corresponds to neglecting the term proportional to ℓ_{GE}. The second limit is the case $\ell_{\text{GE}} \gg r_1, r_2$ and we find

$$R_B^{\text{diss}} \approx \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left(\frac{1}{r_1^2}-\frac{1}{r_2^2}+\frac{1}{2\ell_G^2}\log\left(\frac{r_2}{r_1}\right)\right), \tag{84}$$

which introduces a logarithmic correction of exactly the same form as the bulk resistance R_B of the sample. The final limit is $r_2 - r_1 \ll r_1, r_2, \ell_{\text{GE}}$ where we find the result

$$R_B^{\text{diss}} = \frac{s_L^2 \eta (r_2^2 - r_1^2)}{4\pi e^2 n_L^2 r_1 r_2 s_B^2 \ell_G^2}. \tag{85}$$

If one instead directly takes the limit $\tau_{\text{RE}} \to 0$, and additionally $\ell_G \ll r_1, r_2, r_2 - r_1$ one would obtain

$$R_B^{\text{diss}} \approx \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left[\frac{1}{r_1^2}-\frac{1}{r_2^2}-\frac{1}{2\ell_G}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right]. \tag{86}$$

This is the result for the viscous correction to the resistance at charge neutrality in the setup of Ref. [67].

The plots for $\phi(r)$ and $R = R_B + R_B^{\text{diss}} + R_L^{\text{diss}}$ are shown in Fig. 4 and 5 of the main text respectively. In the case of the potential ϕ we find a logarithmic dependence on the radial position r in both the leads and the sample, where the overall prefactor is however different. In all considered cases, the jump at the interface is in the same direction, which for the second interface is opposite to what Ref. [66] obtains. This is due to the fact, that in our case the contribution of δP is larger than the contributions due to η and η_L alone. The jump is larger, for larger ℓ_{GE}. As seen in Fig. 5 of the main text, the total measured resistance is only slightly changed. The correction shown in the inset of Fig. 5 of the main text is nearly logarithmic for the larger ℓ_{GE}, while is saturates for the smaller ℓ_{GE}.

[1] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).

[2] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[3] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, et al., Science **351**, 1058 (2016).

[4] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).

[5] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[6] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. Krishna Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

[7] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[8] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[9] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).

[10] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[11] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, T. T. K. Watanabe, A. Lucas, A. F. Young, and A. C. Bleszynski Jayich (2020), arXiv:2002.05065.

[12] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).
[13] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, et al., Phys. Rev. B **98**, 241304(R) (2018).
[14] A. Jaoui, B. Fauqué, C. W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D. L. Maslov, C. Felser, et al., npj Quantum Materials **3**, 64 (2018).
[15] M. S. Steinberg, Phys. Rev. **109**, 1486 (1958).
[16] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968), [Usp. Fiz. Nauk **94**, 689 (1968)].
[17] B. Bradlyn, M. Goldstein, and N. Read, Phys. Rev. B **86**, 245309 (2012).
[18] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).
[19] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).
[20] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).
[21] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).
[22] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).
[23] L. S. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).
[24] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).
[25] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).
[26] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).
[27] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).
[28] O. M. Corbino, Nuovo Cimento **1**, 397 (1911).
[29] A. Tomadin, G. Vignale, and M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).
[30] T. Holder, R. Queiroz, and A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).
[31] M. Shavit, A. V. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).
[32] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, et al. (2021), arXiv:2111.06412.
[33] A. Hui, V. Oganesyan, and E.-A. Kim, Phys. Rev. B **103**, 235152 (2021).
[34] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022).
[35] A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, and S. Ilani (2021), arXiv:2110.15369.
[36] O. E. Raichev (2022), arXiv:2202.06623.
[37] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).
[38] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).
[39] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).
[40] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).
[41] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).
[42] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).
[43] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).
[44] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).
[45] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018).
[46] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021).
[47] M. P. Marder, *Condensed Matter Physics* (Wiley, 2010).
[48] B. N. Narozhny and I. V. Gornyi, Frontiers in Physics **9**, 108 (2021).
[49] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).
[50] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **103**, 115402 (2021).
[51] Gall, V. and Narozhny, B. N. and Gornyi, I. V., unpublished.
[52] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021).
[53] M. Kumar, A. Laitinen, and P. Hakonen, Nature Communications **9**, 2776 (2018).
[54] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).
[55] G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).
[56] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1987).
[57] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).
[58] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).
[59] D. Halbertal, J. Cuppens, M. Ben Shalom, L. Embon, N. Shadmi, Y. Anahory, H. R. Naren, J. Sarkar, A. Uri, Y. Ronen, et al., Nature **539**, 407 (2016).
[60] D. Halbertal, M. Ben Shalom, A. Uri, K. Bagani, A. Y. Meltzer, I. Markus, Y. Myasoedov, J. Birkbeck, L. S. Levitov, A. K. Geim, et al., Science **358**, 1303 (2017).
[61] A. Aharon-Steinberg, A. Marguerite, D. J. Perello, K. Bagani, T. Holder, Y. Myasoedov, L. S. Levitov, A. K. Geim, and E. Zeldov, Nature **593**, 528 (2021).
[62] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, et al., Nat. Nanotechnol. **14**, 480 (2019).

[63] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021), URL `https://link.aps.org/doi/10.1103/PhysRevB.104.075443`.
[64] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017), URL `https://link.aps.org/doi/10.1103/PhysRevB.96.195401`.
[65] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021), URL `https://link.aps.org/doi/10.1103/PhysRevB.104.115432`.
[66] M. Shavit, A. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019), URL `https://link.aps.org/doi/10.1103/PhysRevLett.123.026801`.
[67] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022), URL `https://link.aps.org/doi/10.1103/PhysRevB.105.125302`.

Corbino magnetoresistance in neutral graphene

Vanessa Gall,[1, 2] Boris N. Narozhny,[2, 3] and Igor V. Gornyi[1, 2, 4]

[1] *Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2] *Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[3] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[4] *Ioffe Institute, 194021 St. Petersburg, Russia*

(Dated: May 30, 2023)

We explore the magnetohydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. Based on the fully consistent hydrodynamic description derived from a microscopic framework and taking into account all peculiarities of graphene-specific hydrodynamics, we report the results of a comprehensive study of the interplay of viscosity, disorder-induced scattering, recombination, energy relaxation, and interface-induced dissipation. In the clean limit, magnetoresistance of a Corbino sample is determined by viscosity. Hence the Corbino geometry could be used to measure the viscosity coefficient in neutral graphene.

Transport measurements remain one of the most common experimental tools in condensed matter physics. Having dramatically evolved past the original task of establishing bulk material characteristics such as electrical and thermal conductivities, modern experiments often involve samples that are tailor-made to target particular properties or behavior.

In recent years considerable efforts have been devoted to uncovering the collective or hydrodynamic flows of charge carriers in ultraclean materials as predicted theoretically [1–4]. Several dedicated experiments focused on answering two major questions: is the observed electronic flow really hydrodynamic and how to measure electronic viscosity [5–10], the quantity that fascinates physicists beyond the traditional condensed matter physics [11–18]. The hydrodynamic regime is apparently easiest to achieve in graphene [2–4]. This material is especially interesting since it can host two drastically different types of hydrodynamic behavior: (i) "conventional" at relatively high carrier densities [3, 19, 20] and (ii) "unconventional" at charge neutrality [21, 22].

Linearity of the excitation spectrum in graphene leads to the fact that electronic momentum density defines the energy current, $\boldsymbol{j}_E$. In the intermediate temperature window where electron-electron interaction is the dominant scattering process in the system ($\ell_{ee} \ll \ell_{\rm dis}, \ell_{e-ph}, W$, in the self-evident notation) the energy flow becomes hydrodynamic. At high carrier densities (in "doped graphene") the energy current is essentially equivalent to the electric current, $\boldsymbol{j}$, allowing one to formulate a Navier-Stokes-like equation for $\boldsymbol{j}$ [20] as pioneered by Gurzhi [19].

At charge neutrality and in the absence of the external magnetic field ($\boldsymbol{B} = 0$) the energy and electric currents decouple [23]. In the hydrodynamic regime the electric current remains Ohmic [22] (with the "internal" or "quantum" conductivity σ_Q due to electron-electron interaction [24–27]), while the Navier-Stokes-like equation describes the energy current [22, 28, 29]. If external magnetic field is applied, the energy and charge flows become entangled [21–23] allowing for a possibility to detect the hydrodynamic flow in electronic transport experiments. In particular, a bulk (infinite) system is characterized by positive, parabolic magnetoresistance [23, 30] proportional to the disorder mean free time $\tau_{\rm dis}$ (disorder scattering is the only mechanism of momentum relaxation).

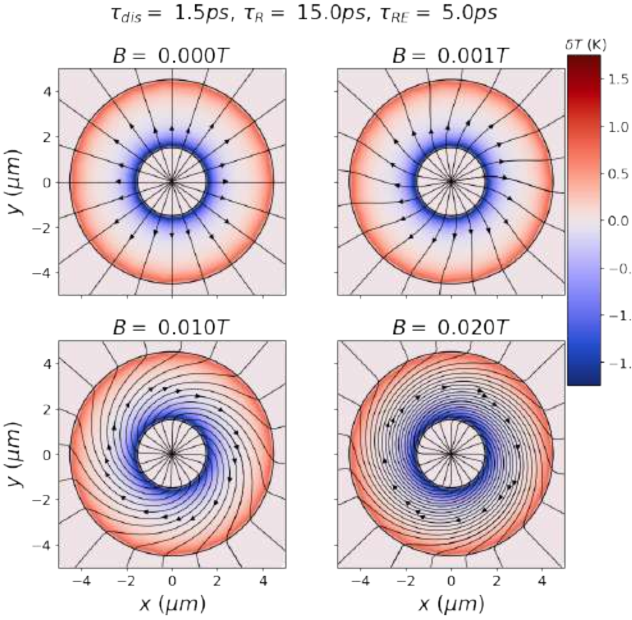


FIG. 1. Hydrodynamic velocity $\boldsymbol{u}$ and temperature δT distribution in the device obtained by solving the hydrodynamic equations at relatively high temperatures where energy relaxation is dominated by supercollisions. Arrows indicate $\boldsymbol{u}$ and the color map shows δT. The quantitative results were computed using the following values of the average temperature $T = 150\,$K, disorder scattering time $\tau_{\rm dis} = 1.5\,$ps (corresponding to the scattering rate $\tau_{\rm dis}^{-1} \approx 0.67\,THz\,\approx 5.1\,$K), recombination time $\tau_R = 15\,$ps, energy relaxation time $\tau_{RE} = 5\,$ps, dimensionless coupling constant in graphene $\alpha = 0.5$, carrier density in the leads $n_L = 5 \times 10^{12}\,$cm^{-2}, and the current passing through the device $I = 1\,\mu$A. The four panels correspond to the indicated values of magnetic field.

The outcome of a given measurement is strongly influenced by the sample size and geometry. Early experiments focused on either the "strip" (or Hall bar) [5–8] or the point contact geometry [9, 10], while more recently data on Corbino disks became available [31].

The simplest viscous phenomenon one can look for in a long (striplike) sample [7–10, 12, 32–52] is the Poiseuille flow [53–55]. This flow is characterized by a parabolic velocity profile with the curvature determined by viscosity. In doped graphene the Poiseuille flow of charge can be detected by imaging the electric current density [8]. In contrast, neutral graphene exhibits the Poiseuille flow of the energy current [56]. Moreover, at relatively high temperatures where hydrodynamic behavior in graphene is observed the electron-phonon interaction (either direct [23, 57, 58] or via "supercollisions" [59–64]) cannot be neglected and hence electronic energy is not conserved. The resulting energy relaxation dwarfs the viscous contribution to the Navier-Stokes [65] equation.

Applying a perpendicular magnetic field to a neutral graphene strip leads to a coupled charge and energy flow with the two currents being orthogonal [23]. The electric current flowing along the strip is accompanied by a neutral quasiparticle flow in the lateral direction resulting in energy and quasiparticle accumulation near the strip boundaries [66, 67]. The accumulation is limited by quasiparticle recombination [67] and energy relaxation processes [59]. As a result, the boundary region's contribution to the resistance is linear in the applied magnetic field [23, 48, 67, 68], in contrast to the standard quadratic magnetoresistance of the bulk system [23, 30]. In classically strong fields the boundary contribution dominates making the linear magnetoresistance directly observable. This effect is not specific to Dirac fermions as shown by experiments in bilayer graphene [69].

The Corbino geometry presents an interesting alternative to the Hall bar experiments [31, 65, 70–78]. In a typical measurement the electric current is passed from the inner to the outer boundary of a Corbino disk. The specific feature of the stationary flow in this geometry is that the magnitude of the radial component of the current is determined by the continuity equation alone. In the absence of the magnetic field the whole current flows radially. Combining the solution of the continuity equation with the hydrodynamic Gurzhi equation (e.g., in doped graphene) leads to an apparent paradox [73]: the current flow appears unaffected by viscosity. However, the dissipated energy is still determined by viscosity leading to the jumps of electric potential at the contacts thus resolving the paradox. In a perpendicular magnetic field the system exhibits parabolic magnetoresistance inverse proportional to the viscosity and independent of the disorder scattering. Applied phenomenologically to neutral graphene (neglecting contact effects) [78] this conclusion stands in sharp contrast to the standard result [23, 30] raising the question of the fate of the disorder-limited bulk magnetoresistance in the Corbino geometry.

In this paper we investigate hydrodynamic flows in neutral graphene in the Corbino disk subjected to the perpendicular magnetic field based on the graphene-specific hydrodynamic theory [2, 22, 59] reporting the results of a careful study of the interplay of viscosity, disorder-induced scattering, recombination, energy relaxation, and interface-induced dissipation. Solving the hydrodynamic equations we find the spatial distribution of the hydrodynamic velocity $\boldsymbol{u}$, temperature (see Fig. 1), electric current, and potential φ (see Fig. 2). Furthermore, we calculate the field-dependent resistance of the whole Corbino sample including the leads. Keeping in mind recent and ongoing experiments, it appears logical to include the effect of the lead resistance in order to achieve a more realistic description of the Corbino device. However, the theoretical limit of "ideal" leads can be considered without any complications.

The main results of this paper are as follows. We show that magnetoresistance of the Corbino device exhibits a crossover from the "hydrodynamic" (viscosity-dominated) to the "bulk" (disorder-limited) behavior with the increasing system size as compared to the Gurzhi length $\ell_G = \sqrt{\nu\tau_{\rm dis}}$ [46–49, 52] (ν is the kinematic viscosity [3, 5, 6, 55, 79] and $\tau_{\rm dis}$ is the disorder mean free time). In the clean limit ($\tau_{\rm dis} \to \infty$) magnetoresistance remains finite and is determined by viscosity offering a way to measure the viscosity coefficient in neutral graphene. In classically strong fields magnetoresistance remain parabolic (in contrast to the linear magnetoresistance in the strip geometry). The "contact magnetoresistance" induced through the dissipation jump is present, but is typically weaker than the bulk contribution.

I. MAGNETOHYDRODYNAMICS IN GRAPHENE

Our arguments are based on the hydrodynamic theory of electronic transport in neutral graphene derived from the kinetic (Boltzmann) equation [21, 22, 59] or from the microscopic Keldysh technique [80]. At charge neutrality both bands contribute to transport on equal footing. A current-carrying state is characterized by the chemical potentials $\mu_\pm$ of each band or by their linear combinations [22, 81]

$$\mu = \frac{\mu_+ + \mu_-}{2}, \qquad \mu_I = \frac{\mu_+ - \mu_-}{2}, \tag{1a}$$

conjugate to the "charge" and "imbalance" (or "total quasiparticle") densities

$$n = n_+ - n_-, \qquad n_I = n_+ + n_-. \tag{1b}$$

In equilibrium $\mu_I = 0$. Any macroscopic current can be expressed as a product of the corresponding density and hydrodynamic velocity $\boldsymbol{u}$ (up to dissipative corrections). Due to the kinematic peculiarity of the Dirac fermions in graphene known as the "collinear scattering singularity" [21, 25] one has to consider the electric, energy, and imbalance, $\boldsymbol{j}_I$ currents defined as

$$\boldsymbol{j} = n\boldsymbol{u} + \delta\boldsymbol{j}, \quad \boldsymbol{j}_I = n_I\boldsymbol{u} + \delta\boldsymbol{j}_I, \quad \boldsymbol{j}_E = \mathcal{W}\boldsymbol{u}, \tag{2}$$

where $\mathcal{W}$ is the enthalpy density and $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ are the dissipative corrections. In the degenerate limit $\mu \gg T$ the dissipative corrections vanish [22, 28] justifying the applicability of the single-band picture to doped graphene. At charge neutrality $n = 0$, the electric and energy currents in Eq. (2) appear to be decoupled [22].

Within linear response, steady-state macroscopic currents obey the linearized hydrodynamic equations [82]. Assuming that the dominant mechanism of energy relaxation is supercollisions [59], the equations have the form

$$\boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{3a}$$

$$n_I\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -\frac{12\ln 2}{\pi^2}\frac{n_I\mu_I}{T\tau_R}, \tag{3b}$$

$$\boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} + \frac{e}{c}\delta\boldsymbol{j}\times\boldsymbol{B} - \frac{3P\boldsymbol{u}}{v_g^2\tau_{\rm dis}}, \tag{3c}$$

$$3P\boldsymbol{\nabla}\cdot\boldsymbol{u} = -\frac{2\delta P}{\tau_{RE}}. \tag{3d}$$

Here Eq. (3a) is the continuity equation; Eq. (3b) is the "imbalance" continuity equation [22, 81] (where v_g is the band velocity in graphene, c is the speed of light, e is the unit charge, and τ_R is the recombination time); Eq. (3c) is the linearized Navier-Stokes equation [22, 29, 82, 83] (with η being the shear viscosity); and Eq. (3d) is the linearized "thermal transport" equation (τ_{RE} is the energy relaxation time [59]). We follow the standard approach [55] where the thermodynamic quantities are replaced by the corresponding equilibrium functions of the hydrodynamic variables. Equilibrium thermodynamic quantities, i.e., the pressure $P = 3\zeta(3)T^3/(\pi v_g^2)$, enthalpy density $\mathcal{W}$, imbalance density, $n_I = \pi T^2/(3v_g^2)$, and energy density are related by the "equation of state", $\mathcal{W} = 3P = 3n_E/2$. Equations (3) should be solved for the unknowns $\boldsymbol{u}$, μ_I, and δP keeping the remaining (thermodynamic) quantities, e.g., n_I, P, and T, constant.

The dissipative corrections to the macroscopic currents can be determined from the underlying microscopic theory [22, 29, 82] and are expressed in terms of the same variables closing the set of hydrodynamic equations (3)

$$\delta\boldsymbol{j} = \frac{1}{e^2\tilde{R}}\left[e\boldsymbol{E} + \omega_B\boldsymbol{e}_B\times\left(\frac{\alpha_1\delta_I\boldsymbol{\nabla}\mu_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}} - \frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right)\right], \tag{4a}$$

$$\delta\boldsymbol{j}_I = -\frac{\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{1}{e^2\tilde{R}}\times \tag{4b}$$

$$\times\left[\alpha_1\omega_B\boldsymbol{e}_B\times\boldsymbol{E} + \frac{2T\ln 2}{\pi}e^2R_0\boldsymbol{\nabla}\mu_I + \alpha_1\omega_B^2\frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right],$$

$$\tilde{R} = R_0 + \alpha_1^2\delta_I\tilde{R}_B. \tag{4c}$$

In Eqs. (4) the following notations are introduced. R_0 is the zero-field bulk resistivity in neutral graphene [23, 30]

$$R_0 = \frac{\pi}{2e^2T\ln 2}\left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\rm dis}}\right) \underset{\tau_{\rm dis}\to\infty}{\longrightarrow} \frac{1}{\sigma_Q}, \tag{5}$$

where $\tau_{11} \propto \alpha_g^{-2}T^{-1}$ describes the appropriate electron-electron collision integral. $\tilde{R}_B$ denotes [65, 82]

$$\tilde{R}_B = \frac{\pi}{2e^2T\ln 2}\frac{\omega_B^2}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}, \tag{6}$$

where $\tau_{22} \propto \alpha_g^{-2}T^{-1}$ describes a component of the collision integral that is qualitatively similar, but quantitatively distinct from τ_{11} and $\delta_I \approx 0.28$. Another numerical factor in Eqs. (4) is $\alpha_1 \approx 2.08$ and $\omega_B = eBv_g^2/(2cT\ln 2)$ is the generalized cyclotron frequency at $\mu = 0$.

The shear viscosity at charge neutrality and in the absence of magnetic field was evaluated in Refs. [22, 79, 83] and has the form

$$\eta(\mu = 0, B = 0) = \mathcal{B}\frac{T^2}{\alpha_g^2v_g^2}, \qquad \mathcal{B} \approx 0.45. \tag{7}$$

Within the renormalization group (RG) approach, α_g is a running coupling constant [56, 83–86]. However, the product α_gv_g remains constant along the RG flow [24, 83]. Hence Eq. (7) gives the correct form of shear viscosity in neutral graphene [84]. Within the kinetic theory approach, the coefficient $\mathcal{B}$ can be expressed in terms of time scales characterizing the collision integral [22, 79]. At neutrality these time scales are qualitatively similar to, but quantitatively distinct from τ_{11} and τ_{22}. The similarity follows from the fact that in general all time scales are functions of the chemical potential and temperature [22, 28, 87]. At neutrality $\mu = 0$ and hence all time scales are inverse proportional to temperature.

As a function of the magnetic field, the viscosity coefficient in neutral graphene exhibit a weak decay until eventually saturating in classically strong fields [79]

$$\eta(\mu = 0, B) = \frac{\mathcal{B}+\mathcal{B}_1\gamma_B^2}{1+\mathcal{B}_2\gamma_B^2}\frac{T^2}{\alpha_g^2v_g^2}, \qquad \gamma_B = \frac{|e|v_g^2B}{\alpha_g^2cT^2}, \tag{8}$$

where

$$\mathcal{B}_1 \approx 0.0037, \qquad \mathcal{B}_2 \approx 0.0274.$$

This behavior should be contrasted with the more conventional Lorentzian decay of field-dependent shear viscosity in doped graphene [6, 45, 46, 79, 88]. However, in weak fields where most present-day experiments are performed this distinction is negligible. Moreover, due to the smallness of the coefficient $\mathcal{B}_1$ and $\mathcal{B}_2$ we disregard the field dependence of η in what follows.

Under the assumptions of the hydrodynamic regime, disorder scattering is characterized by the large mean free time, $\tau_{\rm dis} \gg \tau_{11}, \tau_{22}$, yielding a negligible contribution to Eqs. (5) and (6). Equation (5) describes the uniform bulk current (at $\boldsymbol{B} = 0$) and is independent of viscosity (i.e., in a channel [3, 21, 67, 82]). In contrast, in the Corbino geometry the current flow is necessarily inhomogeneous and hence viscous dissipation must be taken into account.

II. BOUNDARY CONDITIONS

Differential equations (3) should be supplemented by boundary conditions, which should be expressed in terms of the hydrodynamic velocity and macroscopic currents. The statement of the boundary conditions does not imply the validity of the hydrodynamic approximation at the sample edges and generally have to be derived from the underlying microscopic theory. However some of the boundary conditions can be derived based on the conservation laws alone. In the circular Corbino geometry conservation laws can be used to establish boundary conditions for radial components of the currents [65].

A. Radial components of macroscopic currents

A typical experimental setup involves a graphene sample (in our case, at charge neutrality) in the shape of an annulus placed between the inner (a disk of radius r_1) and outer (a ring with the inner radius r_2) metallic contacts (leads). The electric current I is injected into the center of the inner lead preserving the rotational invariance (e.g., through a thin vertical wire attached to the center point) and spreads towards the outer lead, which for concreteness we assume to be grounded. The overall voltage drop U is measured between two points in the two leads (at the radii $r_{\rm in} < r_1$ and $r_{\rm out} > r_2$) yielding the device resistance, $R = U/I$. The only boundaries in the system are between the sample and the external leads.

For simplicity, we assume both leads to be of the same material with a single-band electronic system, e.g., highly doped graphene with the same doping level. In that case, all macroscopic currents in the leads are proportional to the drift velocity and hence are determined by the injected current. In the stationary case, the continuity equation (3a) determines the radial component of the electric current density. In the inner lead this yields $j_r^{\rm in} = I/(2\pi e r)$, defining the radial component of the drift velocity, $u_r^{\rm in} = j_r^{\rm in}/n_L$ (n_L is the carrier density in the inner lead). Assuming charge conservation is not violated at the interface, we find the boundary condition between the inner lead and the sample

$$j_r(r_1-\epsilon) = n_L u_r(r_1-\epsilon) = \delta j_r(r_1+\epsilon), \tag{9a}$$

where $\epsilon > 0$ is infinitesimal and we took into account that in neutral ($n = 0$) graphene $\boldsymbol{j} = \delta\boldsymbol{j}$.

The second hydrodynamic equation, Eq. (3b), is the continuity equation for the imbalance density. Although the total quasiparticle number is not conserved, integrating this equation over an infinitesimally thin region encompassing the boundary yields a similar boundary condition for the imbalance current assuming that the relaxation rate due to quasiparticle recombination is not singular at the boundary

$$j_{I,r}(r_1-\epsilon) = n_L u_r(r_1-\epsilon) = n_I u_r(r_1+\epsilon) + \delta j_{I,r}(r_1+\epsilon). \tag{9b}$$

Here we took into account the fact that in a single-band system $\boldsymbol{j}_I$ is identical with $\boldsymbol{j}$.

Finally, Eq. (3d) is the linearized continuity equation for the entropy density (here we follow the standard practice [55] of replacing the continuity equation for the energy density by the entropy flow equation, also known as the thermal transport equation). Again, assuming the energy relaxation rate is not singular at the interface (i.e., the current flow is not accompanied by energy or excess heat accumulation at the boundary between the sample and the contact) we integrate Eq. (3d) over an infinitesimally thin region encompassing the boundary and arrive at the boundary condition for the entropy current

$$s^{\rm in} u_r(r_1-\epsilon) = s u_r(r_1+\epsilon), \tag{9c}$$

where s and $s^{\rm in}$ are the entropy densities in the sample and inner lead, respectively.

B. Tangential flows in external magnetic field

The above boundary conditions (and the corresponding conditions on the outer lead) are sufficient to solve the hydrodynamic equations in the absence of magnetic field where all currents are radial [65]. An external magnetic field induces the tangential components of the currents due to the classical Hall effect. The continuity equations do not determine the tangential components and hence the boundary conditions have to be derived from a microscopic theory. Generally speaking, the boundary conditions depend on the presence of tangential forces at the interface, usually associated with edge roughness. Typically [2–4, 55, 73], one considers the two limiting cases of either the "no-slip" or "no-stress" boundary conditions corresponding to either the presence or the absence of the drag-like friction across the interface.

For contact interfaces in the Corbino geometry, the boundary conditions corresponding to the above limiting cases differ from the well-known expression of conventional hydrodynamics. The no-slip boundary condition now means that the tangential component of the hydrodynamic velocity is continuous across the interface (written as above for the inner interface)

$$u_{L\vartheta}(r_1-\epsilon) = u_\vartheta(r_1+\epsilon), \tag{10a}$$

in contrast to the common condition of vanishing velocity at the channel boundary (the two are consistent, since in the latter case there is no flow beyond the edge).

The no-stress boundary condition means the absence of any forces along the interface in which case the tangential component of the stress tensor Π^{ij} is continuous. In polar coordinates appropriate for the Corbino geometry one finds

$$\Pi^{\vartheta r}_{L,E}(r_1-\epsilon) = \Pi^{\vartheta r}_{E}(r_1+\epsilon), \tag{10b}$$

The no-stress boundary condition is easy to derive starting from the kinetic equation. Multiplying the kinetic equation by the momentum and summing over all quasiparticle states, one finds an equation featuring the gradient of the stress tensor [22] as well as macroscopic forces in the system. Now the boundary condition can be obtained by integrating that equation over the small volume around the interface. Unless there is a force localized at the interface (with a δ-function-like coordinate dependence on the hydrodynamic scale), this procedure would yield Eq. (10b). Usually, the interfaces are microscopically rough with the roughness providing such a force. As a result, the no-slip boundary condition is more commonly used. In neutral graphene, however, the quasiparticle wavelength typically exceeds any length scale associated with edge roughness leading to specular scattering [82] and Eq. (10b).

In the case of the hard wall edges, the boundary conditions were previously studied theoretically in Ref. [89] and confirmed experimentally in Ref. [8] where a nonzero slip length was proposed indicating a more general Maxwell's boundary condition. However, the specific choice of the boudnary conditions does not lead to qualitatively different results [73]. Here we follow the hydrodynamic tradition and consider both the no-slip and no-stress boundary conditions.

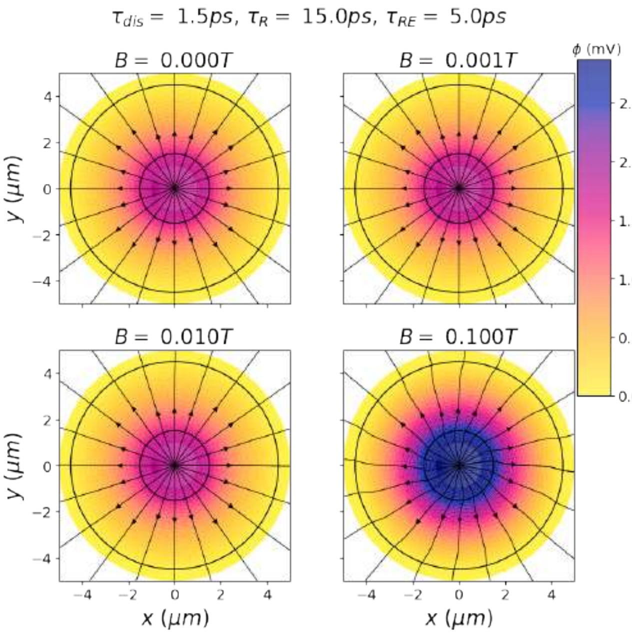


FIG. 2. Electric current density $\boldsymbol{j}$ and potential φ in the device obtained by solving the hydrodynamic equations at relatively high temperatures where energy relaxation is dominated by supercollisions. Arrows indicate $\boldsymbol{j}$ and the color map shows φ. The outer lead is chosen to be grounded. The four panels correspond to the indicated values of magnetic field. For the values of other parameters, see Fig 1.

C. Interface-induced dissipation and jumps of the electric potential

The hydrodynamic theory discussed so far completely describes the energy flow in neutral graphene. In order to establish the device resistance R we have to find the behavior of the electrochemical potential at the interfaces.

The standard description of interfaces between metals or semiconductors in terms of the contact resistance [90] can be carried over to neutral graphene [81]. In graphene, the contact resistance was recently measured in Ref. [8] (see also Refs. [31, 91, 92]). In the diffusive (or Ohmic) case, the contact resistance leads to a voltage drop that is small compared to that in the bulk of the sample and can be neglected. In contrast, in the ballistic case with almost no voltage drop in the bulk, most energy is dissipated at the contacts. Both scenarios neglect interactions.

In the diffusive regime interactions give rise to perturbative corrections to the bulk resistivity [93, 94] and the contact resistance can still be neglected. In ballistic samples electron-electron interaction may lead to the "Knudsen-Poiseuille" crossover [19] and drive the system to the hydrodynamic regime. In this case the Ohmic resistivity of the electronic fluid may remain small, but there exist other channels for dissipation due to viscosity [73] and energy relaxation processes [59]. In neutral graphene the effect is subtle [65], since the electric current is decoupled from the hydrodynamic energy flow. However, both are induced by the current source that provides the energy dissipated through all the above channels. The energy dissipated in the system corresponds to the overall voltage drop. In the bulk of the sample the voltage drop is Ohmic as determined by Eq. (4a), while the additional contribution takes the form of a potential jump at the interface between the sample and leads. At the same time, an excess electric field is induced in a thin Knudsen layer around the interface [73].

The magnitude of the jump in ϕ can be established by considering the flow of energy through the interface as suggested in Ref. [73] and detailed in neutral graphene at $\boldsymbol{B}=0$ in Ref. [65]. Consider the kinetic energy defined by integrating the energy density $n_E(\boldsymbol{u})-n_E(0)$ over the volume

$$\mathcal{E}=\int dV\left[n_E(\boldsymbol{u})-n_E(0)\right]\approx\int dV\frac{6P}{v_g^2}u^2, \tag{11}$$

which we have expanded to the leading order in $\boldsymbol{u}$ (and hence in I). In the stationary state, dissipation is balanced by the work done by the source, such that the time derivative of the kinetic energy vanishes, $\mathcal{A}=\dot{\mathcal{E}}=0$. Using the equations of motion and continuity equations to find time derivatives, one may split $\mathcal{A}$ into the "bulk" and "boundary" contributions, $\mathcal{A}=\mathcal{A}_{\rm bulk}+\mathcal{A}_{\rm edge}$. The former may be interpreted as the bulk dissipation, while $\mathcal{A}_{\rm edge}$ must include the energy brought in (carried away) through the boundary by the incoming (outgoing) flow. The boundary condition is then found under the assumption that energy is not accumulated at the interface.

Assuming the leads' material is highly doped graphene, the equation of motion is the usual Ohm's law where we may combine the diffusion term [95] with a contribution of viscosity η_L due to disorder [96] into the gradient of the stress-energy tensor [23] and hence

$$\begin{aligned}
&\frac{3P_L}{v_g^2}\boldsymbol{u}_L\partial_t\boldsymbol{u}_L = \\
&= u_L^i\left(-\frac{3P_L}{v_g^2}\frac{u_L^i}{\tau_L}-\nabla^j\Pi^{ij}_{L,E}+n_LeE^i+\frac{e}{c}\epsilon^{ijk}j^jB^k\right) \\
&= -\frac{3P_L}{v_g^2}\frac{u_L^2}{\tau_L}+\frac{\partial u_{L,i}}{\partial x_j}\Pi^{ij}_{L,E}+\frac{e}{c}\boldsymbol{u}_L\cdot(\boldsymbol{j}\times\boldsymbol{B})+e\varphi\boldsymbol{\nabla}\cdot\boldsymbol{j} \\
&\qquad -\nabla^i\left(u_L^j\Pi^{ij}_{L,E}+ej^i\varphi\right).
\end{aligned}$$

The last term in this expression determines the boundary contribution. Given that the Lorentz force does not explicitly contribute, the only difference from the expression derived in Ref. [65] at $\boldsymbol{B}=0$ is the nonzero tangential components of the hydrodynamic velocity and the stress tensor (vanishing in the absence of magnetic field). In neutral graphene, we obtain a similar expression from the Navier-Stokes equation, while the Joule heat is determined by $\delta\boldsymbol{j}$. Equating the two contributions we find the jump of the potential in the form

$$\begin{aligned}
&\varphi(r_1-\varepsilon)-\varphi(r_1+\varepsilon) = IR_c + \\
&\quad +\frac{2\pi r_1}{I}\left[\left(u_r\Pi^{rr}_E+u_\vartheta\Pi^{\vartheta r}_E\right)\Big|_{r_1+\varepsilon}\right. \\
&\qquad\qquad \left. -\left(u_r\Pi^{rr}_{L,E}+u_\vartheta\Pi^{\vartheta r}_{L,E}\right)\Big|_{r_1-\varepsilon}\right],
\end{aligned} \tag{12}$$

where R_c is the usual contact resistance [81]. A similar condition holds at the boundary with the outer lead.

III. HYDRODYNAMIC FLOWS IN THE CORBINO GEOMETRY

In polar coordinates and taking into account radial symmetry, the hydrodynamic equations (3) and (4) form two disjoint sets of differential equations. The first one determines the tangential component of the velocity u_ϑ:

$$\frac{1}{r}\frac{\partial(r\delta j_r)}{\partial r}=0, \tag{13a}$$

$$\eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_\vartheta)}{\partial r}\right)-\frac{eB}{c}\delta j_r-\frac{3Pu_\vartheta}{v_g^2\tau_{\rm dis}}=0, \tag{13b}$$

$$\delta j_r=\frac{1}{e^2\tilde{R}}\left[eE_r(r)+\omega_B\frac{2T\ln 2}{v_g^2}u_\vartheta\right], \tag{13c}$$

$$\delta j_{I\vartheta}=-\frac{\alpha_1\delta_I\omega_B}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\delta j_r, \tag{13d}$$

while the second one involves the radial component u_r:

$$\frac{n_I}{r}\frac{\partial(ru_r)}{\partial r}+\frac{1}{r}\frac{\partial(r\delta j_{Ir})}{\partial r}=-\frac{12\ln 2}{\pi^2}\frac{n_I\mu_I(r)}{T\tau_R}, \tag{14a}$$

$$\frac{\partial\delta P}{\partial r}=\eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_r)}{\partial r}\right)+\frac{eB}{c}\delta j_\vartheta-\frac{3Pu_r}{v_g^2\tau_{\rm dis}}, \tag{14b}$$

$$\frac{3P}{r}\frac{\partial(ru_r)}{\partial r}=-\frac{2\delta P(r)}{\tau_{RE}}. \tag{14c}$$

$$\delta j_\vartheta=\frac{\omega_B}{e^2\tilde{R}}\left(\frac{\alpha_1\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{\partial\mu_I}{\partial r}-\frac{2T\ln 2}{v_g^2}u_r\right), \tag{14d}$$

$$\delta j_{Ir}=-\frac{2\delta_I T\ln 2}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\left[\frac{R_0}{\pi\tilde{R}}\frac{\partial\mu_I}{\partial r}+\frac{\alpha_1\omega_B^2}{e^2\tilde{R}}\frac{u_r}{v_g^2}\right]. \tag{14e}$$

A. Tangential component of the velocity and bulk voltage drop

The bulk magnetoresistance can be found by solving Eqs. (13) with the appropriate boundary conditions. Combining Eqs. (13a) and (13b) we find an inhomogeneous Bessel equation for the tangential component of the velocity u_ϑ with the characteristic length scale being the Gurzhi length $\ell_G^2=\eta v_g^2\tau_{\rm dis}/(3P)$. The boundary condition for u_ϑ is determined by microscopic details of viscous drag at the interface and hence is not universal. Here we follow the hydrodynamic tradition and consider both the no-slip and the no-stress boundary conditions, see Sec. II B. Moreover, one can distinguish two different setups where the external magnetic field is applied either to the sample only or to the whole device including the leads. In all these cases we can find an analytic expression for u_ϑ, which can be substituted into of Eq. (13c) to find the electric field in the sample, E_r (the radial component of the current is determined by the continuity equation alone). Similarly, Eq. (13d) determines $\delta j_{I\vartheta}$. Using the obtained electric field we can determine the voltage drop through the bulk of the sample as

$$U=\int\limits_{r_1}^{r_2}E_r dr=\int\limits_{r_1}^{r_2}dr\left(\frac{\tilde{R}I}{2\pi r}-\frac{B}{c}u_\vartheta\right). \tag{15}$$

For the no-slip boundary condition for u_ϑ and allowing the external magnetic field to penetrate the leads, the tangential component of the velocity is given by

$$\begin{aligned}
u_\vartheta = &-\frac{BI\ell_G^2}{2\pi c\eta r}+\frac{BI\left(\eta\ell_L^2-\eta_L\ell_G^2\right)}{2\pi c\eta\eta_L r_1r_2}\times \\
&\times\left[K_1\left(\frac{r}{\ell_G}\right)\frac{r_1I_1\left(\frac{r_1}{\ell_G}\right)-r_2I_1\left(\frac{r_2}{\ell_G}\right)}{K_1\left(\frac{r_1}{\ell_G}\right)I_1\left(\frac{r_2}{\ell_G}\right)-I_1\left(\frac{r_1}{\ell_G}\right)K_1\left(\frac{r_2}{\ell_G}\right)}\right. \\
&\left.+I_1\left(\frac{r}{\ell_G}\right)\frac{r_2K_1\left(\frac{r_2}{\ell_G}\right)-r_1K_1\left(\frac{r_1}{\ell_G}\right)}{K_1\left(\frac{r_1}{\ell_G}\right)I_1\left(\frac{r_2}{\ell_G}\right)-I_1\left(\frac{r_1}{\ell_G}\right)K_1\left(\frac{r_2}{\ell_G}\right)}\right],
\end{aligned} \tag{16}$$

where η_L is the disorder-induced viscosity [96] and $\ell_L^2 = v_g^2\eta_L\tau_L/(2P_L)$ is the Gurzhi length in the leads.

In the limit $\ell_G \gg r_1, r_2$ (i.e., "clean system" with long mean free time $\tau_{\rm dis}\to\infty$) this simplifies to ($p = r_2/r_1$)

$$u_\vartheta \approx -\frac{BI\ell_L^2}{4\pi cr\eta_L}\left[2+\left(\frac{1}{\ell_G^2}-\frac{\eta_L}{\eta\ell_L^2}\right)\times\right. \tag{17}$$
$$\left.\times\,\frac{r^2\ln\left(\frac{r}{r_1}\right)+r^2p^2\ln\left(\frac{r_2}{r}\right)-r_2^2\ln p}{1-p^2}\right].$$

The corresponding voltage drop remains finite

$$U \approx \left(1-\frac{\eta\ell_L^2}{\eta_L\ell_G^2}\right)\frac{B^2Ir_2^2}{4\pi c^2\eta}\frac{(p^2-1)^2-4p^2\ln^2 p}{4p^2(p^2-1)} \tag{18a}$$
$$+\frac{I\ln p}{2\pi}\left(\frac{B^2}{c^2}\frac{v_g^2\tau_L}{3P_L}+\tilde{R}\right),$$

yielding the field-dependent bulk resistance ($R = U/I$)

$$R(B) \approx \frac{\ln p}{2\pi}R_0+\frac{B^2r_2^2}{4\pi c^2\eta}\frac{(p^2-1)^2-4p^2\ln^2 p}{4p^2(p^2-1)} \tag{18b}$$
$$+\frac{B^2v_g^4\ln p}{2c^2T^3}\left[\frac{\alpha_1^2\delta_I}{8\ln^3 2}\frac{1}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}+\frac{T^3}{\mu^3}\tau_L\right],$$

assuming $\eta\ell_L^2/(\eta_L\ell_G^2) = 3P\tau_L/(2P_L\tau_{\rm dis}) \ll 1$ with $P_L = \mu_L^3/(3\pi v_g^2)$. The two field-dependent terms differ in their dependence on temperature, sample size, and coupling constant [35] opening a possibility to separate the two contributions from the experimental data and thus to measure the viscosity coefficient.

If the magnetic field is applied to the sample only (and not to the leads) u_ϑ vanishes in the leads and hence the terms with ℓ_L do not appear in the voltage drop (18). In that case, the field-dependent contribution to U does not contain $\tau_{\rm dis}$ in contrast to the known result in the strip geometry [23, 30].

A similar result can be obtained in the case of no-stress boundary conditions, where the tangential component of the velocity u_ϑ is still expressed in terms of the Bessel functions. In the clean limit ($\ell_G \gg r_1, r_2$) the voltage drop also remains finite

$$U \approx \frac{I}{2\pi}\left(\tilde{R}+\frac{B^2\ell_L^2}{c^2\eta}-\frac{B\eta_H}{ec\eta n_L}\right)\ln p \tag{19}$$
$$+\frac{r_2^2B^2I}{4\pi c^2\eta}\left[\frac{\left(p^2-1\right)\left(p^4+10p^2+1\right)}{12p^2\left(p^2+1\right)^2}-\frac{\ln p}{1+p^2}\right]$$
$$+\frac{I}{2\pi}\left[\frac{B^2}{c^2}\frac{\left(\ell_G^2-\ell_L^2\right)}{\eta}+\frac{B\eta_H}{ec\eta n_L}\right]\frac{p^2-1}{p^2+1},$$

where η_H is the Hall viscosity in the leads, which vanishes if the magnetic field is not allowed in the leads. In that case, the last term in the voltage drop (19) is proportional to $\tau_{\rm dis}$ and independent of viscosity. The second term in Eq. (19) remains similar to Eq. (18) and is inverse proportional to η. This term's dependence on the ratio p is distinct from both Eq. (18) and the third term in Eq. (19) and could be extracted by analyzing the data in a set of Corbino disks with different p.

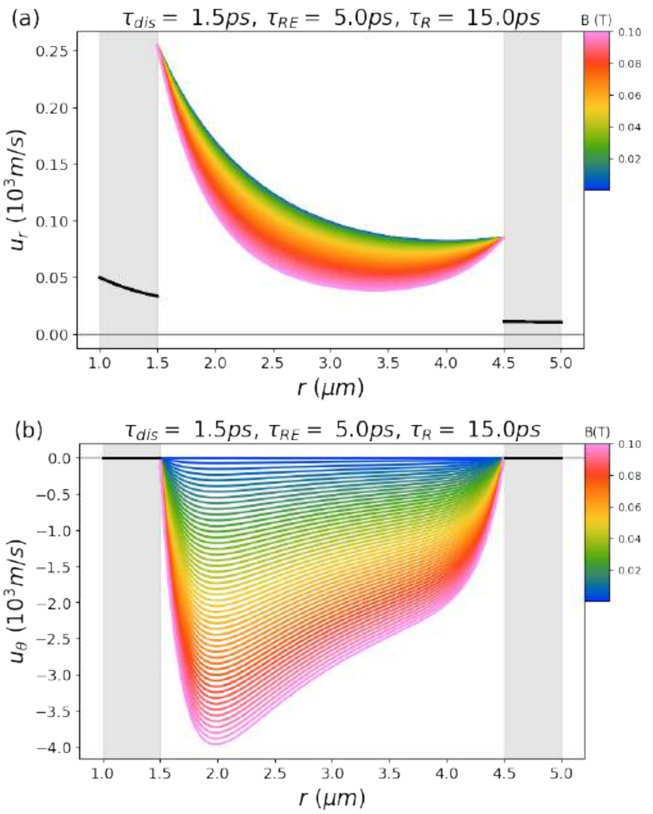


FIG. 3. Radial (top panel) and tangential (bottom panel) components of the hydrodynamic velocity $\boldsymbol{u}$ computed within the "supercollisions model" of energy relaxation. Black lines in the shaded regions show the drift velocity in the leads. Color curves correspond to different values of the external magnetic field according to the shown color coding. The top curve shows values at $\boldsymbol{B} = 0$ and is identical with the results of Ref. [65]. For the parameter values, see Fig 1.

In the opposite limit $\ell_G \ll r_1, r_2$, the leading contribution to the bulk voltage drop is independent of η. For no-slip boundary conditions and in the simplified case where the field is not allowed to penetrate the leads we find for the field-dependent part of U

$$R(B)-R(0) \approx \frac{B^2v_g^2\tau_{\rm dis}\ln p}{6\pi c^2P}+\frac{\ln p}{2\pi}\delta_I\alpha_1^2\tilde{R}_B \propto \tau_{\rm dis}B^2. \tag{20}$$

The voltage drop (20) is proportional to $\tau_{\rm dis}$ similarly to the result in the strip geometry (see Refs. [23, 30]). Of course, in the limit $\ell_G \ll r_1, r_2$ the mean free time $\tau_{\rm dis}$ cannot be arbitrarily large, hence the voltage drop (20) does not diverge. In the limit $\tau_{\rm dis}\to\infty$ the voltage drop crosses over to the above "clean" limit and is given by Eq. (18). However, the limiting expression (20) is independent of viscosity, and hence qualitatively similar to the usual result.

To summarize the results of this section, we have shown that bulk magnetoresistance in neutral graphene in the Corbino geometry exhibits a crossover between the “clean” limit of the large (compared to the disk radius) Gurzhi length to the limit of small Gurzhi length. In the former case, the field-dependent part of the bulk voltage drop is determined by viscosity, while in the latter limit it is proportional to the disorder mean free time similarly to the known result in the strip geometry.

B. Radial component of the velocity and the device resistance

The five equations (14) can be reduced to two coupled differential equations (for similar calculations in the strip geometry see Refs. [23, 48, 68, 82]). To simplify the arguments, we introduce the following notations

$$q=n_Iu_r,\quad p=\delta j_{I,r},\quad x=\frac{2n_I}{3P}\delta P,\quad y=\frac{12\ln 2}{\pi^2}\frac{n_I}{T}\mu_I. \tag{21}$$

In terms of the new variables, Eqs. (14a) and (14c) can be written as

$$\frac{1}{r}\frac{\partial(rq)}{\partial r}+\frac{1}{r}\frac{\partial(rp)}{\partial r}=-\frac{y}{\tau_R}, \tag{22a}$$

$$\frac{1}{r}\frac{\partial(rq)}{\partial r}=-\frac{x}{\tau_{RE}}. \tag{22b}$$

Equation (14e) can be rewritten as

$$\frac{\partial y}{\partial r}=-\frac{6}{\pi}\frac{\tilde{R}n_I}{R_0T^2\tilde{\tau}}\,p-\frac{12\ln 2}{\pi}\frac{\alpha_1\omega_B^2}{e^2v_g^2R_0T}\,q, \tag{23a}$$

where $\tilde{\tau}=\delta_I/(\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1})$. Finally, Eqs. (14b) and (14d) can be combined into

$$\frac{\partial x}{\partial r}=\frac{2\eta}{3P}\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial(rq)}{\partial r}-\frac{2}{v_g^2}\left[\tau_{\rm dis}^{-1}+\frac{\omega_B^2}{e^2\tilde{R}}\frac{4T^2\ln^2 2}{3Pv_g^2}\right]q$$
$$+\alpha_1\tilde{\tau}\frac{\pi^2T^2}{9Pv_g^2}\frac{\omega_B^2}{e^2\tilde{R}}\frac{\partial y}{\partial r}. \tag{23b}$$

Introducing the differential operator

$$\hat{\mathbb{D}}q=\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial(rq)}{\partial r}, \tag{24}$$

we rewrite Eqs. (22) in the matrix form

$$\hat{\mathbb{D}}\begin{pmatrix}q\\p\end{pmatrix}=\widehat{T}_S\begin{pmatrix}\partial x/\partial r\\\partial y/\partial r\end{pmatrix},\qquad \widehat{T}_S=\begin{pmatrix}\frac{1}{\tau_{\rm RE}} & 0\\ -\frac{1}{\tau_{\rm RE}} & \frac{1}{\tau_R}\end{pmatrix}. \tag{25a}$$

Similarly, Eqs. (23) can be written in the matrix form

$$\begin{pmatrix}\partial x/\partial r\\\partial y/\partial r\end{pmatrix}=-\widehat{M}\begin{pmatrix}q\\p\end{pmatrix}+\widehat{V}\hat{\mathbb{D}}\begin{pmatrix}q\\p\end{pmatrix}, \tag{25b}$$

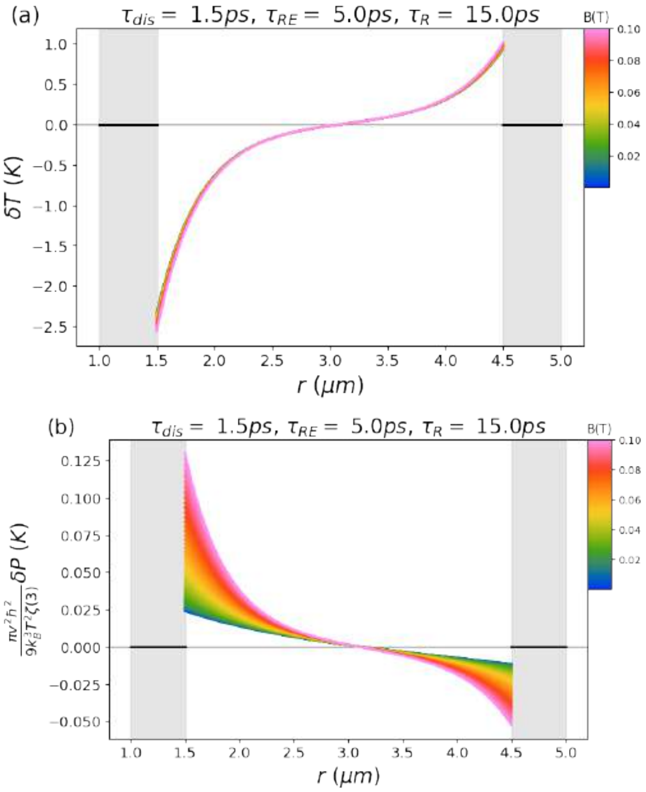


FIG. 4. Local variations of temperature (top panel) and pressure (bottom panel) in the Corbino device computed within the “supercollisions model” of energy relaxation. Black lines in the shaded regions indicate that the leads are at equilibrium. Color curves correspond to different values of the external magnetic field according to the shown color coding. Zero field values are identical with the results of Ref. [65]. For the parameter values, see Fig 1.

where

$$\widehat{V}=\begin{pmatrix}\frac{2\eta}{3P} & 0\\ 0 & 0\end{pmatrix},$$

and

$$\widehat{M}=\begin{pmatrix}\frac{16\ln^3 2}{3\pi}\frac{\delta_I\tilde{R}_BT^3}{v_g^4PR_0\tilde{\tau}}+\frac{2}{v_g^2\tau_{\rm dis}} & \frac{4\ln 2}{3}\frac{\alpha_1\delta_In_I\tilde{R}_BT}{v_g^2PR_0\tilde{\tau}}\\ \frac{24\ln^2 2}{\pi^2}\frac{\alpha_1\delta_I\tilde{R}_B}{v_g^2R_0\tilde{\tau}} & \frac{6}{\pi}\frac{n_I\tilde{R}}{R_0T^2\tilde{\tau}}\end{pmatrix}.$$

Finally, combining Eqs. (25) we find the equation for the variables p and q

$$\hat{\mathbb{D}}\begin{pmatrix}q\\p\end{pmatrix}=\widehat{K}\begin{pmatrix}q\\p\end{pmatrix},\quad \widehat{K}=\left[1-\widehat{T}_S\widehat{V}\right]^{-1}\widehat{T}_S\widehat{M}. \tag{26}$$

The obtained equation should be solved with the boundary conditions (9). The solution is straightforward albeit tedious. The results can be expressed in terms of linear combinations of the Bessel functions. Thus obtained solutions are not particularly instructive, hence we present the results of the calculation in graphical form.

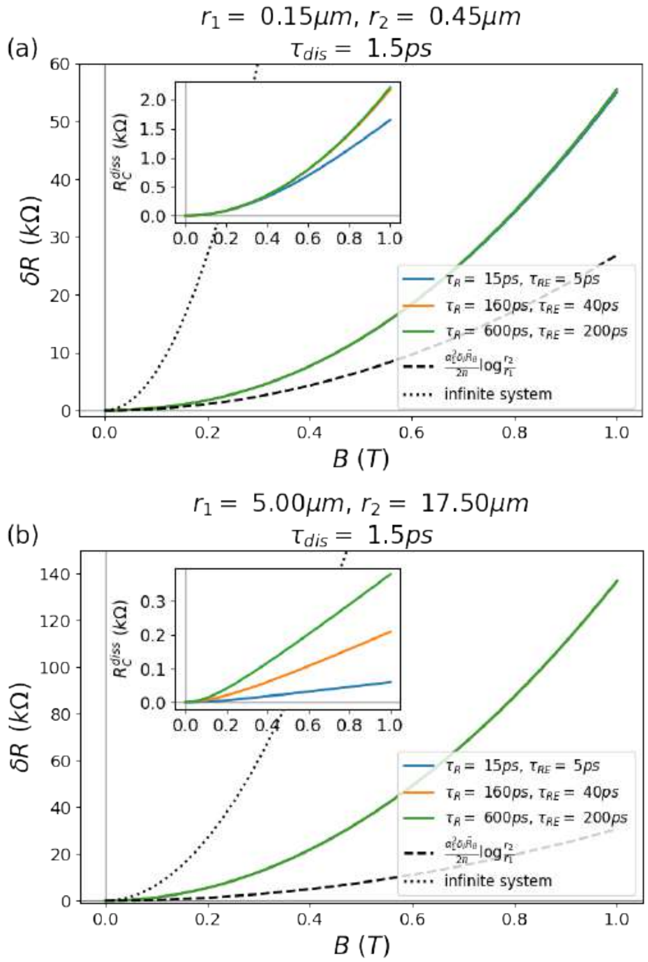


FIG. 5. Magnetoresistance of a small (top panel) and large (bottom panel) Corbino device computed within the "supercollisions model" of energy relaxation. The radii of the Corbino disks are shown above the plots. The black dotted line shows the quantity $\tilde{R}$, which is of the same order of magnitude as the magnetoresistance in the infinite system [23, 82]. Color curves correspond to three different sets of values of the relaxation times. For other parameter values (yielding $\ell_G = 0.2\,\mu$m), see Fig 1. The insets show the contact resistance due to viscous dissipation.

The radial component of the hydrodynamic velocity is shown in the top panel of Fig. 3. The drift velocity in the leads shows the standard Corbino profile, $u_r \propto 1/r$. At each interface, u_r exhibits a jump due to the mismatch of the entropy densities in the sample and leads. For high enough magnetic field, u_r has a sign change close to the interface. However, the corresponding change of direction is hardly seen in the overall flow diagram shown in Fig. 1, since the numerical value of the tangential component u_ϑ is much larger (see the bottom panel of Fig. 3).

The hydrodynamic velocity determines the energy current in the system. The nonuniform energy current results in local variations of the electronic temperature from its equilibrium value (see Fig. 4). The inhomogeneous temperature profile suggests that energy relaxation is less effective in strong magnetic fields. Fig. 1 shows the same data as Fig. 4 but in the form of the color map.

Finally we use the boundary conditions (12) to find the interface jumps of the electric potential which allows us to determine the device resistance. The procedure is the same as in the case of $\boldsymbol{B} = 0$ discussed in Ref. [65]. The results are shown in Fig. 5. For small enough samples (see the top panel in Fig. 5) the device resistance deviates only slightly from $\tilde{R}$ which is of the same order of magnitude as the magnetoresistance in the infinite system [23, 82]. In large samples the deviation is more pronounced and depends on the actual radius of the disk rather than on the ratio p (which is the same in both plots).

The quantitative results shown in this section were computed for a particular choice of the relaxation times. These values are largely phenomenological; however, the magnetoresistance shown in Fig. 5 hardly depends on them, while for larger samples (the bottom panel) the three curves are indistinguishable. However, the values of the relaxation times cannot be completely arbitrary. The point is that the matrix $\widehat{K}$ in Eq. (26) is not guaranteed to have real, positive eigenvalues (although its determinant is positive). In particular, the recombination time τ_R and energy relaxation time τ_{RE} cannot be very different. Within the physical model of supercollisions [59] these time-scales are of the same order of magnitude. Quasiparticle recombination involves supercollision scattering between the bands, while energy relaxation includes an additional contribution of intraband scattering. As a result, the energy relaxation time is shorter than τ_R, but not much shorter since the model does not involve any additional parameter. For such physical values of the relaxation times the eigenvalues of the matrix $\widehat{K}$ are real positive and the resulting magnetoresistance is well accounted for by the curves shown in Fig. 5 where, again, the particular values of τ_R and τ_{RE} do not have a strong quantitative impact on the overall resistance magnitude.

C. Energy relaxation due to electron-phonon interaction

Supercollisions are scattering events involving electron scattering off a phonon and an impurity. As such, this is a next-order process as compared to the direct electron-phonon scattering. The reason supercollisions might be important is that the speed of sound is much smaller than v_g. At high enough temperatures [59, 60] supercollisions indeed dominate, but at lower temperatures the direct electron-phonon scattering cannot be neglected.

Energy relaxation and quasiparticle recombination due to electron-phonon scattering was considered in Ref. [23] within the linear response theory. Since the macroscopic equations of the linear response theory coincide with the linearized hydrodynamic equations [22], we can directly incorporate the corresponding decay terms into our hydrodynamic theory. These decay terms appear in

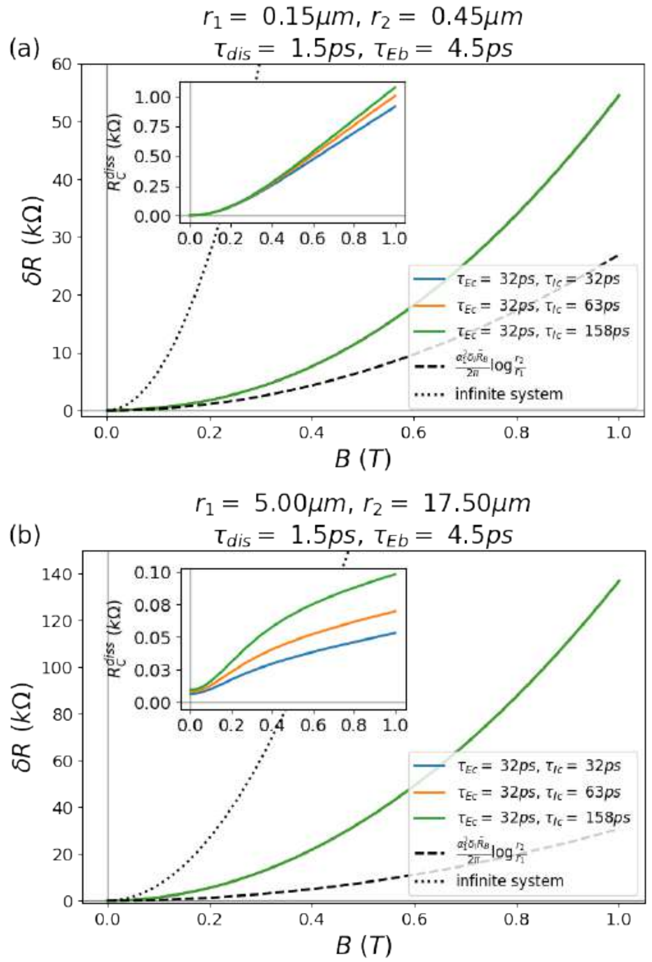


FIG. 6. Magnetoresistance in small (top) and large (bottom) Corbino devices computed within the "electron-phonon model" of energy relaxation (cf. Fig. 5).

Eq. (25a) through the matrix $\widehat{T}_S$. The model of electron-phonon interaction introduced in Ref. [23] corresponds to the following choice of this matrix

$$\widehat{T}_{ep} = -\frac{1}{|\Delta|}\begin{pmatrix} \frac{\gamma}{\tau_{Ec}}+\frac{1}{\tau_{Eb}} & -\frac{\gamma^2}{\mathcal{N}_2\tau_{Eb}}-\frac{\gamma}{\tau_{Ec}} \\ -\frac{\gamma^2\mathcal{N}_2}{\gamma\tau_{Ec}}-\frac{\mathcal{N}_2}{\tau_{Ic}}-\frac{1}{\tau_{Eb}} & \frac{2\gamma}{\tau_{Ec}}+\frac{\gamma^2}{\mathcal{N}_2\tau_{Eb}}+\frac{\mathcal{N}_2}{\tau_{Ic}} \end{pmatrix}, \tag{27}$$

where

$$\gamma = \frac{\pi^2}{12\ln^2 2}, \quad \mathcal{N}_2 = \frac{9\zeta(3)}{8\ln^3 2}, \quad \Delta = \gamma^2 - \mathcal{N}_2,$$

and $\tau_{Eb} \ll \tau_{Ec} \le \tau_{Ic}$ describe the three independent components of the electron-phonon collision integral [23].

Repeating the above calculation with $\widehat{T}_{ep}$ instead of $\widehat{T}_S$, we arrive at the results that are largely similar to those obtained within the supercollision model, but with a few notable differences (see Figs. 6-11). Unless specified in the figure captions, the parameter values used for the quantitative computation are the same as in the case of supercollisions (see the caption to Fig. 1).

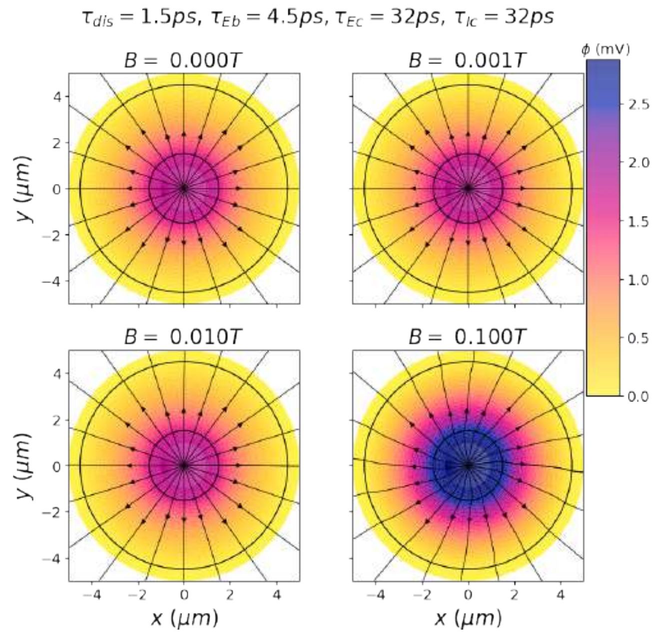


FIG. 7. Electric current density $\boldsymbol{j}$ and potential φ within the electron-phonon model of energy relaxation (cf. Fig. 2).

Magnetoresistance of the device is still positive and parabolic (see Fig. 6). In small devices, it is still largely determined by the quantity $\tilde{R}$ (shown by the black dotted line in Fig. 6 similarly to Fig. 5). In this case, variations of the electron-phonon relaxation rates still do not affect the result in any noticeable way. The results for large devices are also similar to the case of supercollisions: calculated magnetoresistance clearly exceeds $\tilde{R}$ and thus shows a strong dependence on the size of the device (but not on the ratio p).

The electric current density and potential in the device are seen largely the same as in the case of supercollisions, although the deviation of the current from the radial direction (i.e., its tangential component δj_ϑ) is somewhat smaller (see Fig. 7, cf. Fig. 2). This result seems to be consistent with the similarities in the magnetoresistance in the two cases.

The hydrodynamic velocity $\boldsymbol{u}$ is still dominated by its tangential component (see Figs. 8 and 9). The latter shows the behavior that is largely similar to that shown in the bottom panel of Fig. 3, although the magnitude of u_ϑ shows stronger growth with increasing magnetic field. In contrast, the temperature variation is "reversed": now the electronic temperature is increased around the inner contact and decreased close to the outer one (the opposite behavior to that seen in Figs. 1 and 4) (see Fig. 10).

The reversed behavior of the temperature variation corresponds to the change in the radial component of the hydrodynamic velocity u_r. While the jumps at the interfaces with the leads remain the same (insofar u_r on the sample side of the interface is larger than the drift velocity in the leads), the initial slope of u_r as a function

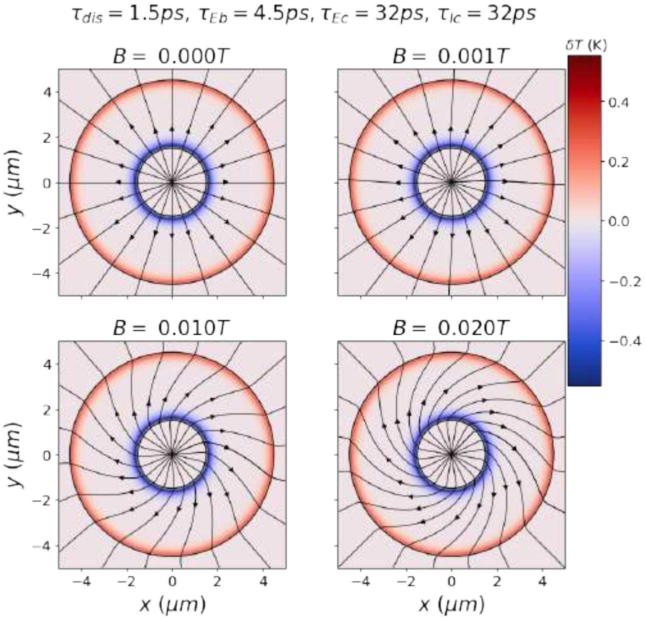


FIG. 8. Hydrodynamic velocity $\boldsymbol{u}$ and temperature δT distribution in the device obtained by solving the hydrodynamic equations at relatively low temperatures where energy relaxation is dominated by direct electron-phonon scattering (cf. Fig. 1).

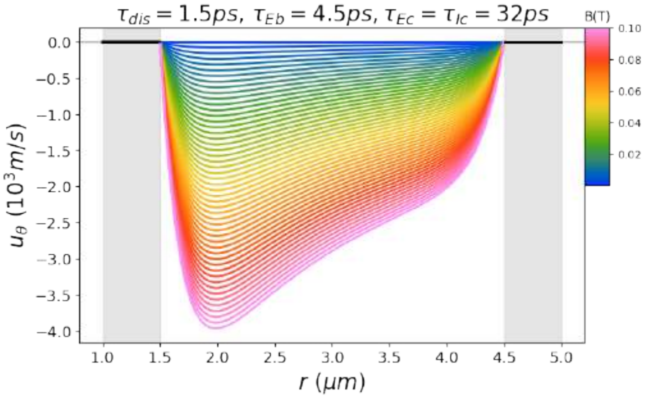


FIG. 9. Tangential component of the hydrodynamic velocity u_ϑ computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

of the radial coordinate has the opposite sign, which does not change with the increase in the magnetic field.

Overall, it is rather natural that the choice of the energy relaxation model mostly affects the energy flow in the device rather than the charge flow. This is a clear consequence of the decoupling of the energy and electric currents in neutral graphene. Although the two currents are being coupled by the magnetic field, the effect appears to be subleading. It is not surprising that the effect of this coupling is most pronounced in strong magnetic fields and large Corbino disks.

Contact resistance induced by viscous dissipation (see insets in Figs. 5 and 6) is also affected by the choice of the energy relaxation model. In the case of supercollisions its qualitative behavior exhibits a strong dependence on the size of the disk (see Fig. 5), while in the model of electron-phonon scattering this dependence is reduced to the magnitude only. The contact resistance is significantly stronger in small devices for both choices of the energy relaxation model as expected on general grounds.

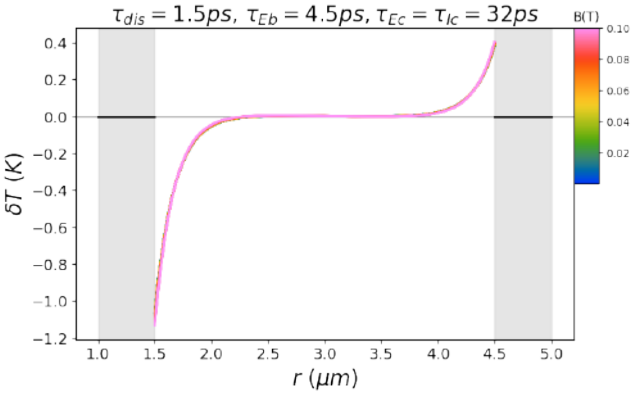


FIG. 10. Local temperature variation computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

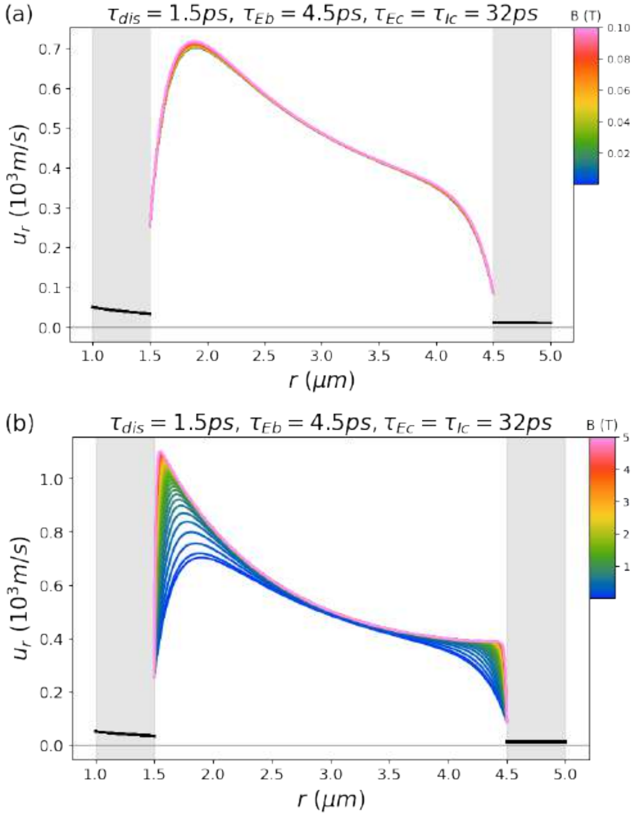


FIG. 11. Radial component of the hydrodynamic velocity u_r computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

IV. SUMMARY

In this paper we considered hydrodynamic flows of charge and energy in neutral graphene Corbino disks. We have shown that the Corbino geometry offers a (in principle realizable) possibility to measure electronic viscosity in neutral graphene, a task that so far has appeared elusive. The viscosity coefficient could be extracted from the magnetoresistance data in the ultra-clean limit where the bulk contribution to the device resistance is independent of the electron-impurity scattering time. The bulk resistance dominates over the contact resistance for larger sized disks and hence can in principle be measured in laboratory experiments.

Corbino magnetoresistance in graphene is illustrated in Figs. 5 and 6, where the calculated magnetoresistance is shown for two models of energy relaxation. In both cases, the dependence $R(B)$ is parabolic, similarly to the known result in the strip geometry. The viscosity coefficient can be *in principle* determined experimentally by analyzing the data in a set of different Corbino disks (see Sec. III A). This is not a straightforward task since the magnetoresistance is given by a sum of viscosity-dependent and viscosity-independent terms. In the clean limit $\ell_G \ll r_1, r_2$ [see Eq. (18)], these terms exhibit distinct dependence on the sample size r_2, the ratio of the radii $p = r_2/r_1$, and temperature, making it possible to extract the viscosity coefficient from the experimental data. In the opposite limit [see Eq. (19)], the dominant contribution to magnetoresistance is independent of viscosity. Existing experiments appear to be in the crossover between these two limits. In this paper we have used parameter values yielding $\ell_G \approx 0.2\,\mu$m. The size of the Corbino disk used in a recent experiment [31] was $r_1 = 2\,\mu$m, $r_2 = 9\,\mu$m, which is closer to the "large Corbino disk" illustrated in panels (b) in Figs. 5 and 6 than to the clean limit. It is fair to say that at present extracting viscosity from Corbino magnetoresistance measurements would be extremely difficult. At the same time, we are not aware of any other way to measure the viscosity coefficient in neutral graphene. We believe that viscosity measurements and more generally experimental observation of purely viscous effects in neutral graphene will be more accessible in the near future with even cleaner samples (increasing $\tau_{\rm dis}$ by an order of magnitude).

The regime of linear magnetoresistance as seen in the strip geometry or infinitely sized models does not exist in the Corbino geometry. This can be easily understood by noting that the origin of linear magnetoresistance is in the accumulation of energy and quasiparticle density in the boundary region of a long strip where the sample edges provide a natural barrier for the lateral neutral flow of quasiparticles induced by the magnetic field. In a Corbino disk there is no such edge. The lateral currents (energy and imbalance) flow freely around the disk without accumulating quasiparticles at any point.

Unlike the case of a single-band conductor (e.g., doped graphene), at charge neutrality the electric field is not expelled from the bulk of the sample. Nevertheless bulk viscous dissipation does lead to a discontinuity of the electric potential at the sample-lead interfaces inducing an additional contact resistance. This resistance however is rather small as compared to the resistance of the whole device and should not have a strong effect on the viscosity measurements.

ACKNOWLEDGMENTS

The authors are grateful to P. Hakonen, V. Kachorovskii, A. Levchenko, A. Mirlin, J. Schmalian, A. Shnirman, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 Program (Project No. 873028 HYDROTRONICS), by the German Research Foundation DFG project NA 1114/5-1 (B.N.N.), and by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant No. I-1505-303.10/2019 (I.V.G.).

---

[1] M. Polini and A. K. Geim, Phys. Today **73**, 28 (2020).
[2] B. N. Narozhny, Riv. Nuovo Cim. **45**, 661–736 (2022).
[3] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).
[4] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).
[5] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).
[6] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).
[7] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).
[8] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).
[9] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).
[10] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, T. T. K. Watanabe, A. Lucas, A. F. Young,

and A. C. Bleszynski Jayich (2020), arXiv:2002.05065.

[11] J. Mayzel, V. Steinberg, and A. Varshney, Nat. Commun. **10**, 937 (2019).

[12] I. Matthaiakakis, D. Rodríguez Fernández, C. Tutschku, E. M. Hankiewicz, J. Erdmenger, and R. Meyer, Phys. Rev. B **101**, 045423 (2020).

[13] P. Rao and B. Bradlyn, Phys. Rev. X **10**, 021005 (2020).

[14] R. Cohen and M. Goldstein, Phys. Rev. B **98**, 235103 (2018).

[15] P. K. Kovtun, D. T. Son, and A. O. Starinets, Phys. Rev. Lett. **94**, 111601 (2005).

[16] T. Schäfer, Annu. Rev. Nucl. Part. Sci. **64**, 125 (2014).

[17] J. Zaanen, Nat. Phys. **9**, 609 (2013).

[18] K. Kaushansky, M. A. Lichtman, J. T. Prchal, M. Levi, L. J. Burns, and D. C. Linch, eds., *Williams Hematology* (McGraw Hill, New York, 2021).

[19] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968), [Usp. Fiz. Nauk **94**, 689 (1968)].

[20] L. S. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).

[21] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).

[22] B. N. Narozhny, Ann. Phys. **411**, 167979 (2019).

[23] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

[24] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).

[25] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[26] M. Müller, L. Fritz, and S. Sachdev, Phys. Rev. B **78**, 115406 (2008).

[27] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).

[28] B. N. Narozhny, Phys. Rev. B **100**, 115434 (2019).

[29] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **103**, 115402 (2021).

[30] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).

[31] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, et al., Nature **609**, 276 (2022).

[32] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[33] U. Vool, A. Hamo, G. Varnavides, Y. Wang, T. X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C. A. C. Garcia, A. T. Pierce, et al., Nat. Phys. **17**, 1216 (2021).

[34] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[35] M. Kim, S. G. Xu, A. I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, et al., Nat. Commun. **11**, 2339 (2020).

[36] G. M. Gusev, A. D. Levin, E. V. Levinson, and A. K. Bakarov, Phys. Rev. B **98**, 161303(R) (2018).

[37] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Sci. Rep. **10**, 7860 (2020).

[38] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Phys. Rev. B **103**, 075303 (2021).

[39] O. E. Raichev, G. M. Gusev, A. D. Levin, and A. K. Bakarov, Phys. Rev. B **101**, 235314 (2020).

[40] A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021).

[41] G. Varnavides, A. S. Jermyn, P. Anikeeva, C. Felser, and P. Narang, Nat. Commun. **11**, 4710 (2020).

[42] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[43] A. Jaoui, B. Fauqué, and K. Behnia, Nat. Commun. **12**, 195 (2021).

[44] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys. Rev. B **92**, 165433 (2015).

[45] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).

[46] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[47] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[48] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).

[49] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).

[50] T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, Phys. Rev. B **100**, 245305 (2019).

[51] S. S. Pershoguba, A. F. Young, and L. I. Glazman, Phys. Rev. B **102**, 125404 (2020).

[52] S. Danz and B. N. Narozhny, 2D Materials **7**, 035001 (2020).

[53] J. L. M. Poiseuille, C. R. Acad. Sci. **11**, 961 (1840).

[54] J. L. M. Poiseuille, Annales de chimie et de physique (Series 3) **21**, 76 (1847).

[55] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1987).

[56] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018).

[57] H.-Y. Xie and M. S. Foster, Phys. Rev. B **93**, 195103 (2016).

[58] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[59] B. N. Narozhny and I. V. Gornyi, Frontiers in Physics **9**, 108 (2021).

[60] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).

[61] M. W. Graham, S.-F. Shi, D. C. Ralph, J. Park, and P. L. McEuen, Nat. Phys. **9**, 103 (2013).

[62] A. C. Betz, S. H. Jhang, E. Pallecchi, R. Ferreira, G. Fève, J.-M. Berroir, and B. Plaçais, Nat. Phys. **9**, 109 (2013).

[63] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **97**, 085415 (2018).

[64] J. F. Kong, L. Levitov, D. Halbertal, and E. Zeldov, Phys. Rev. B **97**, 245416 (2018).

[65] V. Gall, B. N. Narozhny, and I. V. Gornyi, Phys. Rev. B **107**, 045413 (2023).

[66] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[67] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[68] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).

[69] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).

[70] O. M. Corbino, Nuovo Cimento **1**, 397 (1911).

[71] A. Tomadin, G. Vignale, and M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).

[72] T. Holder, R. Queiroz, and A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).

[73] M. Shavit, A. V. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).

[74] A. Hui, V. Oganesyan, and E.-A. Kim, Phys. Rev. B **103**, 235152 (2021).

[75] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022).

[76] A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, and S. Ilani (2021), arXiv:2110.15369.

[77] O. E. Raichev (2022), arXiv:2202.06623.

[78] A. Levchenko, S. Li, and A. V. Andreev, Phys. Rev. B **106**, L201306 (2022).

[79] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).

[80] B. N. Narozhny, Annals of Physics **454**, 169341 (2023).

[81] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[82] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021).

[83] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[84] D. E. Sheehy and J. Schmalian, Phys. Rev. Lett. **99**, 226803 (2007).

[85] J. González, F. Guinea, and M. A. H. Vozmediano, Phys. Rev. B **59**, R2474 (1999).

[86] J. Gonzalez, F. Guinea, and M. A. H. Vozmediano, Nuclear Physics B **424**, 595 (1994).

[87] B. N. Narozhny, M. Titov, I. V. Gornyi, and P. M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).

[88] M. S. Steinberg, Phys. Rev. **109**, 1486 (1958).

[89] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).

[90] M. P. Marder, *Condensed Matter Physics* (Wiley, 2010).

[91] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021).

[92] M. Kumar, A. Laitinen, and P. Hakonen, Nature Communications **9**, 2776 (2018).

[93] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).

[94] G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).

[95] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).

[96] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).

Imaging Hydrodynamic Electrons Flowing Without Landauer-Sharvin Resistance

C. Kumar[1†], J. Birkbeck[1†], J.A. Sulpizio[1†], D. Perello[2,3], T. Taniguchi[4], K. Watanabe[4], O. Reuven[1], T. Scaffidi[5], Ady Stern[1], A.K. Geim[2,3], and S. Ilani[1*]

[1] *Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel.*
[2] *School of Physics & Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom.*
[3] *National Graphene Institute, University of Manchester, Manchester M13 9PL, United Kingdom.*
[4] *National Institute for Materials Science, 1-1 Namiki, Tsukuba, 305-0044 Japan.*
[5] *Department of Physics, University of Toronto, Ontario M5S 1A7, Canada.*
[†] These authors contributed equally to the work.
[*] Correspondence to: shahal.ilani@weizmann.ac.il

Electrical resistance usually originates from lattice imperfections. However, even a perfect lattice has a fundamental resistance limit, given by the Landauer[1] conductance caused by a finite number of propagating electron modes. This resistance, shown by Sharvin[2] to appear at the contacts of electronic devices, sets the ultimate conduction limit of non-interacting electrons. Recent years have seen growing evidence of hydrodynamic electronic phenomena[3–18], prompting recent theories[19,20] to ask whether an electronic fluid can radically break the fundamental Landauer-Sharvin limit. Here, we use single-electron-transistor imaging of electronic flow in high-mobility graphene Corbino disk devices to answer this question. First, by imaging ballistic flows at liquid-helium temperatures, we observe a Landauer-Sharvin resistance that does not appear at the contacts but is instead distributed throughout the bulk. This underpins the phase-space origin of this resistance - as emerging from spatial gradients in the number of conduction modes. At elevated temperatures, by identifying and accounting for electron-phonon scattering, we reveal the details of the purely hydrodynamic flow. Strikingly, we find that electron hydrodynamics eliminates the bulk Landauer-Sharvin resistance. Finally, by imaging spiraling magneto-hydrodynamic Corbino flows, we reveal the key emergent length-scale predicted by hydrodynamic theories – the Gurzhi length. These observations demonstrate that electronic fluids can dramatically transcend the fundamental limitations of ballistic electrons, with important implications for fundamental science and future technologies.

Electrical resistance is synonymous with the back-scattering of electrons. When an electron collides with an impurity, a phonon, or a rough device edge, it loses some of its momentum to the lattice, generating resistance. Therefore, it is very surprising that even when all these back-scattering sources are eliminated, an electronic device still has a non-zero resistance. As shown by Landauer[1], such resistance results from the finite conduction capacity of a channel, given by the number of its conduction modes multiplied by e^2/h (e is the electronic charge and h is Plank's constant). Sharvin[2] realized that this resistance should appear at the interface between the device and its contacts, where the number of conduction modes changes sharply. This Landauer-Sharvin resistance thus sets the ultimate resistance limit for ballistic electrons and gives a performance bound on real-life devices, where electrons are forced to transition frequently between metals and semiconductors.

A growing body of theoretical[21–32] and experimental[3–18] evidence suggests that when the interaction between electrons is sufficiently strong to dominate their scattering, the electronic system behaves as a hydrodynamic fluid. Key hydrodynamic features have been observed in transport[3–8,10,11,15,16] and imaging[9,12–14,17,18] experiments. Interestingly, transport measurements of hydrodynamic electrons flowing through constrictions[7] observed that they conducted up to 15% better than their ballistic counterparts. Theoretically, this was explained[28] by hydrodynamic lubrication of the constriction walls. This raises a question of fundamental and practical importance[19,20]: if the Landauer-Sharvin resistance limits ballistic electrons, what is the ultimate conduction limit for hydrodynamic electrons?

In this work, we show experimentally that hydrodynamic electrons can dramatically outperform ballistic electrons' limitations. By imaging electronic flows in a Corbino disk geometry, we observe that ballistic electrons exhibit roughly half of their Landauer-Sharvin resistance being distributed through the bulk of the device rather than at the contacts' interfaces. At elevated temperatures, we find that electron hydrodynamics efficiently eliminates this 'bulk Landauer-Sharvin' resistance. This observation is consistent with the recent theoretical prediction[20] that hydrodynamic electrons can flow without Landauer-Sharvin resistance. By adding a small magnetic field, we set the electrons into a spiraling motion, generating a viscous boundary layer near the contacts.

This layer provides the first real-space observation of the key emergent length scale of hydrodynamic theories - the Gurzhi length. Our findings demonstrate that hydrodynamics can dramatically modify the well-established rules for electrons obeyed by their ballistic counterparts.

To understand the origin of the Landauer-Sharvin resistance, consider ballistic transport through two device geometries: a straight channel (Fig. 1a) and a Corbino disk (Fig. 1b). In both cases, the total device resistance is given by the Landauer-Sharvin resistance, determined by the number of conduction modes traversing the device. In a straight channel, this resistance is given by $R_{sh} = \frac{\pi h}{4e^2}\frac{1}{k_F W}$, where k_F is the Fermi momentum and W is the channel width. In a Corbino geometry, the width is replaced by the circumference of the inner contact, $2\pi r_{in}$, and its total resistance is $R_{sh}^{in} = \frac{\pi h}{4e^2}\frac{1}{k_F(2\pi r_{in})}$. In a straight channel, this resistance drops only at the contact interfaces, with half dropping at each contact interface and none in the bulk (Fig. 1c). The situation changes in a Corbino geometry (Fig. 1d): here, similar to a straight channel, half of the Landauer-Sharvin resistance drops at the inner contact interface. However, curiously, theory predicts[19,20] (Fig. 1d) that the other half should be distributed across the bulk of the device.

The theoretical prediction that the Landauer-Sharvin resistance can be distributed across the bulk of a device highlights it's geometrical/phase-space origin – this resistance appears whenever there is a spatial gradient in the number of conduction modes. In a straight channel, the number of modes changes sharply at the contact interfaces but is fixed throughout the bulk (Fig. 1e). Consequently, the Landauer-Sharvin resistance appears only at the contact interfaces. In contrast, in a Corbino disk, the number of conduction modes gradually decreases with decreasing radius (Fig. 1f). An electron traveling from the outer to the inner contact thus experiences a gradually shrinking phase space, which should manifest as a distributed bulk Landauer-Sharvin resistance.

Interestingly, recent theory[19] has suggested that for hydrodynamic electrons in a Corbino geometry, the bulk Landauer-Sharvin resistance, which forms about half of the total device resistance, should vanish. A new theory[20] analyzed generalized flow geometries and showed that the Landauer-Sharvin resistance appears whenever there is a gradient in the number of conducting modes, and that it originates from the reflection enforced on electrons whose mode is terminated. Electron-electron scattering allows these electrons to smoothly transfer from a mode that is about to be terminated to a propagating mode, and thus reduces the resistance. Consequently, the theory predicted that for hydrodynamic electrons, resistance should occur only where the number of modes has a non-zero second spatial derivative [20]. Since in a Corbino disk the number of modes increases linearly with distance, this geometry is the ideal testbed to examine the elimination of the ballistic Landauer-Sharvin resistance for hydrodynamic flows.

The devices studied here consist of high mobility hBN-encapsulated monolayer graphene patterned into a Corbino disk geometry, with a graphite back gate that tunes the carrier density, n. The graphene spans a disk between radii r_{in} and r_{out}, where it connects to inner and outer Cr/Au contacts. The line connecting to the inner contact is deposited over the top hBN layer and a patch of crossed-linked resist, so the graphene disk is not perturbed, preserving its full angular symmetry (SI section 1). In the main text, we present data from a device with $r_{in} = 2\ \mu m$ and $r_{out} = 9\ \mu m$ (optical image in Fig. 1g). Similar results were obtained on a second Corbino device with different dimensions (SI section 8).

A major advantage of the Corbino device geometry over the more commonly used Hall-bar devices is the absence of etched edges and lithographically-defined voltage probes. This eliminates spurious scattering at lithographic features and edges, allowing to measure the unperturbed electron flow. This advantage comes at a price: transport experiments can only measure the overall 2-probe resistance of the device, and thus cannot decipher how this resistance is distributed in space. To solve this, we use a nanotube-based scanning single electron transistor[33] (SET) to spatially map the potential drop associated with the electronic current flow[34]. We drive an AC current, I, between the inner and outer Corbino contacts and use the SET to image the local electrostatic potential modulations at this AC frequency (Fig. 1h). This distinguishes the potential drop associated with the current flow from the static disorder potential, which we measure independently in DC in the absence of current. The spatial resolution of the measurement is limited by the scanning height of the SET above the device ($\sim 800\ nm$). In the figures below, we plot the imaged potential normalized by the total current, $R(x, y) = \phi(x, y)/I$, and define the zero of the potential at the outer contact, $\phi(r_{out}) = 0$. The quantity $R(x, y)$ therefore represents the resistance between the point (x, y) and the outer contact.

Fig. 2a shows a typical measured map of $R(x,y)$. Visibly, $R(x,y)$ rises monotonically from the outer to the inner contact. The rise is steeper at the graphene-contact interfaces ($r = r_{in}$ and $r = r_{out}$). Plotted as a colormap (inset), we see that $R(x, y)$ exhibits excellent angular symmetry, attesting to the high quality of the measured device. A similar level of angular symmetry is observed for the different temperatures, densities, and magnetic fields used throughout this work (SI section 3). This symmetry allows us to average over the angular direction and obtain an accurate radial resistance profile, $R(r)$, which conveys the essential information about the nature of the electron flow.

We begin with measurements of the resistance profile at low temperatures, where the transport is expected to be ballistic. Fig. 2b shows $R(r)$ measured at $T = 6\ K$ and $n = 4.5 \times 10^{11}\ cm^{-2}$ (similar phenomenology is also observed at other densities, SI section 7). We see that $R(r)$ starts flat at the outer contact, rises rapidly around $r = r_{out}$, climbs gradually throughout the bulk of the Corbino disk, then rises rapidly again around $r = r_{in}$, and finally flattens out at the inner contact. The overall resistance (the 2-probe resistance that would be measured in transport) can be read out directly from this graph to be $R_{tot} = 19.5\ \Omega$. To compare, the resistance of a completely ballistic Corbino device with perfect contacts is expected to be the Landauer-Sharvin resistance, R_{sh}^{in}. Plotting the same measurement but now normalized by R_{sh}^{in} (blue curve, Fig. 2c), we see that our device is not far from this ideal limit, $R_{tot}/R_{sh}^{in} = 1.42$.

The spatially resolved measurement now allows us to break the total resistance into the constituent contact and bulk components. The resistance of the graphene-contact interfaces leads to step functions of $R(r)$ at $r = r_{in}$ and $r = r_{out}$, but these are slightly smeared due to the finite resolution of our imaging. Using an complementary measurement on the same device, we accurately determine the point-spread-function (PSF) of our imaging (SI section 4). Thus, the only free parameter in these contact step functions, shown by $R_c^{in}(r)$ and $R_c^{out}(r)$ in Fig. 2c, is their height.

Remarkably, if we fit the measured $R(r)$ (blue, Fig. 2c) to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$ (dotted black, Fig. 2c), where $R_{LS}(r) = \frac{R_{sh}^{in}}{\pi} \operatorname{asin}\left(\frac{r_{in}}{r}\right)$ is the theoretically-predicted dependence of the Landauer-Sharvin bulk resistance[19,20], we find extremely close agreement throughout the bulk ($\sim 10\,\%$ difference). Note that the smeared contact functions, $R_c^{in}(r)$ and $R_c^{out}(r)$, penetrate very little into the bulk and therefore do not affect the quality of the fit in the bulk. The close agreement that we see is especially impressive given the fact that the expression for the Landauer-Sharvin bulk resistance has no free parameters. This measurement therefore provides the first real-space evidence of a distributed bulk Landauer-Sharvin resistance, originating from a spatial gradient of the number of conduction modes.

The small difference between the measurement and the ideal Landauer-Sharvin expression can be readily accounted for by weak impurity scattering. Such scattering leads to an ohmic term that depends logarithmically on r, $R_{ohm}(r)/R_{sh}^{in} = \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$, with one free parameter – the momentum-relaxing mean free path, l_{MR}. Adding this to $R_{LS}(r)$, we obtain an excellent fit to the measurement with $l_{MR} = 40\ \mu m$ (dashed red, Fig. 2c). Such a long mean free path is consistent with previous measurements on high mobility graphene devices[14,35], and is much longer than the Corbino channel length ($r_{out} - r_{in} = 7\mu m$), explaining the smallness of the ohmic contribution to the total resistance.

From the above fit we also obtain the magnitude of the graphene-contact interface resistances. The obtained inner contact resistance step height is $0.82R_{sh}^{in}$, larger than the $0.5R_{sh}^{in}$ expected for an ideal contact (Fig. 1d). This difference reflects a contact transparency of $T_{in} \sim 0.75$ (SI section 5), on par with the best transparencies achieved with graphene contacts[35,36]. By subtracting the fitted contact resistance steps from the measured profile, we obtain $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$ (inset, Fig. 2c). This quantity gives the most accurate description of the bulk resistance profile, even very close to the contacts, because it eliminates the smeared tails of the contacts step functions. In the remainder of the paper, we will use this quantity to explore the physics in the bulk.

The measured dependence of $R_{bulk}(r)$ on carrier density is shown in the left inset of Fig. 2d. We see that the total bulk resistance varies strongly with density (~ factor 4 over the measured density range). However, recalling that the number of conduction channels scales as $k_F \sim \sqrt{n}$, and normalizing each curve by R_{sh}^{in} at the corresponding density, we find that all curves collapse to a similar dependence (Fig. 2d, main panel) close to the Landauer-

Sharvin expression (dotted). This demonstrates that at $T = 6\ K$ the Landauer-Sharvin bulk resistance is the dominant contribution over a wide range of carrier densities. With decreasing n, the curves depart further from the ballistic limit, pointing to a growing ohmic component. Using a fit to $R_{LS}(r) + R_{ohm}(r)$, we obtain the n-dependence of l_{MR} (Fig. 2d right inset), in good agreement with previous measurements[12,14,35].

Having established the behavior at low temperatures, we proceed to explore the flow at elevated temperatures. Increased temperature increases both electron-phonon (e-ph) and electron-electron scattering. The former is momentum-relaxing and is thus expected to increase the device's resistance. Since the geometric and ohmic contributions are additive (as shown below), one may expect the total resistance to increase with the added ohmic resistance. For example, at $T{\sim}140\ K$, the previously measured[14,35,37] e-ph mean free path ($\sim 4\ \mu m$) implies that the total resistance should more than double.

The measured temperature dependence in Fig. 3a shows a surprisingly different behavior. The figure plots the measured $R_{bulk}(r)$ at temperatures $T = 6\ K$ to $140\ K$, where in all the curves we subtracted the same contact resistance steps, those obtained from the fit at $T = 6\ K$ (gray traces, Fig 2c). We see that instead of increasing as T increases, the resistance first decreases up to $T \approx 60\ K$ and then only mildly increases ($\sim 20\%$) toward $T = 140\ K$ (Fig. 3a inset). The measured spatial dependence of $R_{bulk}(r)$, plotted on a logarithmic r axis in Fig. 3a, provides a hint for the underlying physics: whereas at low T, the dependence is curved, as expected from an $\mathrm{asin}(r_{in}/r)$ dependence (bottom dashed line), at $T = 140\ K$ the dependence follows a perfectly straight line throughout almost the entire bulk of the device (top dashed line). Namely, at elevated T the resistance follows a pure $\log(r/r_{in})$ dependence. This suggests that as the ohmic e-ph contribution builds up, the contribution of the Landauer-Sharvin bulk resistance disappears, explaining why the resistance doesn't double as naively expected.

Given the unavoidable presence of e-ph scattering at elevated temperatures, even in the cleanest graphene samples, how can we resolve the clean-limit hydrodynamic flow profiles, namely those involving only electron-electron and not momentum-relaxing collisions? Here, the angular symmetry of the Corbino geometry proves advantageous. In the presence of this symmetry and within the relaxation time approximation there is a direct mapping between the flow with ohmic scattering and the clean-limit flow: If $R_{bulk}(r)_{L_{ee}}^{L_{MR}}$

is the profile with momentum-conserving and momentum-relaxing mean free paths L_{ee} and L_{MR}, then one can obtain from it the clean-limit profile, $\tilde{R}_{bulk}(r)$, by a mere subtraction of an ohmic term (the proof is given in SI section 9):

$$\tilde{R}_{bulk}(r)^{L_{MR}=\infty}_{L_{ee}=(l_{ee}^{-1}+l_{MR}^{-1})^{-1}} = R_{bulk}(r)^{L_{MR}=l_{MR}}_{L_{ee}=l_{ee}} - \frac{\hbar}{2e^2 k_F l_{MR}} \log(r/r_{in}) \tag{1}$$

The last term has only one free parameter, l_{MR}, which we obtain directly from independent magneto-hydrodynamic imaging experiments. Before discussing these experiments, we first substitute the obtained l_{MR} into equation (1), giving us directly $\tilde{R}_{bulk}(r)$ without adding any free parameters.

Figure 3b plots the obtained clean-limit hydrodynamic flow profiles, $\tilde{R}_{bulk}(r)$, at the various temperatures. At the lowest temperature, $\tilde{R}_{bulk}(r)$ follows the bulk Landauer-Sharvin dependence. With increasing T, however, this geometrical resistance gradually disappears. Remarkably, at $T = 140\ K$ the profile becomes completely flat throughout most of the bulk of the device, apart from small regions ($< 1\ \mu m$) near the inner and outer contacts.

To understand these measurements in more detail, we performed numerical Boltzmann calculations for a Corbino geometry with an electron–electron mean free path l_{ee}, taken within the relaxation time approximation (SI section 10). Fig. 3c shows the calculated $\tilde{R}_{bulk}(r)/R_{sh}^{in}$ for various l_{ee} values, smeared with the experimental PSF. For $l_{ee} = \infty$ the calculation recovers the Landauer-Sharvin dependence. With decreasing l_{ee} the Landauer-Sharvin resistance is gradually reduced. Once the mean free path becomes much shorter than the channel length, $l_{ee} \ll r_{out} - r_{in}$, the resistive drop occurs only at a distance $\sim l_{ee}$ from the contacts, nicely matching the measurements at elevated temperatures. This is the hydrodynamic buildup distance, over which the electron-electron interactions rearrange the flow from ballistic to hydrodynamic, and its accumulated resistance is the Stokes resistance[19] $\sim l_{ee}/k_F r_{in}^2$. The calculations also reproduce the appearance of a resistive outer contact step with increasing T, albeit stronger than observed experimentally (SI section 11). Most importantly, similar to the experimental result, we see that throughout most of the bulk of the disk, the Landauer-Sharvin geometrical resistance is completely eliminated by the hydrodynamic flow.

Finally, we turn to exploring hydrodynamic flow at non-zero magnetic fields. Corbino geometry is an ideal testbed for magneto-hydrodynamics due to the lack of disruptive physical edges. Fig. 4a shows the evolution of $R_{bulk}(r)$, measured at $T = 140\ K$, with a perpendicular magnetic field, B, ranging from zero to 30 mT. Evidently, the magnetic field increases the resistance throughout the Corbino channel. Plotting R_{bulk} vs. B at several radii (inset), we see positive $\sim B^2$ magnetoresistance (dashed lines).

Figure 4b presents the ratio between the radial electric field, obtained via a numerical derivative of the measured potential, $E_r = \frac{d\phi}{dr}$, and the radial current density, $j_r = I/2\pi r$. In contrast to a Hall bar geometry, where such a ratio between longitudinal field and current density gives the longitudinal resistivity, ρ_{xx}, in a Corbino geometry, this ratio yields the inverse longitudinal conductivity, σ_{xx}^{-1}. This is because, in the latter, due to angular symmetry the transverse field rather than the transverse Hall current is zero. We see that at $B = 0$, σ_{xx}^{-1} is independent of r throughout most of the disk's bulk. With increasing B the magnitude of σ_{xx}^{-1} increases. Interestingly, this increase is not constant in space, but is larger at the center of the conducting channel and smaller at its sides (dashed arrows).

Recalling that $\sigma_{xx}^{-1} = \rho_{xx} + \rho_{xy}^2/\rho_{xx}$ and $\rho_{xy} \sim B$, we see that the inverse conductivity, while being equal to the resistivity at $B = 0$, acquires an additional Hall component at finite B. This term arises from the appearance of an angular current density, j_θ, and a corresponding Hall angle, $\tan(\theta_{Hall}) = \rho_{xy}/\rho_{xx} = j_\theta/j_r$ (Fig. 4c inset), both of which grow linearly with B. By fitting $\sigma_{xx}^{-1}(r, B)$ to $\rho_{xx}(r) + a(r)B^2$ we get directly from $a(r)$ the dependence of $\tan(\theta_{Hall})/B$ on r, plotted in Fig. 4c. Visibly, $\tan(\theta_{Hall})$ is maximal at the center of the channel, but drops gradually toward the contacts. The length scale of the drop is much longer than our resolution limit (black line, Fig. 4c). Measurement of $\tan(\theta_{Hall})$ in the full 2D plane (Fig. 4d), shows a similar behavior: the Hall angle is largest at the center of the conducting channel, and gradually drops toward the contacts. Thus, in a finite magnetic field we observe a length scale that does not exists at $B = 0$, and describes the spatial change of θ_{Hall}.

Such an emergent scale was proposed by recent theories of magneto-hydrodynamic flow in a Corbino geometry[19,31], whose predicted flow lines are reproduced in Fig. 4e.

These lines are slanted with respect to the radial electric field by the local Hall angle, $\tan(\theta_{\mathrm{Hall}}(r)) = \frac{j_\theta}{j_r}$. At the center of the conducting channel, the Lorentz force and electric field balance to give the standard expression for the Hall angle, $\tan(\theta_{\mathrm{Hall}}) = \frac{l_{MR}}{R_c}$ (R_c is the cyclotron radius). However, near the contacts, θ_{Hall} must go to zero, since the electrons are injected from the contact isotropically, namely, $j_\theta = 0$. Formally, the theory gives:

$$\tan(\theta_H) = \frac{l_{MR}}{R_C} + C_1 r I_1\left(\frac{r}{l_G}\right) + C_2 r K_1\left(\frac{r}{l_G}\right), \tag{2}$$

where I_1,K_1 are the modified Bessel functions, C_1,C_2 are constants chosen such that $\tan(\theta_H) = 0$ at $r = r_{in}, r_{out}$, and $l_G = \sqrt{l_{MR} l_{ee}/4}$ is the Gurzhi length. The red curve in Fig. 4c plots the profile for $l_{MR} = 4.35\ \mu m$ and $l_{ee} = 1.3\ \mu m$, which agrees best with the experiment. In clean samples and elevated temperatures, l_{MR} is dominated by electron-phonon coupling and is thus rather universal. Indeed the l_{MR} we find here is in full agreement with earlier experiments[12,14,35]. Moreover, the l_{ee} obtained from the $B = 0$ experiment (Fig. 3) and from the magneto-hydrodynamic experiment (Fig. 4) closely agree, although they have completely different manifestations in these two flow regimes: in the former, l_{ee} gives the hydrodynamic buildup length and the Stokes resistance of the inner contact (Figs. 3b and 3c). In the latter, l_{ee} enters only through its geometrical average with l_{MR} to give the spatial scale of the θ_H gradient. Finally, we note that this is the first time that this key emergent length of electron hydrodynamics theory, the Gurzhi length, has been observed directly.

Our experiments demonstrate the intimate connection between the Landauer-Sharvin resistance and the spatial gradient in the number of conduction modes. Corbino devices have such gradients naturally, by virtue of their geometry. It is interesting to note that these devices are mathematically equivalent to devices with a simple (e.g., rectangular) geometry, in which the spatial gradient in the number of modes is caused by chemical doping rather than geometry[20]. This means that the physics discussed in this paper might even be relevant for future real-world devices. The observation that hydrodynamic electrons can dramatically outstrip the fundamental bounds of their ballistic counterparts, is thus of fundamental as well as technological importance.

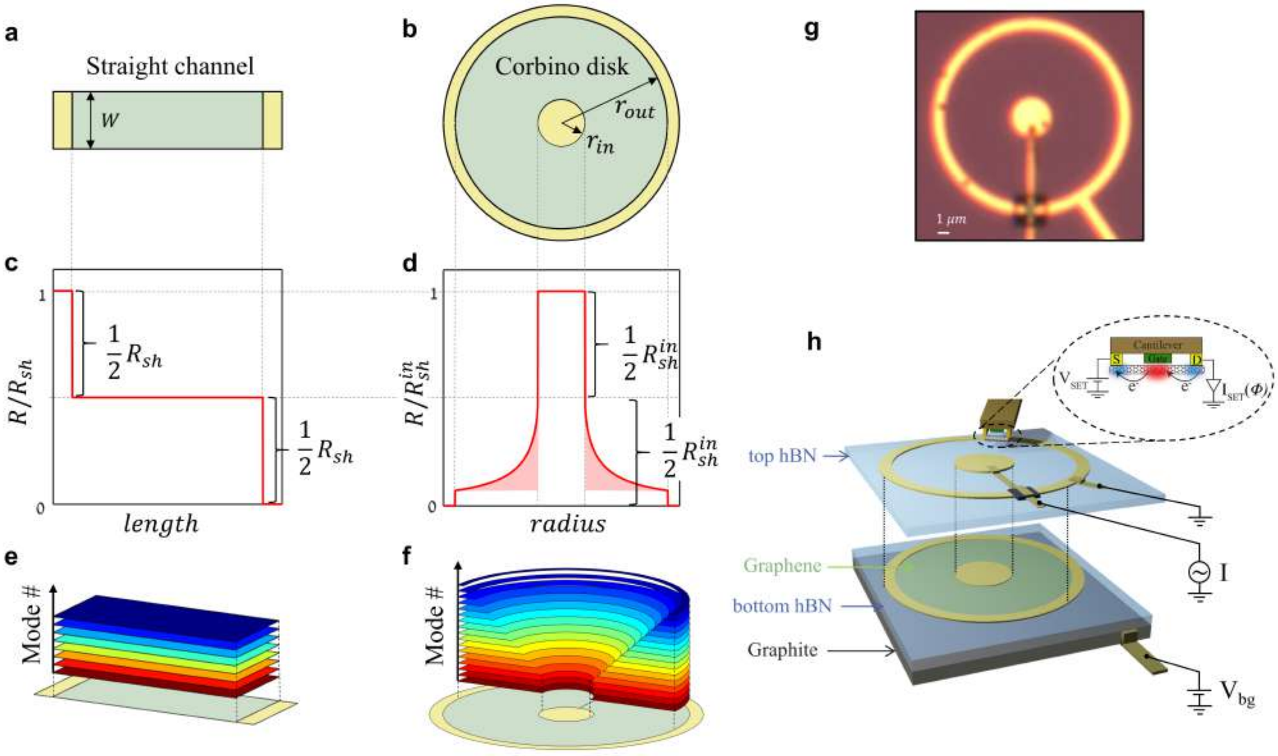


Figure 1: Landauer-Sharvin bulk geometrical resistance and the experimental setup for its measurement. Comparing two channel geometries: **a**. A straight channel (channel - green, contacts - yellow) **b**. Corbino disk channel (inner and outer radii are r_{in} and r_{out}, respectively). In the ballistic limit (no scattering, perfect edges) the total (2-probe) resistance of both geometries is inversely proportional to the number of conduction modes traversing the device, multiplied by e^2/h. For a straight graphene channel this gives $R_{sh} = \frac{\pi h}{4e^2}\frac{1}{k_F W}$ (W is the channel width and k_F is the Fermi momentum). For a Corbino disk, W is replaced by the inner contact circumference, giving a total resistance of $R_{sh}^{in} = \frac{\pi h}{4e^2}\frac{1}{k_F(2\pi r_{in})}$. **c.** In the straight channel, half of the Landauer-Sharvin resistance drops sharply at each contact interface and there is no resistive drop in the bulk. **d**. In the Corbino disk, similarly, half the Landauer-Sharvin resistance drops on the inner contact. However, the other half is distributed throughout the bulk of the Corbino disk, falling off as $R(r) = \frac{1}{2}R_{sh}^{in}\frac{2}{\pi}\mathrm{asin}(r_{in}/r)$. The Landauer-Sharvin resistance has a fundamental geometrical / phase-space origin – it appears whenever there is a spatial gradient in the number of conduction modes: **e**. In a straight channel the number of modes is constant throughout the bulk but changes sharply at the interfaces with the metallic contacts that effectively have an infinite number of modes. **f**. In a Corbino disk, there is a similar sharp change in the mode number at the inner contact interface, but throughout the bulk there is a linear increase of mode number with the radius, leading to the bulk Landauer-Sharvin resistance. **g**. Optical image of one of the studied devices ($r_{in} = 2\ \mu m$, $r_{out} = 9\ \mu m$). **h**. The device is composed of top hBN, graphene, bottom hBN and a graphite back gate, with inner (circular) and outer (ring) contacts. The carrier density is tuned by voltage V_{bg} on the graphite back-gate. We use a nanotube-based single electron transistor (SET) (inset) to image the potential in the device while flowing a current I between the contacts.

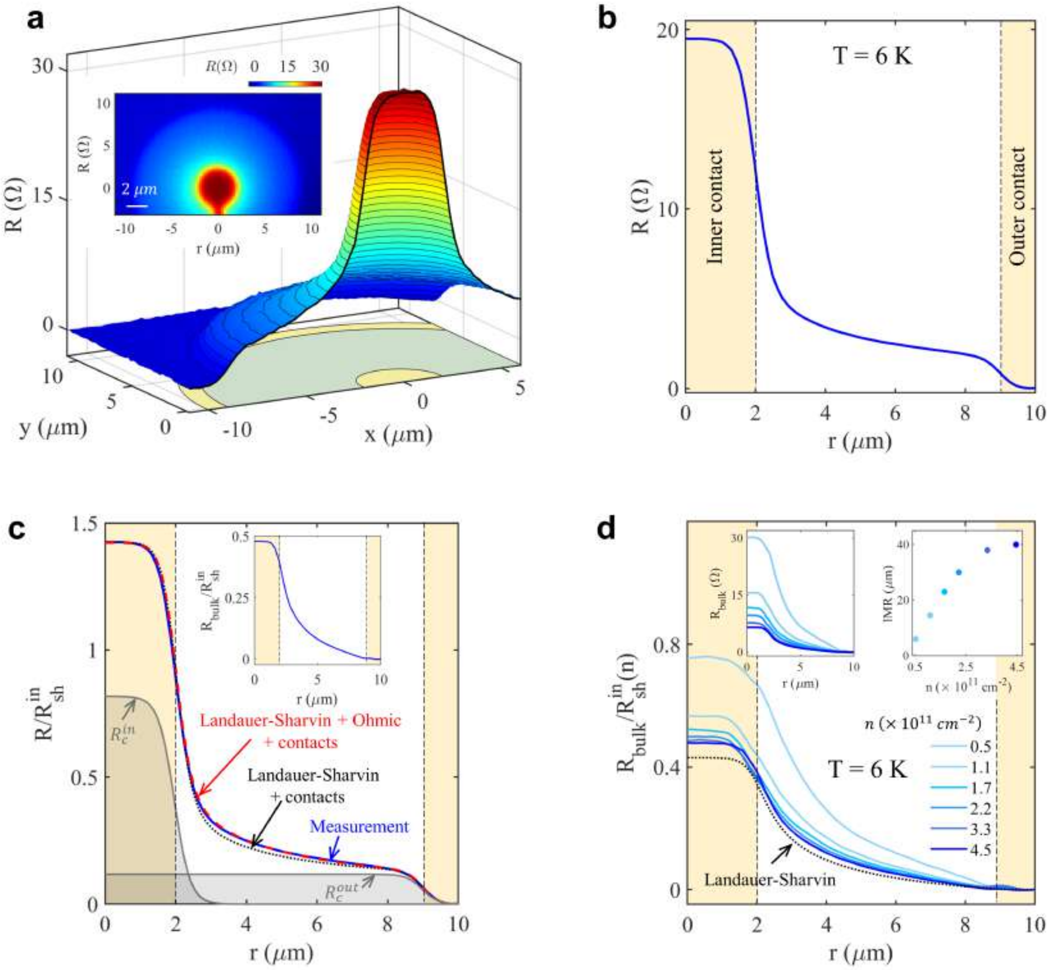


Figure 2: Imaging the Landauer-Sharvin bulk resistance in a Corbino disk. **a**. Spatially resolved image of the resistance $R(x,y) = \phi(x,y)/I$, where $\phi(x,y)$ and I are the measured potential and current, respectively, displayed above the schematic of the Corbino device. ($T = 140\ K$, $n = 4.5 \times 10^{11}\ cm^{-2}$). Inset: the same measurement presented as a colormap. The measurement exhibits excellent angular symmetry, allowing us to average along the angular direction and obtain $R(r)$. **b.** The measured radial dependence of the resistance, $R(r)$, at $T = 6\ K$, $n = 4.5 \times 10^{11}\ cm^{-2}$. Contacts are marked yellow and their interface with the graphene by dashed lines. **c**. Disentangling the different components of the resistance. The blue curve is the same measurement as in panel b, but now plotted normalized by R_{sh}^{in}. We fit this curve with a function that include bulk and contact dependence. The graphene-contact interface resistances are described by $R_c^{in}(r)$ and $R_c^{out}(r)$ (gray curves), which are step functions at $r = r_{in}$ and $r = r_{out}$, smeared by the point-spread-function (PSF) of our imaging experiment, which is measured separately in an complementary experiment (SI section 4). The dashed dotted line shows a fit of the measurement to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$, where $R_{LS}(r) = \frac{R_{sh}^{in}}{\pi} \operatorname{asin}\left(\frac{r_{in}}{r}\right)$ is the theoretically-predicted Landauer-Sharvin bulk geometrical resistance that has no free parameters. The dashed red line is a fit to a similar function which includes in addition an ohmic term, $R_{ohm}(r)/R_{sh}^{in} = \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$, with a momentum-relaxing mean free path of $l_{MR} = 40\ \mu m$. Inset: The bulk component of the resistance, obtained from the measured $R(r)$ by subtracting away the fitted contact resistance curves, $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$. **d**. Left inset: Measured $R_{bulk}(r)$ at various carrier densities, n (see key). Main panel: Same curves, but normalizing each curve by the Sharvin resistance at the corresponding density, $R_{sh}^{in}(n)$. Dotted line, theoretical Landauer-Sharvin bulk dependence. Right inset: l_{MR} vs. n obtained from fitting the graphs in the main panel to $R_{LS}(r) + R_{ohm}(r)$.

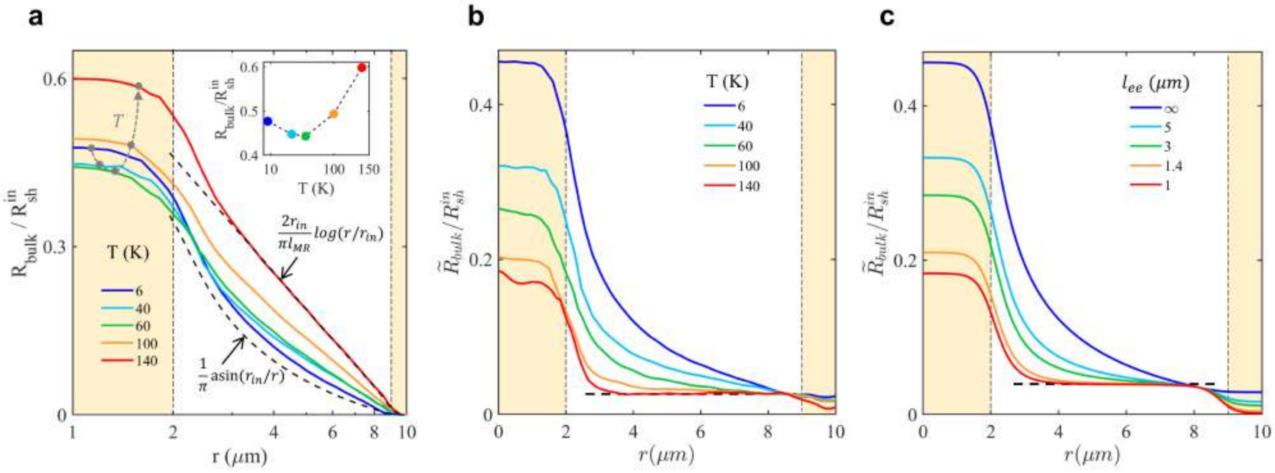


Figure 3. Observation of the perfect elimination of Landauer-Sharvin bulk resistance by hydrodynamic electron flow. **a**. Measured $R_{bulk}(r)$ at various temperatures, T, (see key) normalized by R_{sh}^{in} and plotted on a logarithmic r axis. Similar to Fig 2, $R_{bulk}(r)$ is obtained from the measured $R(r)$ by removing the contact resistance contribution. Note that we removed from the curves at the different temperatures the same contact resistance traces, those obtained from the fit at $T = 6\ K$ ($R_c^{in}(r)$ and $R_c^{out}(r)$ shown in gray in Fig. 2c). We see that the total bulk resistance first decreases with increasing T and then only mildly increases (gray dashed line). Inset: total bulk resistance normalized by R_{sh}^{in} as a function of T. The r dependence of the resistance evolves from Landauer-Sharvin dependence (bottom dashed line), at $T = 6\ K$, to a purely logarithmic dependence (top dashed line), at $T = 140\ K$, with a small deviation only very close to the inner contact ($< 1\ \mu m$). **b**. The clean-limit hydrodynamic flow profiles, $\tilde{R}_{bulk}(r)$, at various T's, obtained from the data in panel a using equation (1) and a momentum-relaxing mean-free path, l_{MR}, measured by a completely independent experiment on the same device (see main text). Note that in the bulk of the device these curves involve no free parameters. Notably, the distributed Landauer-Sharvin bulk resistance at $T = 6$ K is fully eliminated throughout most of the disk's bulk once the temperature has reached $T = 140$ K (the horizontal dashed black line is a guide to the eye). A small resistive component remains only close ($< 1\ \mu m$) to the inner contact. In addition, we see a slight increase of outer contact resistance with T **c**. Theoretical clean-limit hydrodynamic profiles, calculated using Boltzmann equations in the Corbino geometry. Different traces correspond to different electron-electron scattering length, l_{ee} (see key). Similar to the subtraction of the $T = 6\ K$ contact resistance steps in the experimental data, here we subtracted from all the curves the contact resistance steps calculated at $l_{ee} = \infty$. To allow for quantitative comparison with the experiment, the theory curves are smeared with the measured PSF of the imaging experiment. We observe a close correspondence between experiments and theory in the detailed profile shapes (see text).

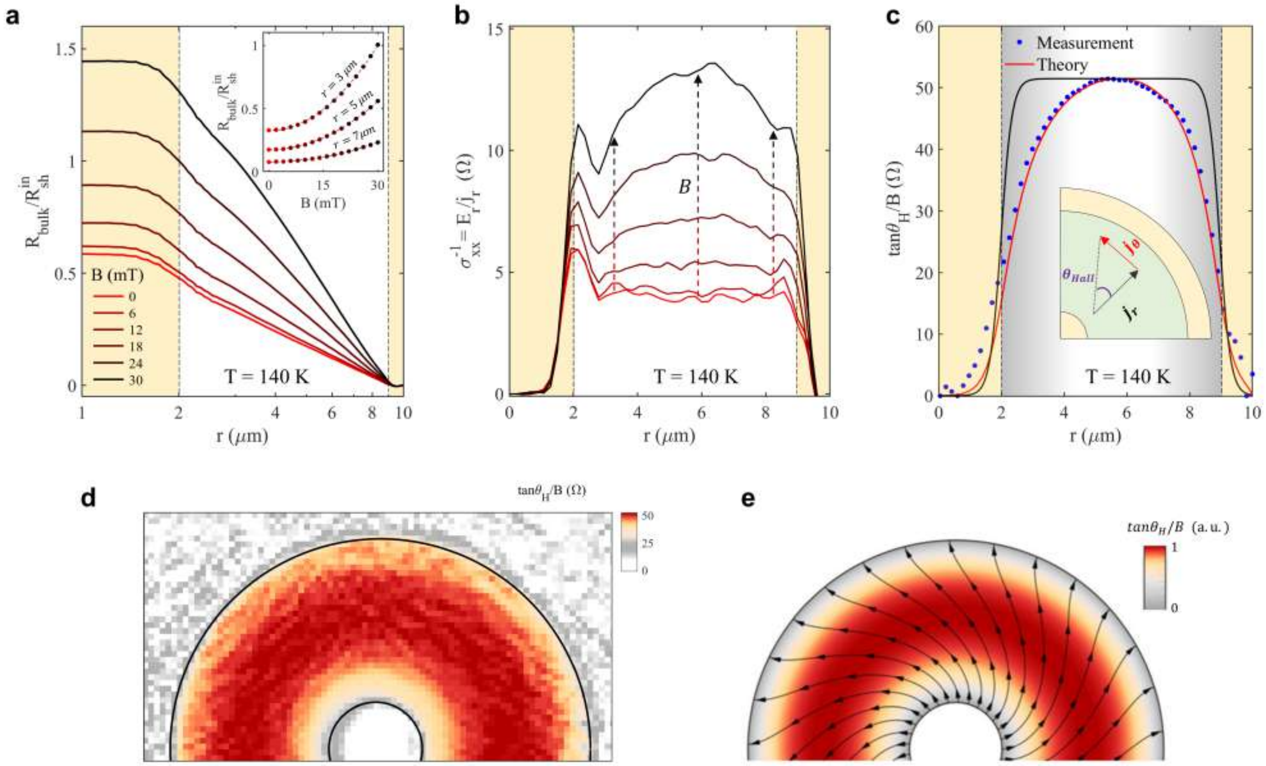


Figure 4. Imaging spiraling magneto-hydrodynamic electron flow and its Gurzhi boundary layer. **a.** $R_{bulk}(r)$, normalized by R_{sh}^{in}, measured at $T = 140\ K$ and various perpendicular magnetic fields, B (see key). Inset: R_{bulk}/R_{sh}^{in} at three spatial locations, $r = 3\ ,5\ ,7\mu m$, measured as a function of B. Dashed lines are parabolic fits. **b.** The inverse conductivity, $\sigma_{xx}^{-1} = E_r/j_r$, vs. r at various B's (same key as in panel a). E_r is the radial component of the electric field, obtained by numerically differentiating the measured potential, $E_r = d\phi/dr$, and $j_r = I/2\pi r$ is the current density obtained from the measured total current, I. Note that, while σ_{xx}^{-1} is independent of r throughout most of the disk's bulk at $B = 0$, the component added to σ_{xx}^{-1} at non-zero B is largest in the center of the Corbino channel and decays gradually toward the contacts (dashed arrows). **c**. Spatial dependence of the Hall angle, θ_H, obtained from fitting the quadratic-in-B term from panel b. The figure plots $\tan(\theta_H)\ /B$ (since θ_H increases linearly with B) as a function of r. While θ_H plateaus at the center of the channel, it drops gradually toward zero at the contacts. The drop happens over the Gurzhi length (gray shading). This length scale is considerably longer than our imaging resolution as is apparent from the black curve, which is a rectangular function convolved with our imaging PSF. The red curve is a fit to hydrodynamic theory (see text. Inset: diagram sketching the radial current density, j_r, the angular current density, j_θ, and the Hall angle, θ_H. **d**. $\tan(\theta_H)\ /B$, but now shown in a full 2D imaged spatial map. To obtain this image we measured $R(x, y)$ maps at $B = 0, 6, 12, 18, 24, 30\ mT$, then for each (x, y) point determined the quadratic-in-B term in the resistance, from which we obtain the local $\tan(\theta_H)\ /B$. Similar to panel c, which is the angular average of this measurement, even in this spatially resolved map we can clearly see that θ_H plateaus at the center of the channel, but drops gradually to zero at the contacts, over a rather long spatial scale. **e**. The calculated map of $\tan(\theta_H)\ /B$ using Navier-Stokes magneto-hydrodynamic equations and the parameters of the experiment in panels c and d. Overlaid are the flow lines. At the center of the channel, the flow lines are skewed from the radial direction following the standard expression for the Hall angle $\tan(\theta_H) = l_{MR}/R_c$. ($R_c$ is the cyclotron radius). The boundary condition at the contacts dictates that $j_\theta = 0$ and thus $\theta_H = 0$. The climb of θ_H from zero to its bulk value occurs over the Gurzhi length, $l_G = \sqrt{l_{ee}l_{MR}/4}$, corresponding closely to the length scale that emerged in the experiment.

Methods:

Device fabrication: Scanning SET devices were fabricated using a nanoscale assembly technique[38]. The graphene/hBN devices were fabricated using electron-beam lithography and standard etching and nanofabrication procedures[35] to define the channels and evaporation of Cr/Au (S4) to deposit contact electrodes.

Measurements: The measurements are performed on multiple graphene devices in home-built, variable temperature, Attocube-based scanning probe microscopes. The microscopes operate in vacuum inside liquid helium dewar with superconducting magnets, and are mechanically stabilized using Newport laminar flow isolators. A local resistive SMD heater is used to heat the samples under study from $T = 7.5$ K to $T = 150$ K, and a DT-670-BR bare chip diode thermometer mounted proximally to the samples and on the same printed circuit boards is used for precise temperature control. The voltage imaging technique employed is presented in reference[34]. Voltages and currents (for both the SET and sample under study) are sourced using a home-built DAC array, and measured using a home-built, software-based audio-frequency lock-in amplifier consisting of 1uV accurate DC+AC sources and a Femto DPLCA-200 current amplifier and NI-9239 ADC. The local gate voltage of the SET is dynamically adjusted via custom feedback electronics employing a least squares regression algorithm to prevent disruption of the SET's working point during scanning and ensure reliable measurements.

The voltage excitations applied to the graphene channels were as follows: 1 mV at $T = 6$ K, and 8 mV at $T = 140$ K. The magnetic fields applied ranged between ± 30 mT .

Acknowledgements: We thank L. Ella, G. Falkovich, L. Levitov, M. Polini, M. Shavit, A. Rozen, A. V. Shytov and U. Zondiner for useful discussions. Work was supported by the Leona M. and Harry B. Helmsley Charitable Trust grant, ISF grant (1182/21), Minerva grant (713237). TS acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), in particular the Discovery Grant [RGPIN-2020-05842], the Accelerator Supplement [RGPAS-2020-00060], and the Discovery Launch Supplement [DGECR-2020-00222]. During the preparation of this manuscript, we became

aware of a partially related STM work[17], which images voltage drops in flows across a constriction.

Contributions: C.K., J.B., J.A.S, A.K.G. and S.I. designed the experiment. C.K., J.B., J.A.S. performed the experiments. J.B., D.P. fabricated the devices. C.K., J.B., J.A.S, and S.I. analyzed the data. T.S. A.S. and S.I. wrote the theoretical model. K.W. and T.T. supplied the hBN crystals. C.K., J.B., A.S. and S.I. wrote the manuscript with input from other authors.

Data availability: The data that support the plots and other analysis in this work are available from the corresponding author upon request.

References

1. Landauer, R. Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction. *IBM Journal of Research and Development* **1**, 223–231 (1957).
2. Sharvin, Yu. v. A possible method for studying Fermi surfaces. *Sov. Phys. JETP* **21**, 655 (1965).
3. de Jong, M. J. M. & Molenkamp, L. W. Hydrodynamic electron flow in high-mobility wires. *Physical Review B* **51**, 389–402 (1995).
4. Bandurin, D. A. *et al.* Negative local resistance caused by viscous electron backflow in graphene. *Science* **351**, 1055–1058 (2016).
5. Crossno, J. *et al.* Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene. *Science* **351**, 1058–1061 (2016).
6. Moll, P. J. W., Kushwaha, P., Nandi, N., Schmidt, B. & Mackenzie, A. P. Evidence for hydrodynamic electron flow in PdCoO 2. *Science* **351**, 1061–1064 (2016).
7. Krishna Kumar, R. *et al.* Superballistic flow of viscous electron fluid through graphene constrictions. *Nature Physics* **13**, 1182–1185 (2017).

8. Gooth, J. *et al.* Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide. *Nature Communications* **9**, 4093 (2018).
9. Braem, B. A. *et al.* Scanning gate microscopy in a viscous electron fluid. *Physical Review B* **98**, 241304 (2018).
10. Berdyugin, A. I. *et al.* Measuring Hall viscosity of graphene's electron fluid. *Science (New York, N.Y.)* **364**, 162–165 (2019).
11. Tan, C. *et al.* Realization of a universal hydrodynamic semiconductor in ultra-clean dual-gated bilayer graphene. *arXiv:1908.10921* (2019).
12. Sulpizio, J. A. *et al.* Visualizing Poiseuille flow of hydrodynamic electrons. *Nature* **576**, 75–79 (2019).
13. Ku, M. J. H. *et al.* Imaging viscous flow of the Dirac fluid in graphene. *Nature* **583**, 537–541 (2020).
14. Jenkins, A. *et al.* Imaging the breakdown of ohmic transport in graphene. *arXiv:2002.05065* (2020).
15. Keser, A. C. *et al.* Geometric Control of Universal Hydrodynamic Flow in a Two-Dimensional Electron Fluid. *Physical Review X* **11**, 031030 (2021).
16. Gupta, A. *et al.* Hydrodynamic and Ballistic Transport over Large Length Scales in GaAs/AlGaAs. *Physical Review Letters* **126**, 076803 (2021).
17. Krebs, Z. J. *et al.* Imaging the breaking of electrostatic dams in graphene for ballistic and viscous fluids. *arXiv:2106.07212* (2021).
18. Vool, U. *et al.* Imaging phonon-mediated hydrodynamic flow in WTe2. *Nature Physics* 1–5 (2021) doi:10.1038/s41567-021-01341-w.
19. Shavit, M., Shytov, A. & Falkovich, G. Freely Flowing Currents and Electric Field Expulsion in Viscous Electronics. *Physical Review Letters* **123**, 026801 (2019).
20. Stern, A. *et al.* Spread and erase - How electron hydrodynamics can eliminate the Landauer-Sharvin resistance. *arXiv: 2110.15369* (2021).
21. Gurzhi, R. N. Minimum of Resistance in Impurity Free Conductors. *Sov. Phys. JETP* **17**, 521 (1963).
22. Nagaev, K. E. & Ayvazyan, O. S. Effects of electron-electron scattering in wide ballistic microcontacts. *Physical Review Letters* **101**, 1–4 (2008).

23. Nagaev, K. E. & Kostyuchenko, T. v. Electron-electron scattering and magnetoresistance of ballistic microcontacts. *Physical Review B - Condensed Matter and Materials Physics* **81**, 1–9 (2010).
24. Andreev, A. v., Kivelson, S. A. & Spivak, B. Hydrodynamic Description of Transport in Strongly Correlated Electron Systems. *Physical Review Letters* **106**, 256804 (2011).
25. Torre, I., Tomadin, A., Geim, A. K. & Polini, M. Nonlocal transport and the hydrodynamic shear viscosity in graphene. *Physical Review B* **92**, 165433 (2015).
26. Levitov, L. & Falkovich, G. Electron viscosity, current vortices and negative nonlocal resistance in graphene. *Nature Physics* **12**, 672–676 (2016).
27. Scaffidi, T., Nandi, N., Schmidt, B., Mackenzie, A. P. & Moore, J. E. Hydrodynamic Electron Flow and Hall Viscosity. *Physical Review Letters* **118**, 226601 (2017).
28. Guo, H., Ilseven, E., Falkovich, G. & Levitov, L. S. Higher-than-ballistic conduction of viscous electron flows. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 3068–3073 (2017).
29. Narozhny, B. N., Gornyi, I. v., Mirlin, A. D. & Schmalian, J. Hydrodynamic Approach to Electronic Transport in Graphene. *Annalen der Physik* **529**, 1700043 (2017).
30. Holder, T. *et al.* Ballistic and hydrodynamic magnetotransport in narrow channels. *Physical Review B* **10**, 245305 (2019).
31. Levchenko, A. & Schmalian, J. Transport properties of strongly coupled electron–phonon liquids. *Annals of Physics* **419**, 168218 (2020).
32. Hong, Q., Davydova, M., Ledwith, P. J. & Levitov, L. Superscreening by a Retroreflected Hole Backflow in Tomographic Electron Fluids. *arXiv: 2012.03840* **1**, 1–8 (2020).
33. Honig, M. *et al.* Local electrostatic imaging of striped domain order in LaAlO3/SrTiO3. *Nature materials* **12**, 1112–8 (2013).
34. Ella, L. *et al.* Simultaneous voltage and current density imaging of flowing electrons in two dimensions. *Nature Nanotechnology* **14**, 480–487 (2019).
35. Wang, L. *et al.* One-Dimensional Electrical Contact to a Two-Dimensional Material. *Science* **342**, 614–617 (2013).

36. Ben Shalom, M. *et al.* Quantum oscillations of the critical current and high-field superconducting proximity in ballistic graphene. *Nature Physics* **12**, 318–322 (2016).

37. Efetov, D. K. & Kim, P. Controlling Electron-Phonon Interactions in Graphene at Ultrahigh Carrier Densities. *Physical Review Letters* **105**, 256805 (2010).

38. Waissman, J. *et al.* Realization of pristine and locally tunable one-dimensional electron systems in carbon nanotubes. *Nature Nanotechnology* **8**, 569–574 (2013).

Supplementary materials for:

Imaging Hydrodynamic Electrons Flowing Without Landauer-Sharvin Resistance

Chandan Kumar[†], John Birkbeck[†], Joseph A. Sulpizio[†], David J. Perello, Takashi Taniguchi, Kenji Watanabe, Oren Reuven, Thomas Scaffidi, Ady Stern, Andre K. Geim, and Shahal Ilani[*]

Contents

S1. Device fabrication

Our devices consist of monolayer graphene encapsulated between two hexagonal boron nitride (hBN) layers, with a graphite flake underneath acting as a back-gate. These devices are assembled using the standard dry transfer technique[1]. Briefly, a polypropylene carbonate (PPC) coated polydimethylsiloxane (PDMS) was used to pick up the top hBN, monolayer graphene, and the bottom hBN, and this stack was then dropped on a graphite flake. After assembly, the heterostructure was annealed under ultra-high vacuum at 500 °C for ~ 8 hrs, followed by ironing with an AFM tip in contact mode. Both steps helped clean any residues from the surface of the top hBN and improve the sample's homogeneity by removing bubbles and ripples at the hBN-graphene interface[2]. Fig. S1a shows the optical image of the heterostructure. The monolayer graphene area (marked with a white dashed line) is free of bubbles/wrinkles. Next, we defined the Corbino disk using a standard electron beam lithography process, followed by reactive ion etching (RIE) with $CHF_3 + O_2$ and metal deposition of Cr (2 nm) / Au (80 nm) (Fig. S1b).

For making contact to the inner circular disk without electrically shorting to the outer metal ring, a bridge was defined on a small section of outer contact (highlighted by the dotted black line in Fig. S1c). The bridge was made by cross-linking polymethyl methacrylate (PMMA), using a high dose (x1000 regular dose) during electron beam lithography. Finally, in another electron beam lithography step, we defined a lead that passes over the cross-linked PMMA, ensuring that the inner circular disk and the outer contact ring are electrically isolated.

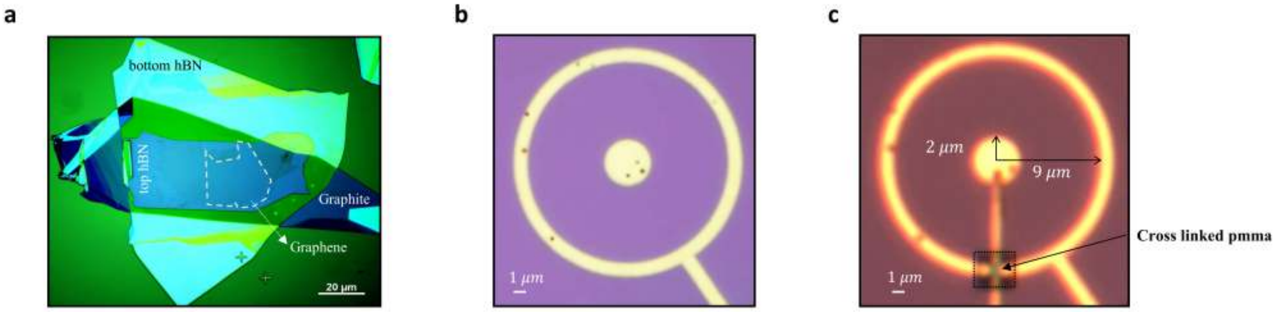


Fig. S1 | Device fabrication. **a.** Optical image of the heterostructure (hBN/Graphene/hBN/Graphite) after annealing and AFM ironing. The graphene region is demarcated with a white dashed line and is free of bubbles and wrinkles. The bottom graphite acts as a back-gate. The thicknesses of the top and bottom hBN are $40\,nm$ and $76\,nm$, respectively. **b.** Optical image of the heterostructure after making 1D contacts that define the Corbino disk geometry. **c.** Optical image of the Corbino disk after contacting the inner circular disk. The contact to the inner circular disk passes over cross-linked PMMA, defined over a small segment of outer contact (dashed black), electrically isolating the inner contact from the outer ring. The inner and outer radius of the graphene disk in this device are $r_{in} = 2\ \mu m$ and $r_{out} = 9\ \mu m$ (arrows), respectively.

S2. Transport measurements

We use standard lock-in techniques to measure the two-probe resistance between the inner and outer contacts of the Corbino disks as a function of carrier density, n, and temperature, T (from $6\ K$ to $140\ K$). We find that a significant component of the resistance comes from the lithographic lines that lead to the device. Using the Scanning SET imaging we accurately determine these line resistances by imaging the potential drop between the voltage source and the actual potential of the metal contact, measured by the SET. This line resistance was measured as a function of temperature and back-gate voltage, V_{BG} (which controls the carrier density) and found that it depends on T but not on V_{BG}. The measured total line resistance at $T = 6\ K$ is $R_{lines} = 515\ \Omega$ and it increases with increasing temperature, reaching $R_{lines} = 615\ \Omega$ at $T = 140\ K$. We subtract this SET-measured line resistance from the transport-measured two-probe resistance to obtain an effective four-probe resistance. Note that this four-probe resistance still includes the metal-graphene contact resistance, which includes both the fundamental Landauer-Sharvin component, and the component due to imperfect contacts. Fig. S2a plots this effective four-probe resistance as a function of carrier density, n, at various temperatures (see legend). Independently, we can determine the total device resistance directly from the SET measurements by reading out from the imaged $R(r)$ (e.g. Fig. 2b, main text) the difference in its value between the outer and inner contacts. The inset to Fig. S2a compares the SET measured device resistance (red dots) and the effective four-probe transport measured device resistance (blue) at $T = 6$ K. We can see an excellent agreement between the two measurements. (Note that the removed line resistance is independent of density and has excellent agreement over the entire density range). Plotting the conductance at $T = 6\ K$ as a function of $\sqrt{n}$ we can estimate the charge inhomogeneity in our samples from the flat region of conductivity near the Dirac point, which comes out to be $\delta n \sim 5 \times 10^{9}\ cm^{-2}$ (Fig. S2b). In the main text, data is presented for the electron-doped region. Most of the data (apart from Fig. 2d) are taken at a gate voltage of, $V_{BG} = 2\ V$ ($n = 4.5 \times 10^{11}\ cm^{-2}$), though we see similar phenomenology also at other densities (see e.g. section S7 below).

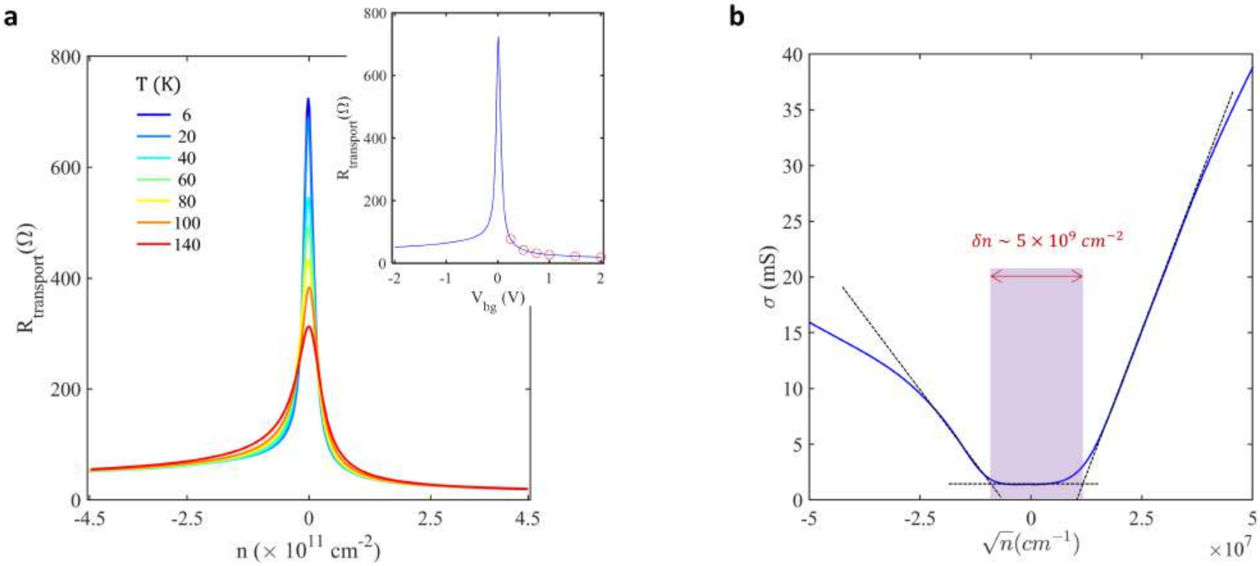


Fig. S2 | Transport measurements of the Corbino device in the main text. Main panel: effective four-probe transport resistance, $R_{transport}$, as a function of density, n, and at various temperatures, T (see legend). To obtain $R_{transport}$, we use standard lock-in measurements of the two-probe resistance between the Corbino disk's contacts. From this we subtract the lines resistance imaged using the scanning SET. We find that the line resistance depends on temperature but not on the back-gate voltage, V_{BG}. Inset: Corbino device resistance as a function of V_{BG} at $T = 6\ \mathrm{K}$. The blue line was taken from the main panel (transport measurement with line resistance subtracted) the red dots are obtained from imaging measurements of $R(r)$ using the scanning SET (e.g. Fig. 2b, main text). **b**. Transport measured conductivity, σ, at $T = 6\ \mathrm{K}$, plotted as a function of $\sqrt{n}$. The width of the plateau at the center provides an estimate for the charge disorder in the sample: $\delta n \sim 5 \times 10^9\ cm^{-2}$.

S3. Angular symmetry of the measured flow

In Figs. 2,3,4 of the main text we plot the resistance profile, $R(r)$, obtained by angular averaging of two-dimensional SET images of $R(x,y)$. This procedure is valid as long as the physics has a high degree of angular symmetry. We demonstrate this symmetry below using a specific measurement and note that a similar level of symmetry exists also for the data measured at other carrier densities and temperatures.

Figs. S3 plots a spatial scan of $R\ (x,y)$ measured at $B = 30\ mT, V_{BG} = 2\ V$ and $T = 100\ K$. Similar to the scan in Fig. 2a in the main text, which was performed at a different temperature and magnetic field, also here we see that the measurement exhibits excellent angular symmetry. Panels a and b show the same spatial scan, but with different averaging regions marked by a shaded "pizza" slice (outlined in black and red, respectively, in the two panels). The two resulting $R(r)$ profiles are shown in panel c with the corresponding colors. We can see that the two profiles are practically identical.

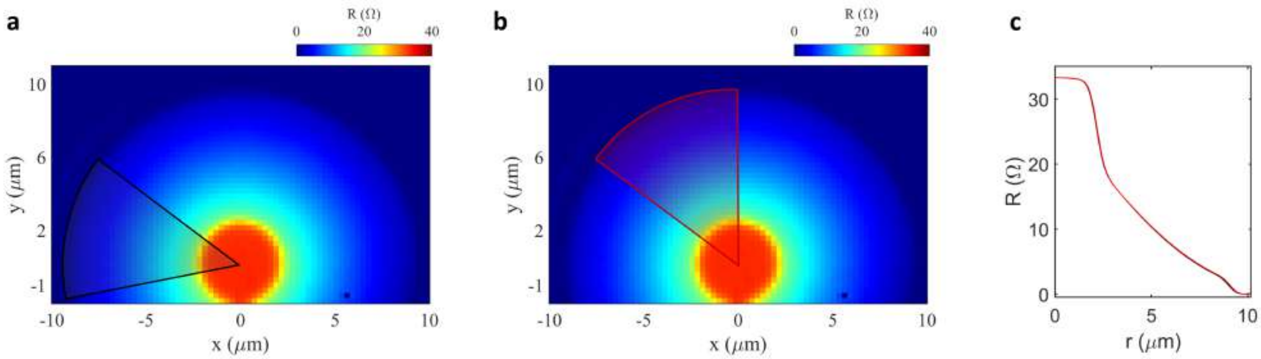


Fig. S3: **Angular symmetry of the measured data. a,b** In both panels we show the same spatial map of the resistance, $R(x,y)$, measured at $B = 30\ mT$, $V_{BG} = 2\ \text{V}$ and $T = 100\ \text{K}$. The shaded slices with black (panel **a**) and red (panel **b**) outline marks the regions used for the angular averaging. **c.** The resistance profiles, $R(r)$, obtained from averaging over the slices in a and b, plotted with corresponding black and red colors.

S4. Measurement of the point spread function (PSF) of the imaging experiments.

The scanning SET measurements have a finite spatial resolution, determined by the scanning height of the SET above the graphene. This manifests in our scans as a spatial smearing with a point spread function (PSF) that depends on our scanning height. To accurately compare our measurements with the theory we extract the PSF from an independent imaging experiment and convolve the theory curves with this PSF. To determine the PSF we use an experiment that images the workfunction, $W(x,y)$, shown in Fig. S4a, and follow the recipe discussed in our previous work[3]. In contrast to the measurements of $R(x,y)$ which probe the potential that is generated by an electronic current (out of equilibrium), measurement of $W(x,y)$ probe the static (equilibrium) potential, in the absence of current. These are therefore measurements of completely independent properties.

In Fig. S4a we can see that the workfunction is constant throughout most of the graphene disk. On the central gold contact the workfunction is also constant, but with a different value. The transition between these two constants occurs very sharply, over lithographic scales. The workfunction image thus yields a sharp rise, ideal for determining the imaging PSF.

Fig. S4b plots the measured workfunction (blue dots) along the blue dashed radial line in Fig. S4a. Compared to an ideal step function positioned at $r_{in} = 2\ \mu m$ (black line), we see that the measurement is spatially smeared. The red curve shows the step function convolved with a PSF given by $g(r) = 1/\cosh^2\left(1.76\frac{r}{\sigma}\right)$. In our previous work[3] we demonstrated that this PSF describes well the smearing in the experiment and that σ corresponds to our scanning height. We find a good fit between the measurement and the step function PSF-smeared with a height value of $\sigma = 0.85\ \mu m$. This is the PSF used in the main text.

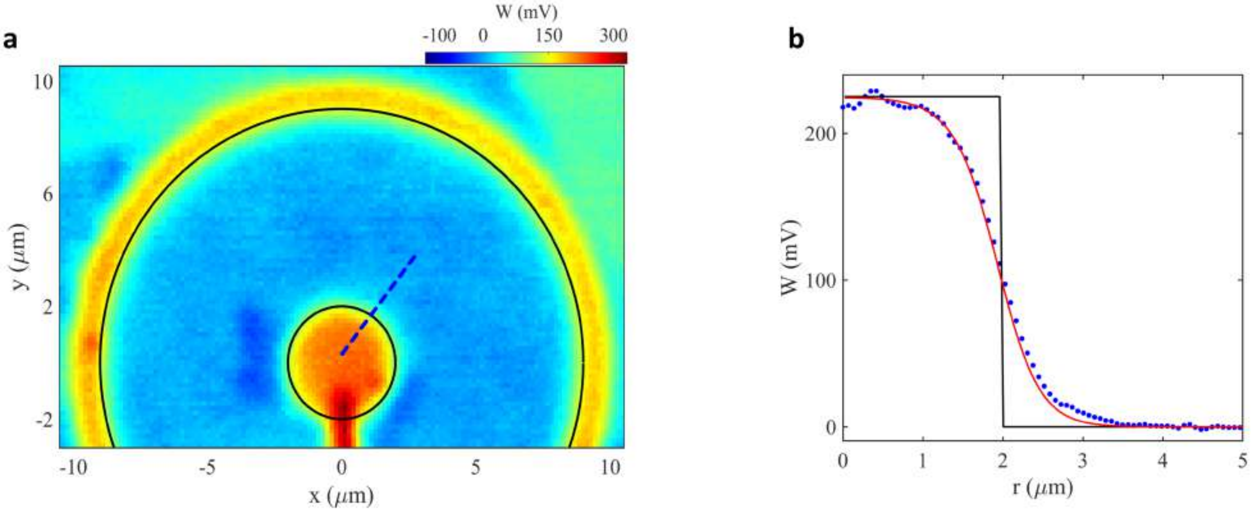


Fig. S4: **Obtaining the measurement's PSF from imaging of the workfunction. a.** Colormap of the measured workfunction, $W(x, y)$, over the Corbino disk. The solid black lines indicate the inner and outer radius of the graphene disk. **b.** Measured W (blue dots) along the dashed blue line in panel a. The black line shows a step function at the radius of the inner contact. The red line shows this profile smeared with the PSF $g(r) = 1/\cosh^2\left(1.76\frac{r}{\sigma}\right)$, where $\sigma = 0.85\ \mu m$. The smeared step function agrees well with the measured data.

S5. Determining the contact transparency from the measured resistance profile

As discussed in the main text, the two-probe resistance of a Corbino disk in the ballistic regime and with perfectly transmitting contacts, (transmission coefficient $T = 1$), is equal to the Sharvin resistance that corresponds to the radius of its inner contact, $R_{sh}^{in} = \frac{\pi h}{4e^2 K_F(2\pi r_{in})}$. Fig. 2b of the main text presents the radial dependence of resistance, $R(r)$, in the ballistic regime ($T = 6\ K$). Our measurement showed that the overall resistance (equivalent to two probe transport) is $19.5\ \Omega$, somewhat larger than $R_{sh}^{in} \sim\ 13.67\ \Omega$ at this density. We showed that a small part of this difference appears in the bulk and is due to a finite mean free path ($l_{MR} = 40\ \mu m$) that leads to a small ohmic bulk resistance. However, most of this difference happens at the contacts. For example, from the fitting in Fig. 2c we found that inner contact resistance step height is $0.82R_{sh}^{in}$, larger than the $0.5R_{sh}^{in}$ expected for an ideal contact. We will focus here only on the inner contact, because at low temperatures the physics there is simpler than at the outer contact (see section SI11 below). This increased contact resistance reflects a contact transmission that is smaller than one. Following Landauer, we know that a finite transmission leads to a resistance of $\frac{1-T}{T} R_{sh}^{in}$. Adding this to the half Sharvin resistance expected for an ideal contact, we get:

$$R_{contact} = \frac{R_{sh}^{in}}{2} + \frac{1-T}{T} R_{sh}^{in} = R_{sh}^{in} \frac{(2-T)}{2T} \tag{S5.1}$$

Rearranging the above expression, we get:

$$T = \frac{2}{2R_{contact}/R_{sh}^{in} + 1} \tag{S5.2}$$

Fig. S5 plots the contact transparency of the inner contact at $T = 6\ K$ as a function of carrier density, obtained by using equation S5.2 and $R_{contact}$ deduced from similar fits as in Fig. 2c, but for the resistance profiles measured at the different densities. We see that over the entire carrier density range the contact transparency is high, on par with the best transparencies achieved with graphene contacts[1,4].

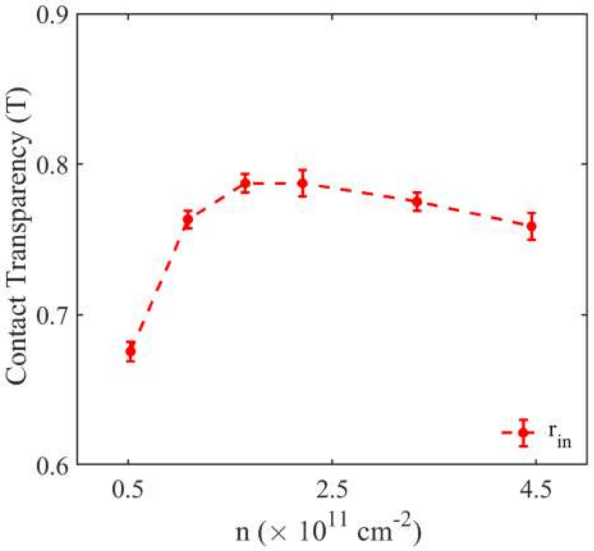


Fig. S5: **Density dependence of the inner contact transparency**. We fit $R(r)$ measured at $T = 6\ K$ and different carrier densities and obtain the height of the resistive step at the inner contact (see main text). Using equation S5.2 we obtain the corresponding contact transparency, presented in the figure as a function of the carrier density.

S6. Determining the momentum relaxing mean free path over the full temperature range.

As explained in the main text, when the temperature is high enough such that the electron flow is hydrodynamic, we can determine the momentum relaxing mean free path, l_{MR}, directly from measurements at finite magnetic fields (Fig. 4 in the main text). At low temperature, we can also determine l_{MR} from fitting the imaged resistance profile to Landauer-Sharvin + ohmic dependence + contact resistance, $R_{LS}(r) + R_{ohm}(r) + R_c^{in}(r) + R_c^{out}(r)$ (see Fig. 2c and corresponding text). In this section we use the theoretical expression for the temperature dependence of electron-phonon scattering to interpolate between these measurements and obtain the full temperature dependence of l_{MR}. This l_{MR} vs. T curve is then used together with Eq. (1) in the main text to obtain the traces in Fig. 3b.

In general, the momentum relaxing mean free path in graphene is determined by the scattering by disorder and phonons. We will term the former 'impurity scattering', although one has to keep in mind that disorder is not dominated few isolated impurities but is rather by a smooth potential modulation with long spatial scale caused by a distribution of many impurities spaced from the graphene by an hBN spacer. The corresponding disorder and electron-phono mean free paths are l_{imp} and l_{e-ph}. The total momentum relaxing mean free path is then given by the Matthiessen sum rule of these two processes:

$$l_{mr} = (l_{imp}^{-1} + l_{e-ph}^{-1})^{-1} \tag{6.1}$$

The density and temperature dependence of the resistance due to phonon scattering is given by[1,5]:

$$\rho_{e-ph}(n,T) = \frac{8D_A^2 k_F}{e^2 \rho_m v_s v_F^2} f_s\left(\frac{\theta_{BG}}{T}\right) \tag{S6.2}$$

where $k_F = \sqrt{\pi n}$ is the Fermi momentum, D_A is the acoustic deformation potential, ρ_m is the mass density of graphene, v_F is the Fermi velocity, v_s is the longitudinal acoustic phonon velocity and f_s is the Bloch Gruneisen function, given by $f_s(z) = \int_0^1 \frac{zx^4\sqrt{1-x^2}\, e^{zx}}{(e^{zx}-1)^2}\, dx$. The resistivity is translated to the e-ph mean free path using the standard relation valid for graphene:

$$l_{e-ph} = \frac{1}{\rho_{e-ph}(n,T)}\left(\frac{h}{2e^2 k_F}\right) \tag{S6.3}$$

We obtain the density dependence of the impurity mean free path at $T = 6\ K$ directly from our measurements (inset to Fig. 2d in the main text). These measurements are reproduced in blue dots in Fig. S6a together with a polynomial interpolation (blue line). At this temperature, the contribution of electron-phonon scattering is negligible, and thus this curve represents $l_{imp}(n)$. We further assume that the temperature dependence comes entirely from the temperature dependence of the e-ph scattering. If we add to $l_{imp}(n)$ the $l_{e-ph}(n,T)$ from equations (S6.1) - (S6.3) and use the parameters given in Ref[5] ($D_A = 21\ eV$, $\rho_m = 7.6e-7\ Kgm^{-2}$, $v_F = 1\times10^6\ ms^{-1}$ and $v_s = 2\times10^4\ ms^{-1}$) we obtain the total l_{MR} at elevated temperatures, shown by the different colored traces in Fig. S6a.

The red and orange dots in Fig. S6a correspond to the l_{mr} obtained from magneto-resistance measurements (as described in the main text) at a density of $n = 4.5\times10^{11}\ cm^{-2}$ and $T = 100\ \mathrm{K}$ (orange dot) and at the same density and $T = 140\ \mathrm{K}$ (red dot). We see that these measurements nicely fit the above expression. This can be also seen when we plot l_{MR} vs. temperature (Fig. S6b). The above expression (black line) fits well the three experimentally measured points. This suggest that we can use this expression for obtaining the l_{MR} at temperatures between our low and high temperatures data points.

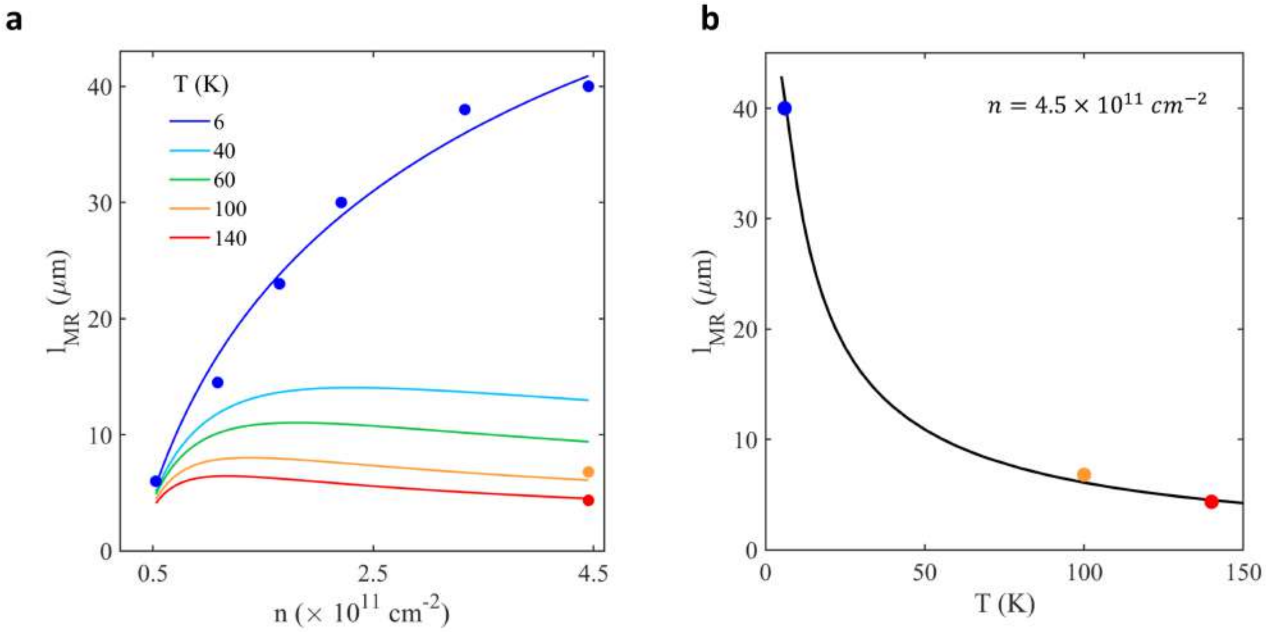


Fig. S6: l_{MR} vs. temperature and density. a. Measured mean free path as a function of carrier density at $T = 6\ \mathrm{K}$ (blue dots), $100\ \mathrm{K}$ (orange dot) and $140\ \mathrm{K}$ (red dot) (see text). The blue line is a polynomial fit through the data at 6 K. The lines corresponding to higher temperatures (see legend) include in addition the electron phonon scattering term (see text in this section). **b**. l_{MR} as a function of temperature at $n = 4.5\times10^{11} cm^{-2}$. The measured l_{MR} at $T = 6\ \mathrm{K}$ (blue dot) $100\ \mathrm{K}$ (orange dot) and $140\ \mathrm{K}$ (red dot) are in good agreement with the theoretical expression.

S7. Additional data at a different carrier density

In the main text, we present data at $V_{BG} = 2\ V$, corresponding to $n = 4.5 \times 10^{11}\ cm^{-2}$. Here, we present additional data from another carrier density ($V_{BG} = 1.5\ V,\ n = 3.3 \times 10^{11}\ cm^{-2}$), demonstrating similar behavior to that presented in the main text.

Fig. S7a shows the imaged $R(r)$ at $T = 6\ \mathrm{K}$ and $n = 3.3 \times 10^{11}\ cm^{-2}$. This curve has similar characteristics as that in Fig. 2b. In Fig. S7b we use a similar fit as in the main text. The dashed line shows a fit to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$, demonstrating that the measured bulk resistance is predominantly given by the Landauer-Sharvin resistance. Adding the ohmic term $R_{ohm}(r)\ = R_{sh}^{in} \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$ with $l_{mr} = 38\ \mu m$ we obtain the excellent fit shown the dashed red curve.

Fig. S7c shows the bulk resistance $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$, as a function of temperature. We see similar trend to the measurements in the main text: the total bulk resistance first decreased with increasing temperature up to $\sim T = 60\ \mathrm{K}$ and then goes slightly up. Moreover, the spatial dependence within the bulk shows a similar evolution to the one shown in the main text: from a curved resistance profile at low temperatures (when plotted on a logarithmic r axis), $\sim \mathrm{asin}(r_{in}/r)$, to a linear profile, namely $\sim \log(r/r_{in})$, at high temperatures.

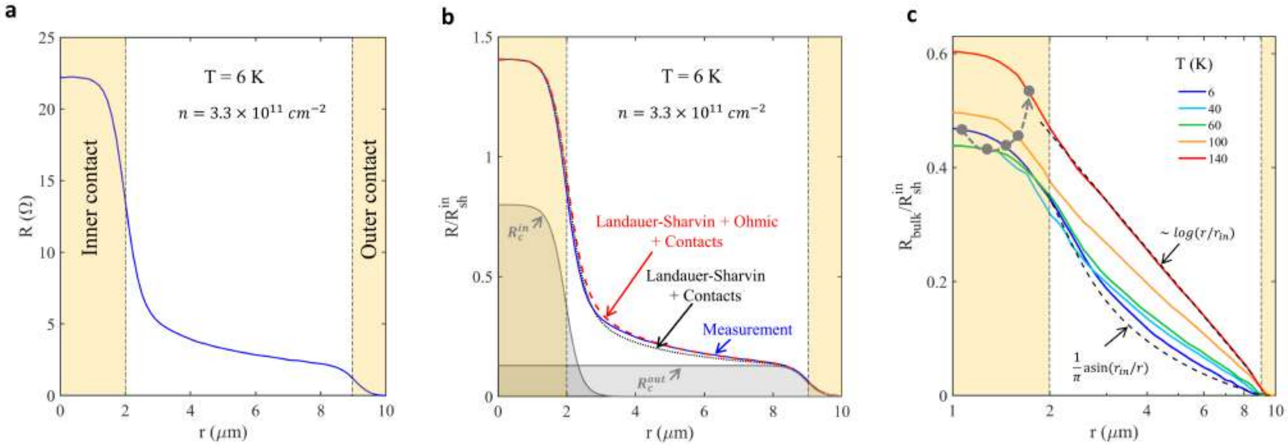


Fig. S7 | Measured resistance profiles at $n = 3.3 \times 10^{11}\ cm^{-2}$. a. Measured $R(r)$, at $T = 6\ K$. **b.** Breaking up the resistance to bulk and contact components. The PSF smeared contact resistances, $R_c^{in}(r)$ and $R_c^{out}(r)$ (grey), are as described in the main text. Dotted black line: Fit to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$, where , $R_{LS}(r)\ = R_{sh}^{in} \frac{1}{\pi} \mathrm{asin}\left(\frac{r_{in}}{r}\right)$ is the theoretically predicted Landauer-Sharvin bulk geometrical resistance. Red dashed line shows a fit to the same expression with the addition of $R_{ohm}(r)\ = R_{sh}^{in} \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$, with $l_{MR} = 38\ \mu m$. **c.** Measured $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$ at various temperatures, T, normalized by R_{sh}^{in} and plotted with a logarithmic r axis.

S8. Imaging measurements on a second Corbino device

We performed measurements on a second Corbino disk with different dimensions, $r_{in} = 1\ \mu m$ and $r_{out} = 6\ \mu m$. The inset in Fig S8a shows the optical image of this device, fabricated using similar procedure as discussed in SI section 1.

Fig. S8a presents the measured $R(r)$ at $T = 6\ \mathrm{K}$ and $n = 4.5 \times 10^{11}\ cm^{-2}$, where the transport is ballistic. This curve has the same characteristics to that in Fig. 2b. In Fig. S8b we use a similar fit as in the main text. The dashed line shows a fit to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$, demonstrating that the measured bulk resistance is predominantly given by the Landauer-Sharvin resistance. Adding the ohmic term $R_{ohm}(r) = R_{sh}^{in} \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$ with $l_{mr} = 27\ \mu m$ we obtain the excellent fit shown the dashed red curve.

Fig. S7c shows the bulk resistance $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$, as a function of temperature. We see similar trend to the measurements in the main text that: the resistance evolves from a curved resistance profile at low temperatures (when plotted on a logarithmic r axis), $\sim \mathrm{asin}(r_{in}/r)$, to a linear profile, namely $\sim \log(r/r_{in})$, at high temperatures.

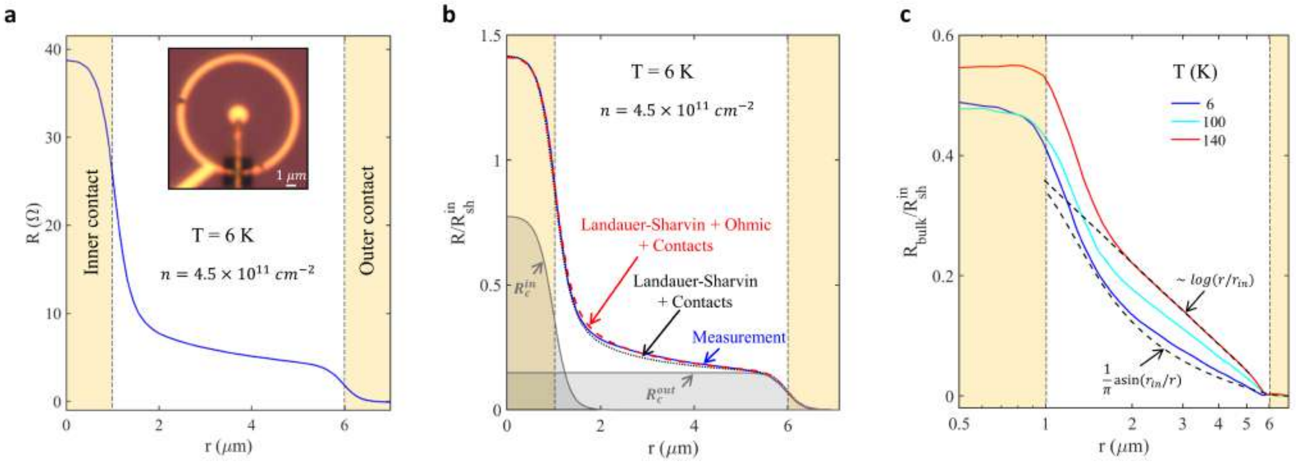


Fig. S8 | Imaging of a second Corbino device. **a**. Inset: optical image of the device. Main panel: measured $R(r)$, at $T = 6\,K$. **b.** Breaking up the resistance to bulk and contact components. The PSF smeared contact resistances, $R_c^{in}(r)$ and $R_c^{out}(r)$ (grey), are as described in the main text. Dotted black line: Fit to $R_{LS}(r) + R_c^{in}(r) + R_c^{out}(r)$, where , $R_{LS}(r) = R_{sh}^{in} \frac{1}{\pi} \mathrm{asin}\left(\frac{r_{in}}{r}\right)$ is the theoretically predicted Landauer-Sharvin bulk geometrical resistance. Red dashed line shows a fit to the same expression with the addition of $R_{ohm}(r) = R_{sh}^{in} \frac{2r_{in}}{\pi l_{MR}} \log\left(\frac{r}{r_{in}}\right)$, with $l_{MR} = 27\ \mu m$. **c.** Measured $R_{bulk}(r) = R(r) - \left(R_c^{in}(r) + R_c^{out}(r)\right)$ at various temperatures, T, normalized by R_{sh}^{in} and plotted with a logarithmic r axis.

S9. Derivation of Equation 1 in the main text

In this section we prove the identity:

$$R(r, L_{ee} = (l_{ee}^{-1} + l_{MR}^{-1})^{-1}, L_{MR} = \infty) = R(r, L_{ee} = lee, L_{MR} = l_{MR}) - \frac{\hbar}{2e^{2}k_{F}l_{MR}}\log(r/r_{in}) \tag{s9.1}$$

which is equation (1) in the main text, used to obtain the clean-limit hydrodynamic flow profile.

In the equation above $R(r, L_{ee}, L_{MR})$ is the bulk resistance profile with momentum-conserving and momentum-relaxing mean free paths of L_{ee} and L_{MR} correspondingly.

The full Boltzmann equation reads as:

$$\vec{v}.\nabla\chi + (l_{ee}^{-1} + l_{MR}^{-1})(\chi - \bar{\chi}) - l_{ee}^{-1}2\cos(\theta)\,\overline{\chi\cos(\theta)} = 0 \tag{S9.2}$$

where $\chi(r, \theta)$ is the Boltzmann function describing the distribution of the carrier's momenta direction, given by θ at location r, and $\bar{\chi}$ is its averaging over θ.

We rewrite equation (S9.2) as,

$$\vec{v}.\nabla\chi + (l_{ee}^{-1} + l_{MR}^{-1})(\chi - \bar{\chi}) - (l_{ee}^{-1} + l_{MR}^{-1})2\cos(\theta)\,\overline{\chi\cos(\theta)} + l_{MR}^{-1}2\cos(\theta)\,\overline{\chi\cos(\theta)} = 0 \tag{S9.3}$$

or,

$$L_{0}\chi + l_{MR}^{-1}2\cos(\theta)\,\overline{\chi\cos(\theta)} = 0 \tag{S9.4}$$

where L_0 is the Boltzmann equation for a modified problem with $L_{MR} = \infty$ and $L_{ee} = (l_{ee}^{-1} + l_{MR}^{-1})^{-1}$.

Denoting the solution of L_0as χ_0, we look for a solution of the form $\chi = \chi_0 + \delta\chi(r)$.

We find

$$L_{0}\delta\chi + l_{MR}^{-1}2\cos(\phi\theta)\,\overline{\chi\cos(\theta)} = 0$$

$$\cos(\theta)\,\partial_{r}\delta\chi + l_{MR}^{-1}2\cos(\theta)\,\overline{\chi\cos(\theta)} = 0$$

$$\partial_{r}\delta\chi + l_{MR}^{-1}j = 0 \tag{S9.5}$$

and thus

$$\delta\chi = -l_{MR}^{-1}\frac{I}{2\pi}\left(\log(r) - \log(r_{in})\right) \tag{S9.6}$$

S10. Boltzmann simulations of interacting flow in a Corbino geometry

To model electron flow through the graphene channels, we employ an approach based on the Boltzmann equation[6–9] that incorporates the effects of both electron-impurity and electron-phonon scattering as well as electron-electron interactions:

$$\partial_t f + \vec{v} \cdot \nabla_{\vec{r}} f = \frac{\partial f}{\partial t}|_{scatt}, \tag{S10.1}$$

where the scattering integral,

$$\frac{\partial f(\vec{r}, \vec{v})}{\partial t}|_{scatt} = -\frac{f(\vec{r}, \vec{v}) - n(\vec{r})}{\tau} + \frac{2}{\tau_{ee}} \vec{v} \cdot \vec{j}(\vec{r}), \tag{S10.2}$$

has two contributions: one from momentum-relaxing scattering, with a rate $\frac{1}{\tau_{MR}}$, and one from momentum-conserving, electron-electron scattering, with a rate $\frac{1}{\tau_{ee}}$. This equation describes the evolution of the semiclassical occupation number $f(\vec{r}, \vec{v})$ for a wave packet at position $\vec{r}$ and velocity $\vec{v}$, where $n(\vec{r}) = \langle f \rangle_{\vec{v}}$ is the local charge density, $\vec{j}(\vec{r}) = \langle f \vec{v} \rangle_{\vec{v}}$ the local current density, $\langle \ldots \rangle_{\vec{v}}$ is the momentum average, and $\frac{1}{\tau} = \frac{1}{\tau_{MR}} + \frac{1}{\tau_{ee}}$. For the sake of simplicity, we consider the case of a circular Fermi surface with $\vec{v} = v_F \hat{\rho}(\theta)$, where $\hat{\rho}$ is the radial unit vector at angle θ. Mean free paths are then simply defined as $l_{MR(ee)} = v_F \cdot \tau_{MR(ee)}$. The term proportional to τ_{ee}^{-1} is the simplest momentum-conserving scattering term that can be written, assuming that the electrons relax to a Fermi-Dirac distribution shifted by the drift velocity[6,10–12].

The sample is a Corbino disk with inner radius r_{in} and outer radius r_{out}. We use polar coordinates in real space as well, with radius r and angle ϕ. Thanks to a rotational symmetry, the ϕ variable drops out of the calculation. Combining everything, the Boltzmann equation takes the form:

$$cos(\theta)\partial_r f - sin(\theta)\frac{1}{r}\partial_\theta f = -\frac{f(\mathrm{r}, \theta) - n(\mathrm{r})}{l} + \frac{2}{l_{ee}} \vec{v} \cdot \vec{j}(\mathrm{r})$$

with $= v_F \tau$, $n(r) = \frac{1}{2\pi} \int d\theta\, f(r, \theta)$ and where $j_r(r) = \frac{1}{2\pi} \int d\theta\, f(r, \theta) \cos(\theta)$ is the radial current (the azimuthal current is zero by symmetry).

The rapid transition from a practically-infinite density of states in the metal contact, to the finite density of states at the graphene channel next to it, imposes the following boundary condition:

$$\begin{aligned} f(r = r_{min}, -\pi/2 \le \theta \le \pi/2) = f_{in} \\ f(r = r_{max}, \pi/2 \le \theta \le 3\pi/2) = 0, \end{aligned} \qquad \text{(S10.3)}$$

where f_{in} is a constant whose value is set to fix the total current.

The resulting integrodifferential equation is solved numerically using the method of characteristics[13] to invert the differential part of the equation, and an iterative method to solve the integral part.

Based on the solution for f, one finds the total current as $I = 2\pi j_r$ and the electrochemical potential as the electron density $n(r)$ divided by the density of states at the Fermi level.

Further, the contact resistance can be deduced by assuming the following form for f at the two contacts: $f(r = r_{min} - \varepsilon, \theta) = f_{in}$ at the inner contact, and $f(r = r_{out} + \varepsilon, \theta) = 0$ at the outer contact, where ε is infinitesimal.

S11. Temperature dependence of the outer contact resistance

The experiments described in Fig. 3b of the main text shows that as the temperature increases, a small step gradually builds up near the outer contact. We recall that in all the curves in that figure (corresponding to measurement temperatures from $T = 6\ K$ to $140\ K$) we subtracted the **same** contact step functions $R_c^{in}(r)$ and $R_c^{out}(r)$, those obtained from fitting the resistance profile at $T = 6\ K$ (Fig. 2c). This means that the contact steps that we see in Fig. 3b reflect the **difference** between the contact resistance at finite temperatures and that at $T = 6\ K$. The fact that we observe a finite outer contact step at high T therefore suggests that the outer contact resistance is larger at higher T.

In our Boltzmann numerical calculations (Fig. 3c) we observed a similar effect, but even stronger. There upon decreasing of l_{ee} (which corresponds to increasing T in the experiment) we see the buildup of a resistance step at the outer contact. We note that we use a similar procedure to the one used for the experimental curves - subtracting from all the Boltzmann curves the same contact resistance steps, the one obtained in the calculation with $l_{ee} = \infty$. The buildup of contact resistance step with decreasing l_{ee} in Fig. 3c therefore implies that the outer contact resistance increases with decreasing l_{ee}.

The Boltzmann calculation allows us to identify that this phenomenon originates from the transition between highly non-local ballistic flow at $l_{ee} = \infty$ and a locally equilibrated hydrodynamic flow at small l_{ee}. In a ballistic flow, the angular distribution of carriers in graphene, just outside the inner contact, has half of the angles populated by hot carriers emitted from the inner contact. Going toward the outer contact, the hot electrons get collimated to a smaller angular spread (the simple analogy would be the angle distribution of light rays reaching from the sun to an observer. As the observer gets further away from the sun, the distribution of light rays (/hot electrons) becomes more collimated). At a radius r the hot electrons are collimated to an angular spread of $\Delta\theta = 2\,\mathrm{asin}\left(\frac{r_{in}}{r}\right)$. As we have shown in the paper, the corresponding r dependence of the resistance is $R(r) = R_{sh}^{in}\frac{1}{\pi}\mathrm{asin}\left(\frac{r_{in}}{r}\right)$. For $r = r_{out} \gg r_{in}$ the collimated beam is very narrow, $\Delta\theta \approx 2\frac{r_{in}}{r_{out}}$. The outer contact resistance can be readily obtained from the value of the resistance function at r_{out}, since at the outer contact itself it equals zero. Namely, the theoretically predicted outer contact resistance in the ballistic regime is:

(ballistic)
$$R_{contact}^{out} = R_{sh}^{in}\frac{1}{\pi}\mathrm{asin}\left(\frac{r_{in}}{r_{out}}\right) \approx R_{sh}^{in}\frac{1}{\pi}\frac{r_{in}}{r_{out}} \quad \text{(S11.1)}$$

Indeed, this is the value that we obtain in the Boltzmann calculations in the ballistic regime ($l_{ee} = \infty$).

When we examine the Boltzmann calculations deep in the hydrodynamic case ($l_{ee} \ll r_{in}, r_{out}, r_{out} - r_{in}$) we see that we obtain a different outer contact resistance. In fact, in this case we observe that this resistance roughly equals the resistance one would obtain in the diffusive regime, namely, half the Sharvin resistance corresponding to r_{out}, $R_{sh}^{out} = \frac{\pi h}{4e^2} \frac{1}{k_F(2\pi r_{out})}$. Consequently,

$$\text{(hydrodynamic)} \qquad R_{contact}^{out} = \frac{1}{2} R_{sh}^{out} = R_{sh}^{in} \frac{1}{2} \frac{r_{in}}{r_{out}} \qquad \text{(S11.2)}$$

At the same time, in the hydrodynamic regime, the angular distribution of the electrons near the outer contact is $\sim\cos(\theta)$. Comparing equations (11.1) and (11.2) we see that $R_{contact}^{out}$ in the hydrodynamic regime is larger by a factor $\pi/2$ than that in the ballistic regime (we ignore here higher orders in l_{ee}). This corresponds to the increase of outer contact resistance with decreasing l_{ee} seen in Fig. 3c. This factor reflects the fact that the current density carried by a highly collimated angular distribution is larger by a factor of $\pi/2$ than the current density carried by a $\cos(\theta)$ distribution.

In the experiment we observe similar effect to that in the Boltzmann numerics, although the increase of the outer contact resistance with temperature is somewhat smaller. We note also that this feature is approaching the noise level of our experiment. It is likely that the reduced amplitude of the effect in the experiment, compared to the ideal theory, reflects the fact that the collimation effect described above is much more sensitive to experimental imperfections, such as roughness of the outer contact, more than the other effects reported in this paper that display very good agreement with the theory.

References:

1. Wang, L. *et al.* One-Dimensional Electrical Contact to a Two-Dimensional Material. *Science* **342**, 614–617 (2013).
2. Kim, Y., Herlinger, P., Taniguchi, T., Watanabe, K. & Smet, J. H. Reliable Postprocessing Improvement of van der Waals Heterostructures. *ACS Nano* **13**, 14182–14190 (2019).
3. Ella, L. *et al.* Simultaneous voltage and current density imaging of flowing electrons in two dimensions. *Nature Nanotechnology* **14**, 480–487 (2019).
4. ben Shalom, M. *et al.* Quantum oscillations of the critical current and high-field superconducting proximity in ballistic graphene. *Nature Physics* **12**, 318–322 (2016).
5. Efetov, D. K. & Kim, P. Controlling Electron-Phonon Interactions in Graphene at Ultrahigh Carrier Densities. *Physical Review Letters* **105**, 256805 (2010).
6. de Jong, M. J. M. & Molenkamp, L. W. Hydrodynamic electron flow in high-mobility wires. *Physical Review B* **51**, 389–402 (1995).
7. Ditlefsen, E. & Lothe, J. Theory of size effects in electrical conductivity. *Philosophical Magazine* **14**, 759–773 (1966).
8. Moll, P. J. W., Kushwaha, P., Nandi, N., Schmidt, B. & Mackenzie, A. P. Evidence for hydrodynamic electron flow in PdCoO 2. *Science* **351**, 1061–1064 (2016).
9. Scaffidi, T., Nandi, N., Schmidt, B., Mackenzie, A. P. & Moore, J. E. Hydrodynamic Electron Flow and Hall Viscosity. *Physical Review Letters* **118**, 226601 (2017).
10. Callaway, J. Model for lattice thermal conductivity at low temperatures. *Physical Review* **113**, 1046–1051 (1959).
11. Gennaro, S. de & Rettori, A. The low-temperature electrical resistivity of potassium: size effects and the role of normal electron-electron scattering. *Journal of Physics F: Metal Physics* **14**, L237–L242 (1984).
12. Gennaro, S. de & Rettori, A. On the electron-electron contribution to the low-temperature electrical resistivity of thin wires. *Journal of Physics F: Metal Physics* **15**, 2177–2188 (1985).
13. Courant, R. & Hilbert, D. Methods of Mathematical Physics. *Methods of Mathematical Physics* **2**, 1–830 (2008).

arXiv:1912.12341v3 [cond-mat.mes-hall] 11 Feb 2020

Giant nonlocality in nearly compensated 2D semimetals

S. Danz,[1] M. Titov,[2, 3] and B.N. Narozhny[1, 4]
[1]*Institut für Theorie der kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[2]*Radboud University Nijmegen, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*
[3]*ITMO University, 197101 St. Petersburg, Russia*
[4]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*

(Dated: February 12, 2020)

In compensated two-component systems in confined, two-dimensional geometries, nonlocal response may appear due to external magnetic field. Within a phenomenological two-fluid framework, we demonstrate the evolution of charge flow profiles and the emergence of a giant nonlocal pattern dominating charge transport in magnetic field. Applying our approach to the specific case of intrinsic graphene, we suggest a simple physical explanation for the experimental observation of giant nonlocality. Our results provide an intuitive way to predict the outcome of future experiments exploring the rich physics of many-body electron systems in confined geometries as well as to design possible applications.

The trend towards miniaturization of electronic devices requires a deeper understanding of the electron flow in confined geometries. In contrast to the electric current in household wiring, charge flow in small chips with multiple leads may exhibit complex spatial distribution patterns depending on the external bias, electrostatic environment, chip geometry, and magnetic field. Traditionally, such patterns were detected using nonlocal transport measurements [1–7], i.e. by measuring voltage drops between various leads other than the source and drain. Devised to study ballistic propagation of charge carriers in mesoscopic systems, these techniques were recently applied to investigate possible hydrodynamic behavior in ultra-pure conductors [8–12], where the unusual behavior of the nonlocal resistance is often associated with viscosity of the electronic system [13–17].

Nonlocal resistance measurements have also been used to study edge states accompanying the quantum Hall effect [18–23]. While the exact nature of the edge states has been a subject of an intense debate, the nonlocal resistance, R_{NL}, appears to be an intuitively clear consequence of the fact that the electric current flows along the sample edges and not through the bulk. Such a current would not be subject to exponential decay [24] exhibited by the bulk charge propagation leading to a much stronger nonlocal resistance.

In recent years the focus of the experimental work on electronic transport has been gradually shifting towards measurements at nearly room temperatures [6, 8–10, 21]. A particularly detailed analysis of the nonlocal resistance in a wide range of temperatures, carrier densities, and magnetic fields was performed on graphene samples [21]. Remarkably, the nonlocal resistance measured at charge neutrality remained strong well beyond the quantum Hall regime, with the peak value $R_{NL} \approx 1.5\,\mathrm{k}\Omega$ at $B = 12\,\mathrm{T}$ and $T = 300\,\mathrm{K}$, three times higher than that at $T = 10\,\mathrm{K}$.

In this Letter, we argue that the giant nonlocality observed in intrinsic graphene at high temperatures can be attributed to the presence of two types of charge carriers (electrons and holes): at the neutrality point, the two bands (the conductance and valence bands) touch creating a two-component electronic system. Physics of such systems is much richer than in their single-component counterparts. Observed phenomena that are directly related to the two-band structure of the neutrality point include giant magnetodrag in graphene [25, 26] and linear magnetoresistance [27, 28]. Both effects have been explained within a phenomenological framework [26, 27] allowing for a two-component (electron-hole) system coupled by the external magnetic field. We generalize this approach to investigate evolution of the spatial distribution of the electron current density in the experimentally relevant Hall bar geometry. In sufficiently strong magnetic fields, the current density forms a giant nonlocal pattern where the current is flowing not only in the bulk, but also along the boundaries leading to strong nonlocal resistance, see Fig. 1. Such patterns can be directly observed in laboratory experiments using the modern imaging techniques [29–31]. Tuning the model parameters to the specific values available for graphene, we arrive at a quantitative estimate of the nonlocal resistance [21].

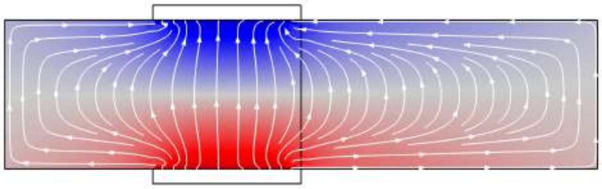


FIG. 1: Giant nonlocality in a compensated semimetal in magnetic field. The arrows indicate the current flow and the color map shows the electrochemical potential (see the main text and Figs. 2 and 3 for specific parameters).

To highlight the difference between the one- and two-component systems, we briefly recall the macroscopic description of electronic transport in the standard (former) case. Allowing for nonuniform charge density, the linear

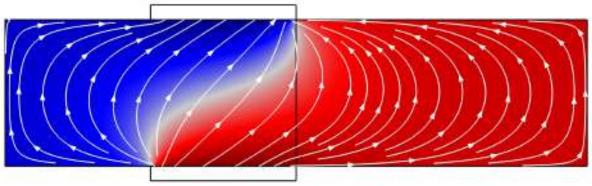


FIG. 2: Classical Hall effect in a one-component electronic system. The current density (shown by the arrows) and the electrochemical potential (shown by the color map) were obtained from Eqs. (1) for a sample of the width $W = 1\,\mu$m and length $L = 4\,\mu$m with the carrier density $n = 10^{12}\,\text{cm}^{-2}$ at the temperature $T = 240\,$K and in magnetic field $B = 0.2\,$T.

relation between the electric current $\boldsymbol{J}$ and the external fields $\boldsymbol{E}$, $\boldsymbol{B}$ could be formulated as [17, 32, 33]

$$r_0 \boldsymbol{J} = \boldsymbol{E} + r_H \boldsymbol{e_B} \times \boldsymbol{J} + \frac{1}{e\nu_0} \boldsymbol{\nabla} n, \tag{1a}$$

where $e > 0$ is the unit charge, ν_0 is the density of states (DoS), n is the carrier density, $\boldsymbol{e_B}$ is the unit vector in the direction of the magnetic field, and r_0 and r_H are the longitudinal and Hall resistivities. Within the Drude-like description, $r_H = \omega_c \tau r_0$ (ω_c is the cyclotron frequency and τ is the mean free path). The relation Eq. (1a) is applicable to a wide range of electronic systems from simple metals [34, 35] to doped graphene [11, 36]. The transport coefficients r_0 and r_H could be treated as phenomenological or could be derived from the underlying kinetic theory [11, 32, 37].

In addition to Eq.(1a), the electric current satisfies the continuity equation, which for stationary currents reads

$$\boldsymbol{\nabla} \cdot \boldsymbol{J} = 0. \tag{1b}$$

Charge density inhomogeneity induces electric field, so that Eq. (1a) should be combined with the corresponding electrostatic problem. Most recent experiments were performed in gated structures, where the relation between the electric field and charge density simplifies [27, 38]. In two-dimensional (2D) samples

$$\boldsymbol{E} = \boldsymbol{E}_0 - \frac{e}{C} \boldsymbol{\nabla} n, \tag{1c}$$

where $C = \epsilon/(4\pi d)$ is the gate-to-sample capacitance per unit area, d is the distance to the gate, ϵ is the dielectric constant, and $\boldsymbol{E}_0$ is the external field.

In a two-terminal (slab) geometry, solution of Eqs. (1) is a textbook problem. In the absence of magnetic field, the resulting electrochemical potential is governed by the relation of the mean free path to the system size, exhibiting either a flat (in short, ballistic samples) or linear (in long, diffusive samples) spatial profile. Most recently, these solutions were used as benchmarks in the imaging experiment [29] and the numerical solution of the hydrodynamic equations in doped graphene [17]. In external magnetic field, the system exhibits the classical Hall effect, which in short samples is accompanied by nontrivial current flow patterns [39].

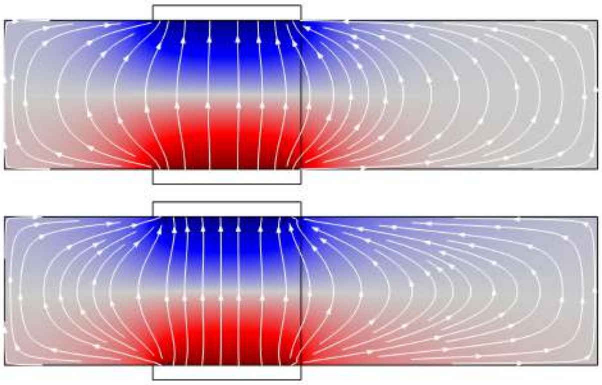


FIG. 3: Charge flow in compensated semimetals. Top: Ohmic flow in the absence of magnetic field. Bottom: emergent nonlocality in weak magnetic field $B = 0.2\,$T. The associated potential on the sample boundaries grows with the increasing field, see Fig. 1 for the pattern at $B = 2\,$T. Stronger fields expel the current from the bulk such that it flows along the boundary.

In a four-terminal Hall bar geometry, the electric current still fills the whole sample, but decays exponentially [24] away from the direct path between source and drain. The resulting flow pattern was calculated (in the context of doped graphene) in Refs. [14, 15, 17]. In magnetic field, the pattern gets skewed due to the classical Hall effect, but exhibits no qualitatively new features, see Fig. 2.

Let us now extend the transport equations (1) to a two-component system. Keeping in mind applications to graphene, we re-write Eq. (1a) for the quasiparticles in the conduction band ("electrons") in the form

$$-\boldsymbol{j}_e = eD\nu_e \boldsymbol{E} + \omega_c \tau \boldsymbol{j}_e \times \boldsymbol{e_B} + D \boldsymbol{\nabla} n_e, \tag{2a}$$

where $\boldsymbol{j}_e$ is the electron flow density (carrying the electric current $\boldsymbol{J}_e = -e\boldsymbol{j}_e$) and ν_e is DoS. The "holes" (i.e., the quasiparticles in the valence band) are described by

$$-\boldsymbol{j}_h = -eD\nu_h \boldsymbol{E} - \omega_c \tau \boldsymbol{j}_h \times \boldsymbol{e_B} + D \boldsymbol{\nabla} n_h. \tag{2b}$$

Here the electric current carried by the holes is $\boldsymbol{J}_h = e\boldsymbol{j}_h$ and DoS may differ from that of electrons, $\nu_h \neq \nu_e$. For simplicity, we assume that the the cyclotron frequency, mean free time, and diffusion constant for the two bands coincide (a generalization is straightforward, but doesn't lead to qualitatively new physics).

The total electric current in the two component system is given by $\boldsymbol{J} = -e\boldsymbol{j}$, where $\boldsymbol{j} = \boldsymbol{j}_e - \boldsymbol{j}_h$. Introducing also the total quasiparticle flow $\boldsymbol{j}_I = \boldsymbol{j}_e + \boldsymbol{j}_h$, we find (cf. Ref. [37])

$$\boldsymbol{j} + eD(\nu_e + \nu_h)\boldsymbol{E} + \omega_c \tau \boldsymbol{j}_I \times \boldsymbol{e_B} + D \boldsymbol{\nabla} n = 0, \tag{3a}$$

$$\boldsymbol{j}_I + eD(\nu_e - \nu_h)\boldsymbol{E} + \omega_c \tau \boldsymbol{j} \times \boldsymbol{e_B} + D \boldsymbol{\nabla} \rho = 0, \tag{3b}$$

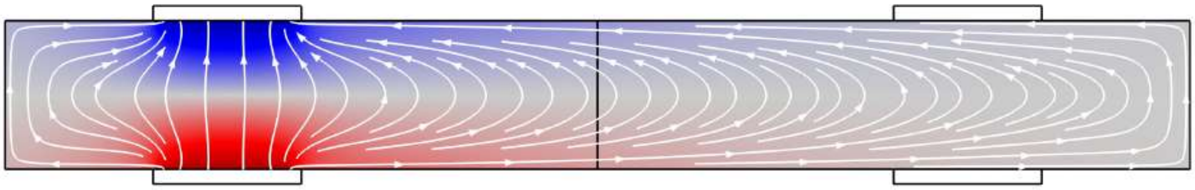


FIG. 4: Giant nonlocality in the Hall bar geometry. The sample has a width $W = 1\,\mu$m and length $8\,\mu$m, with the distance between contacts $L = 5\,\mu$m. The driving current is $I = 0.1\,\mu$A. The flow pattern was computed for $B = 0.8\,$T, cf. Fig. 1.

where $n = n_e - n_h$ is the carrier density per unit charge (the charge density being $-en$) and $\rho = n_e + n_h$ is the total quasiparticle density. The transport equations have to be supplemented by continuity equations reflecting the particle number conservation. The electric current satisfies Eq. (1b), but the total number of quasiparticles [40] can be affected by electron-hole recombination processes leading to a weak decay term in the continuity equation

$$\boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\delta\rho/\tau_R, \tag{3c}$$

where $\delta\rho$ is the deviation of the quasiparticle density from its equilibrium value and τ_R is the recombination time.

Under the assumption of electron-hole symmetry (e.g., at the charge neutrality point in graphene), $\nu_e = \nu_h$, we recover the phenomenological model of Ref. [27]. In the two-terminal geometry this model yields unsaturating linear magnetoresistance in classically strong fields [28].

Now we analyze the behavior of the phenomenological model (3) in the four-terminal Hall bar geometry. In the absence of the magnetic field, the system exhibits a typical Ohmic flow [14, 15, 17], see the top panel in Fig. 3. Applying the field we find a qualitative change in the flow pattern – the emergence of a boundary flow and the associated electrochemical potential at the sample edges. Increasing the field leads to the nonlocal pattern growing until it fills the whole sample, see Figs. 1 and 4. Stronger fields essentially expel the current from the bulk with the charge flow being concentrated along the sample boundaries, which leads to strong nonlocal resistance.

The nonlocal flow pattern emerging in magnetic field, see Figs. 1, 3 and 4, has to be contrasted with the vortices appearing in the viscous hydrodynamic flow (e.g., in doped graphene [14, 15, 17, 41]). In the latter case, vorticity appears due to the constrained geometry of the flow and the particular boundary conditions [15, 17, 42]: neglecting Ohmic effects, the solution of the hydrodynamic equations can be obtained by introducing the stream function, which obeys a biharmonic equation independent of viscosity, which however affects the distribution of the electrochemical potential. In contrast, within the model (3) the "Ohmic" scattering represents the only source of dissipation and hence cannot be omitted. One can still introduce the stream function, but now it is determined not only by the sample geometry, but also by the Ohmic scattering and magnetic field. As a result, the flow pattern does not exhibit vortices, unlike those suggested recently for the hydrodynamic flow in intrinsic graphene [41] (in the absence of magnetic field).

Nonlocal resistance in graphene subjected to external magnetic field was studied experimentally in Ref. [21]. At high enough temperatures where signatures of the quantum Hall effect are washed out, strong (or "giant") nonlocality was observed at the neutrality point. The effect vanishes in zero field as well as with doping away from neutrality. Both features are consistent with the model (3): in zero field the model exhibits usual Ohmic flow patterns, see Fig. 3, while at sufficiently high doping levels the effects of the second band are suppressed – the two equations (3a) and (3b) become identical showing the response typical of one-component systems, see Fig. 2.

Having discussed the qualitative features of the charge flow in two-component systems, we now turn to a quantitative calculation of nonlocal resistance in graphene. Although the model (3) is applicable to any semimetal, graphene is a by far better studied material with readily available experimental values for model parameters. Here we use the data measured in Refs. [8, 9, 21, 26, 43] and theoretical calculations of Refs. [11, 12, 26, 37, 41].

DoS of the quasiparticles in graphene has been evaluated in, e.g., Refs. [11, 12, 36, 37], and has the form

$$\nu_e + \nu_h = 2\mathcal{T}/(\pi v_g^2), \qquad \nu_e - \nu_h = 2\mu/(\pi v_g^2), \tag{4}$$

where μ is the chemical potential, v_g is the quasiparticle velocity in graphene, and $\mathcal{T} = 2T\ln[2\cosh(\mu/2T)]$. The generalized cyclotron frequency is $\omega_c = eBv_g^2/(c\mathcal{T})$ and the diffusion coefficient has the usual form $D = v_g^2\tau/2$. At charge neutrality, $\mu = 0$ and $\mathcal{T} = 2T\ln 2$, while in the degenerate regime $\mathcal{T}(\mu \gg T) = \mu$. The latter confirms that all coefficients in Eqs. (3a) and (3b) become identical with doping. Similarly, the continuity equations (1b) and (3c) should coincide in the degenerate regime. In graphene this happens by means of the fast decay of the recombination rate [26]. Close to neutrality we assume

$$\tau_R^{-1} = g^2T/\cosh(\mu/T), \tag{5}$$

where g is determined by the corresponding matrix element. The above expression [26] reflects the exponential

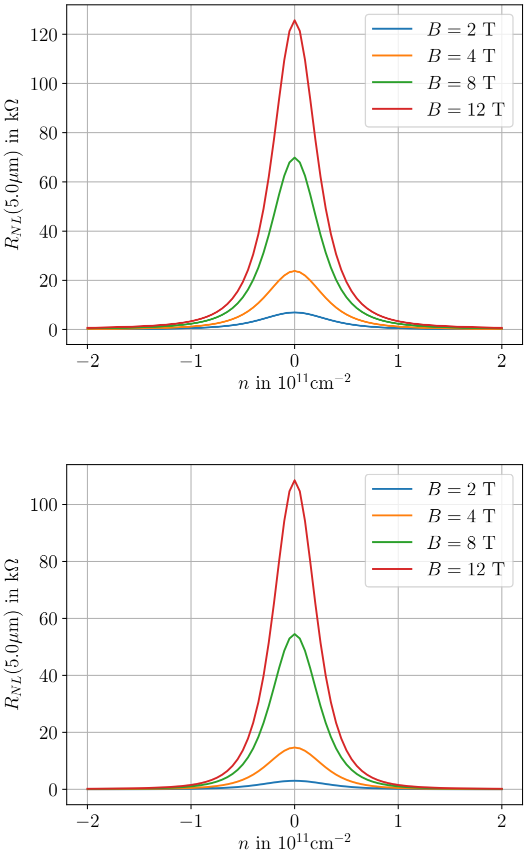


FIG. 5: Nonlocal resistance measured in the Hall bar geometry, see Fig. 4, as a function of carrier density. Top: Coulomb scatterers; bottom: short-ranged impurities. The impurity model parameters are chosen to represent the mobility at $n = 10^{11}$ cm^{-2} reported in Ref. [21]. The range of magnetic fields and carrier densities as well as the distance to the gate ($d = 50$ nm) is taken from Ref. [21], see Fig.2.

decay of the two-band physics away from charge neutrality, which is responsible for the fast decay of R_{NL} as a function of carrier density [21], see Fig. 5. Finally, the mean-free time, τ, in graphene is a non-trivial function of temperature and carrier density [11, 12, 36, 43, 44], which strongly depends on the model of the impurity potential [45–50]. However, these dependencies are typically not exponential and hence do not affect the exponential decay of the nonlocal resistance.

In Fig. 5 we demonstrate the decay of R_{NL} for two impurity models – the Coulomb scatterers and short-ranged impurities – showing nearly identical behavior. Such robustness of the model (3) with respect of the functional dependence of the mean free time justifies the inaccuracy of our description of electronic transport in graphene, where close to charge neutrality the resistivity is strongly affected by electron-electron interaction. The data shown in Fig. 5 were obtained by solving Eqs. (3) in the Hall bar geometry of Fig. 4 using the estimate [41] for the recombination length scale, $\ell_R = v_g\tau_R \approx 10\,\mu$m (a previous calculation of Ref. [26] put it at a smaller value $1.2\,\mu$m), which leads to similar results for the nonlocal resistance, but with a smaller peak value at charge neutrality.

The results for R_{NL} shown in Fig. 5 are extremely similar to those reported in Ref. [21] with the exception of the values at neutrality, which are grossly exaggerated. There are several reasons for this behavior. Firstly, by ignoring the effects of electron-electron interaction, we strongly underestimate the usual resistivity of intrinsic graphene. Secondly, we ignore viscous effects. Furthermore, DoS in real graphene never really vanishes “at neutrality” due to electrostatic potential fluctuations [51]. As a result, the minimal carrier concentration is often as high as 10^{10}cm^{-2}, essentially cutting off the lower density range around the peak in Fig. 5. Finally, Eq. (5) is a rather crude estimate that needs to be improved.

To conclude, we have argued that the observed giant nonlocality in neutral graphene in non-quantizing magnetic fields at relatively high temperatures observed in Ref. [21] is a direct consequence of the two-band nature of the quasiparticle spectrum in graphene. As such, this effect is not specific to graphene and should be observable in any compensated two-component system. Our theory does not involve spin-related phenomena including the effect of Zeeman splitting invoked in Ref. [21]. The latter should be independent of the field direction, however, the effect was not observed in the nearly parallel field studied in Ref. [51]. Assuming the g-factor to be equal to 2, we estimate the Zeeman splitting as $E_z \approx 0.35$ meV≈ 4 K at $B = 10$ T. The corresponding residual quasiparticle density (at $T = 0$) is given by $\rho_Q = E_z^2/(4\pi v_g^2) \approx 2.2 \times 10^6$ cm^{-2}. As a result, we expect the effects of Zeeman splitting to be observable at temperatures and carrier densities much lower than those typical to nonlocal measurements discussed here.

With material-specific parameters, our phenomenological model is capable of a quantitative description of the effect. For graphene, a more precise calculation involving solution of the full system of hydrodynamic equations near charge neutrality is required to reach perfect agreement with the data, however the present approach shows that the effect is more general and does not require additional assumptions of electronic hydrodynamics.

The authors are grateful to I.V. Gornyi, A.D. Mirlin, J. Schmalian, J.A. Sulpizio, M. Schütt, A. Shnirman, and Y. Tserkovnyak for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 program (Project 873028 HYDROTRONICS), and by the Russian Science Foundation Project No. 17-12-0 (MT). BNN acknowledges the support by the MEPhI Academic Excellence Project, Contract No. 02.a03.21.0005.

---

[1] W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, Phys. Rev. Lett. **58**, 2347 (1987).

[2] H. van Houten, C. W. J. Beenakker, J. G. Williamson, M. E. I. Broekaart, P. H. M. van Loosdrecht, B. J. van Wees, J. E. Mooij, C. T. Foxon, and J. J. Harris, Phys. Rev. B **39**, 8556 (1989).

[3] A. K. Geim, P. C. Main, P. H. Beton, P. Streda, L. Eaves, C. D. W. Wilkinson, and S. P. Beaumont, Phys. Rev. Lett. **67**, 3014 (1991).

[4] K. L. Shepard, M. L. Roukes, and B. P. Van der Gaag, Phys. Rev. Lett. **68**, 2660 (1992).

[5] Y. Hirayama, A. D. Wieck, T. Bever, K. von Klitzing, and K. Ploog, Phys. Rev. B **46**, 4035 (1992).

[6] G. Mihajlović, J. E. Pearson, M. A. Garcia, S. D. Bader, and A. Hoffmann, Phys. Rev. Lett. **103**, 166601 (2009).

[7] R. V. Gorbachev, J. C. W. Song, G. L. Yu, A. V. Kretinin, F. Withers, Y. Cao, A. Mishchenko, I. V. Grigorieva, K. S. Novoselov, L. S. Levitov, et al., Science **346**, 448 (2014).

[8] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[9] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

[10] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. K. Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[11] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[12] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).

[13] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys. Rev. B **92**, 165433 (2015).

[14] L. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).

[15] G. Falkovich and L. Levitov, Phys. Rev. Lett. **119**, 066601 (2017).

[16] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[17] S. Danz and B. N. Narozhny (2019), arXiv:1910.14473.

[18] P. L. McEuen, A. Szafer, C. A. Richter, B. W. Alphenaar, J. K. Jain, A. D. Stone, R. G. Wheeler, and R. N. Sacks, Phys. Rev. Lett. **64**, 2062 (1990).

[19] J. K. Wang and V. J. Goldman, Phys. Rev. B **45**, 13479 (1992).

[20] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[21] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, A. S. Mayorov, M. I. Katsnelson, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levitov, et al., Science **332**, 328 (2011).

[22] X.-P. Zhang, C. Huang, and M. A. Cazalilla, 2D Materials **4**, 024007 (2017).

[23] K. Komatsu, Y. Morita, E. Watanabe, D. Tsuya, K. Watanabe, T. Taniguchi, and S. Moriyama, Science Advances **4**, eaaq0194 (2018).

[24] L. J. van der Pauw, Philips Tech. Rev. **20**, 223 (1958).

[25] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, et al., Nat. Phys. **8**, 896 (2012).

[26] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[27] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[28] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).

[29] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, et al., Nat. Nanotechnol. **14**, 480 (2019).

[30] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, et al. (2019), arXiv:1905.10791.

[31] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[32] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).

[33] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).

[34] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).

[35] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[36] M. I. Katsnelson, *Graphene* (Cambridge University Press, 2012).

[37] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[38] I. L. Aleiner and B. I. Shklovskii, Phys. Rev. B **49**, 13721 (1994).

[39] A. Shik, J. Phys. Condens. Matter **5**, 8963 (1993).

[40] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[41] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[42] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).

[43] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[44] B. N. Narozhny, M. Titov, I. V. Gornyi, and P. M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).

[45] T. Ando, J. Phys. Soc. Jpn. **75**, 074716 (2006).

[46] N. Shon and T. Ando, J. Phys. Soc. Jpn. **67**, 2421 (1998).

[47] K. Nomura and A. H. MacDonald, Phys. Rev. Lett. **96**, 256602 (2006).

[48] V. V. Cheianov and V. I. Fal'ko, Phys. Rev. Lett. **97**, 226801 (2006).

[49] I. L. Aleiner and K. B. Efetov, Phys. Rev. Lett. **97**, 236801 (2006).

[50] P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **74**, 235443 (2006).

[51] F. Chiappini, S. Wiedmann, M. Titov, A. K. Geim, R. V. Gorbachev, E. Khestanova, A. Mishchenko, K. S. Novoselov, J. C. Maan, and U. Zeitler, Phys. Rev. B **94**, 085302 (2016).

Hydrodynamic collective modes in graphene

B.N. Narozhny,[1, 2] I.V. Gornyi,[1, 3, 4] and M. Titov[5]
[1] *Institut for Theoretical Condensed Matter Physics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[2] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[3] *Institut for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[4] *Ioffe Institute, 194021 St. Petersburg, Russia*
[5] *Radboud University Nijmegen, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*

(Dated: March 22, 2021)

Collective behavior is one of the most intriguing aspects of the hydrodynamic approach to electronic transport. Here we provide a consistent, unified calculation of the dispersion relations of the hydrodynamic collective modes in graphene. Taking into account viscous effects, we show that the hydrodynamic sound mode in graphene becomes overdamped at sufficiently large momentum scales. Extending the linearized theory beyond the hydrodynamic regime, we connect the diffusive hydrodynamic charge density fluctuations with plasmons.

Electronic hydrodynamics is quickly growing into a mature field of solid state physics [1–17]. Similarly to the usual hydrodynamics [18], this approach offers a universal, long-wavelength description of collective flows in interacting many-electron systems. Such flows have been experimentally confirmed [6] to be more efficient than the usual single-electron (ballistic or diffusive) transport.

In graphene, hydrodynamic collective modes have been considered by many authors [2, 15, 19–26]. All of them agree that at charge neutrality, the ideal electronic fluid (i.e., neglecting all dissipative processes) allows for a sound-like collective mode (which has been referred to as either the "cosmic sound" [20] or the "second sound" [25]) with the dispersion relation

$$\omega = v_g q/\sqrt{2}, \tag{1}$$

where v_g is the quasiparticle velocity in graphene. Taking into account dissipation changes the above dispersion relation giving rise to damping. To the best of our knowledge, no consensus on the latter effect has been reached so far with several contradicting results available in the literature [15, 23].

The hydrodynamic approach to electronic systems is applicable in an intermediate parameter regime [1, 2]. In particular, the underlying gradient expansion is valid at length scales much larger than the typical length scale ℓ_{ee} describing the energy- and momentum-conserving interaction (responsible for equilibration of the system). At smaller length scales, one can study more traditional collective excitations in interacting many-electron systems, including plasmons [15, 21, 24–42], which behavior is well established both theoretically and experimentally.

In this paper we provide a consistent, unified calculation of the dispersion relations of the hydrodynamic collective modes in graphene. While the true hydrodynamics is universal (as long as no symmetries are broken), graphene is somewhat unique in the sense that there are two length scales associated with electron-electron interaction that are parametrically different in the weak coupling limit [2, 25, 43, 44]. This allows us to extend the results of the linearized hydrodynamic theory [15, 43, 45] to the length scales smaller that ℓ_{ee} (going beyond the small-momentum expansion of Ref. 15). At that point the sound mode (1) in neutral graphene (see Fig. 1) becomes overdamped due to the high viscosity [3, 9, 46]

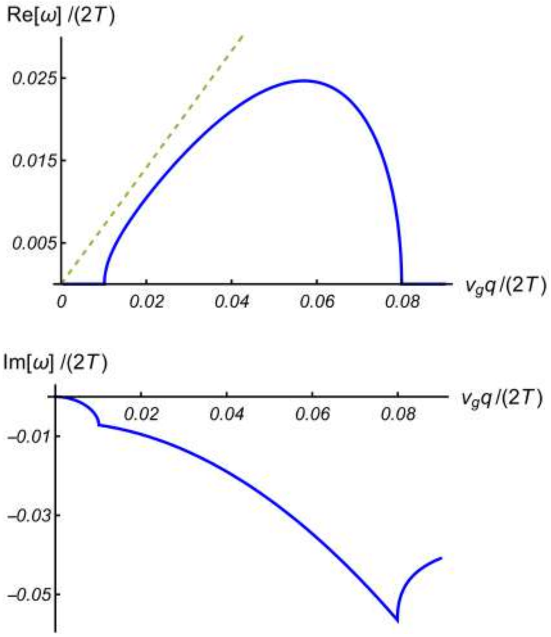


FIG. 1. Real (top) and imaginary (bottom) parts of the hydrodynamic sound dispersion in neutral graphene taking into account viscosity and weak disorder, Eq. (2). The numerical values were computed with the realistic parameter values taken from Refs. 3, 9, and 12; see the main text. The dispersion acquires a finite real part at the threshold value of momentum determined by dissipation. The mode becomes overdamped at small enough momenta, still in the region of the growing real part. The dashed line shows the ideal dispersion, Eq. (1).

$$\omega = \sqrt{\frac{v_g^2 q^2}{2} - \frac{\left(1+q^2\ell_G^2\right)^2}{4\tau_{\mathrm{dis}}^2}} - i\,\frac{1+q^2\ell_G^2}{2\tau_{\mathrm{dis}}}, \tag{2}$$

where τ_{dis} is the disorder mean free time and ℓ_G is the so-called Gurzhi length [47–51] (here ν stands for the kinematic viscosity [2, 3, 9, 46])

$$\ell_G = \sqrt{\nu\tau_{\rm dis}}. \tag{3}$$

This mode describes energy fluctuations and is completely decoupled from charge fluctuations. The latter are purely diffusive within the hydrodynamic approach, where dissipation is described by the momentum- and frequency-independent coefficients, including the electrical conductivity and viscosity.

Extending the linearized theory beyond the hydrodynamic regime, we are able to connect the charge fluctuations with the more conventional plasmons by taking into account the frequency and momentum dependence of conductivity. At charge neutrality we find the plasmon mode

$$\omega = \sqrt{\frac{v_g^2\varkappa q}{2}\left(1+\frac{q}{\varkappa}\right) - \frac{v_g^4\varkappa^2}{64\pi^2\sigma_0^2}} - i\frac{v_g^2\varkappa}{8\pi\sigma_0}, \tag{4}$$

where σ_0 is the conductivity in neutral graphene [1, 2, 16, 52]

$$\sigma_0 = \frac{2e^2T\ln 2}{\pi}\frac{\tau_{11}\tau_{\rm dis}}{\tau_{11}+\tau_{\rm dis}}, \tag{5}$$

and $\varkappa$ is the inverse Thomas-Fermi screening length. Neglecting dissipation and for small momenta, the dispersion (4) coincides with the result of Ref. 19.

Finally, we extend our results over the whole range of carrier densities up to the degenerate ("Fermi-liquid") regime. Given the weak density dependence of the kinematic viscosity in graphene [3, 9, 46] the sound dispersion remains qualitatively similar to that shown in Fig. 1 at all doping levels.

I. HYDRODYNAMIC THEORY OF ELECTRONIC TRANSPORT IN GRAPHENE

In this Section, we briefly review the hydrodynamic theory of electronic transport in graphene.

A. Nonlinear hydrodynamic equations

The complete set of hydrodynamic equations includes the generalized Navier-Stokes equation [16, 17]

$$W(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} + v_g^2\boldsymbol{\nabla}P + \boldsymbol{u}\partial_t P + e(\boldsymbol{E}\cdot\boldsymbol{j})\boldsymbol{u} = \tag{6a}$$

$$= v_g^2\left[\eta\Delta\boldsymbol{u} - \eta_H\Delta\boldsymbol{u}\times\boldsymbol{e}_B + en\boldsymbol{E} + \frac{e}{c}\boldsymbol{j}\times\boldsymbol{B}\right] - \frac{W\boldsymbol{u}}{\tau_{\rm dis}},$$

the continuity equations [1, 2, 16]

$$\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0, \tag{6b}$$

$$\partial_t n_I + \boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\frac{n_I - n_{I,0}}{\tau_R}, \tag{6c}$$

and the generalized "heat transport" equation [53–55] (we follow the usual approach [18] using the entropy flow equation instead of the continuity equation for energy).

$$T\left[\frac{\partial s}{\partial t} + \boldsymbol{\nabla}\cdot\frac{3P\boldsymbol{u} - \mu\boldsymbol{j} - \mu_I\boldsymbol{j}_I}{T}\right] = \tag{6d}$$

$$= \delta\boldsymbol{j}\cdot\left[e\boldsymbol{E} + \frac{e}{c}\boldsymbol{u}\times\boldsymbol{B} - T\boldsymbol{\nabla}\frac{\mu}{T}\right] - T\delta\boldsymbol{j}_I\cdot\boldsymbol{\nabla}\frac{\mu_I}{T}$$

$$+\frac{\eta}{2}\left(\nabla_\alpha u_\beta + \nabla_\beta u_\alpha - \delta_{\alpha\beta}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right)^2$$

$$-\frac{n_E - n_{E,0}}{\tau_{RE}} + \mu_I\frac{n_I - n_{I,0}}{\tau_R} + \frac{W\boldsymbol{u}^2}{v_g^2\tau_{\rm dis}}.$$

Here $\boldsymbol{u}$ is the hydrodynamic velocity, c is the speed of light, and n and n_I are the carrier and imbalance densities ($n_{I,0}$ is the equilibrium value), related to the quasiparticle densities in each of the two bands by

$$n = n_+ - n_-, \qquad n_I = n_+ + n_-.$$

The carrier density n differs from the charge density by a multiplicative factor of the electric charge, e. Similarly, we define the two quasiparticle currents, $\boldsymbol{j}$ and $\boldsymbol{j}_I$,

$$\boldsymbol{j} = \boldsymbol{j}_+ - \boldsymbol{j}_-, \qquad \boldsymbol{j}_I = \boldsymbol{j}_+ + \boldsymbol{j}_-,$$

with the electric current $\boldsymbol{J} = e\boldsymbol{j}$. We also define the two chemical potentials, μ and μ_I,

$$\mu = (\mu_+ + \mu_-)/2, \qquad \mu_I = (\mu_+ - \mu_-)/2,$$

allowing for the two independent chemical potentials for each band out of equilibrium [53] (hence the term "imbalance"). The remaining vector quantities in Eqs. (6) are the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$. The thermodynamic quantities are the enthalpy density W, pressure P, entropy density s, and temperature T. Finally, η and η_H are the shear and Hall viscosities, τ_R is the recombination time [53] [the recombination term in Eq. (6c) agrees with Ref. 54, whereas Ref. 53 suggests a slightly different term that is proportional to μ_I instead of the δn_I], and τ_{RE} is the energy relaxation time [55]. In equilibrium, $\mu_I = 0$.

In comparison to the usual hydrodynamics [18], the electronic system in graphene is characterized by one additional variable describing the second band. Traditional ideal fluid is described by two thermodynamic variables, e.g., density and pressure, and the velocity field. As a result, in two dimensions one needs four equations to describe the dynamics of the flow. Two of these are given by the Euler equation, the third is the continuity equation, while the fourth can be either the continuity equation for energy or the adiabaticity equation (i.e., the continuity equation for entropy). In graphene these are Eqs. (6a), (6b), and (6d) in the absence of dissipation. The additional continuity equation (6c) for the quasiparticle density n_I appears exactly due to the presence of the second band, which is why the overall number of hydrodynamic equations as well as independent variables in graphene is five. As the additional variable one can choose either n_I or the corresponding chemical potential μ_I.

The entropy flow equation (6d) should be compared to the corresponding equations in Refs. 2, 53, and 54. The four equations contain mostly the same terms (up to trivial notation changes) with the following exceptions. Equation (54) of Ref. 2 is written in the relativistic notation omitting the imbalance mode, quasiparticle recombination, and disorder scattering, all of which are discussed separately elsewhere in Ref. 2. Reference 53 was the first to focus on the imbalance mode with Eq. (2.6) containing all the terms of Eq. (6d) except for the viscous term. Finally, Eq. (1c) of Ref. 54 contains all of the terms in Eq. (6d) and in addition contains a term describing energy relaxation due to electron-phonon scattering that is neglected in this paper (generalization of the resulting theory is straightforward).

Weak disorder scattering is described in Eqs. (6a) and (6d) by the mean free time $\tau_{\rm dis}$. The disorder contribution to the hydrodynamic equations was derived in Ref. 16 using the simplest τ-approximation to the kinetic equation. A better version of the disorder collision integral in graphene should involve the Dirac factors suppressing backscattering [56] which would lead to the similar approximation but with the transport scattering time. In graphene, this brings about a factor of 2. In this paper, we treat $\tau_{\rm dis}$ as a phenomenological parameter adopting the approach of Ref. 12.

The imbalance density n_I appears under the assumption of the approximate conservation of the number of particles in each individual band. The processes that break this conservation (i.e., mix electrons and holes) involve the three-particle scattering, Auger processes [53], and most importantly, impurity assisted electron-phonon coupling [57]. These effects are described in Eq. (6c) by the phenomenological [58, 59] recombination time [60], τ_R, as well as the energy relaxation time τ_{RE} in Eq. (6d).

B. Dissipative corrections to quasiparticle currents

The usual hydrodynamic flow [18] is a mass flow where dissipative processes lead to a correction to the energy flux as described by the thermal conductivity. Consequently the flow is characterized by three dissipative coefficients, the thermal conductivity $\varkappa$ and two viscosities η and ζ. In contrast, electronic hydrodynamics in graphene describes an energy flow where the quasiparticle currents acquire dissipative corrections. The energy flow is proportional to the momentum density and hence can only be affected by disorder, which is "extrinsic" to the hydrodynamic theory. As a result, the dissipative coefficients include the electrical conductivity σ and viscosity, while the thermal conductivity has to be computed by solving the linear response equations (similarly to the electrical conductivity in the standard theory). Within the three-mode approximation of Ref. 16, the bulk viscosity vanishes, $\zeta=0$. In the absence of the magnetic field the dissipative corrections are related to external bias by means of a "conductivity matrix" [16, 53, 54]

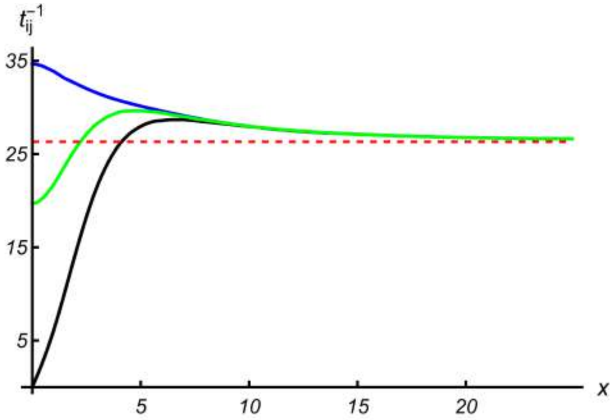


FIG. 2. Dimensionless scattering rates comprising the matrix $\widehat{\mathfrak{T}}$: t_{11}^{-1}, t_{12}^{-1}, t_{22}^{-1} (blue, black, and green, respectively). The red dashed line indicates the "Fermi-liquid" limit, Eq. (11).

$$\begin{pmatrix}\delta\boldsymbol{j}\\ \delta\boldsymbol{j}_I\end{pmatrix}=\widehat{\Sigma}\begin{pmatrix}e\boldsymbol{E}-T\boldsymbol{\nabla}(\mu/T)\\ -T\boldsymbol{\nabla}(\mu_I/T)\end{pmatrix}. \tag{7}$$

In particular, at the Dirac point $\mu=\mu_I=0$ the matrix $\widehat{\Sigma}$ is diagonal with the upper diagonal element defining (in the absence of disorder) the "quantum" or "intrinsic" conductivity [1, 2, 16, 53, 54]

$$\sigma_Q=e^2\Sigma_{11}(0). \tag{8}$$

In the hydrodynamic theory of graphene, the elements of the matrix $\widehat{\Sigma}$ play the role that is equivalent to that of the thermal conductivity $\varkappa$ in the usual hydrodynamics. The matrix nature of $\widehat{\Sigma}$ reflects the band structure of graphene. In the case of strong recombination, the imbalance mode becomes irrelevant and one is left with the single dissipative coefficient σ_Q, see Ref. 2.

1. *Macroscopic currents within the three-mode approximation*

Within the three-mode approximation of Ref. 16, one defines three macroscopic currents (using $\bar{W}=3\bar{n}_E/2$)

$$\boldsymbol{j}=\bar{n}\boldsymbol{u}+\delta\boldsymbol{j},\quad \boldsymbol{j}_I=\bar{n}_I\boldsymbol{u}+\delta\boldsymbol{j}_I,\quad \boldsymbol{j}_E=\frac{3}{2}\bar{n}_E\boldsymbol{u}, \tag{9}$$

where $\bar{n}$, $\bar{n}_I$, and $\bar{n}_E$ are the equilibrium values of the carrier, imbalance, and energy densities, respectively. The linear response theory relates the dissipative corrections $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ to the external bias by Eq. (7). The dimensionless conductivity matrix (at $\boldsymbol{B}=0$) is given by [16]

$$\widehat{\Sigma}=\widehat{\mathfrak{M}}\widehat{\mathfrak{S}}_{xx}^{-1}\widehat{\mathfrak{M}},\quad \widehat{\mathfrak{S}}_{xx}=\frac{\alpha_g^2T^2}{2\mathcal{T}^2}\widehat{\mathfrak{T}}+\frac{\pi}{\mathcal{T}\tau_{\rm dis}}\widehat{\mathfrak{M}}, \tag{10a}$$

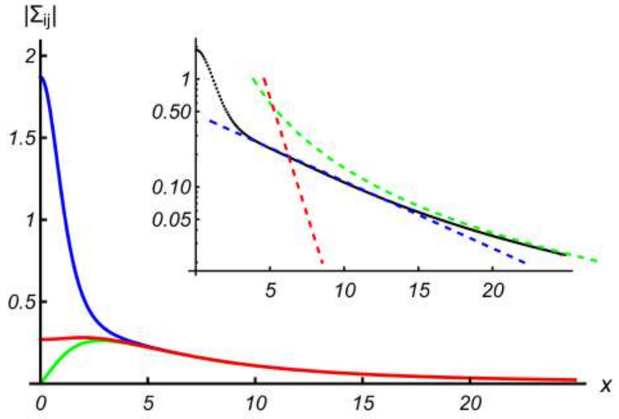


FIG. 3. Matrix elements of $\widehat{\Sigma}$. The blue, red, and green curves correspond to Σ_{11}, $|\Sigma_{12}|$, $|\Sigma_{22}|$, respectively (notice, that $\Sigma_{21}=-\Sigma_{12}$). The inset shows the log plot of Σ_{11}, where the red and blue lines indicate the exponential decay, while the green line is the power law $\sim x^{-2}$.

where

$$\widehat{\mathfrak{M}}=\begin{pmatrix} 1-\frac{2\tilde{n}^2}{3\tilde{n}_E}\frac{T}{\mathcal{T}} & \frac{xT}{\mathcal{T}}-\frac{2\tilde{n}\tilde{n}_I T}{3\tilde{n}_E \mathcal{T}} \\ \frac{xT}{\mathcal{T}}-\frac{2\tilde{n}\tilde{n}_I}{3\tilde{n}_E}\frac{T}{\mathcal{T}} & 1-\frac{2\tilde{n}_I^2}{3\tilde{n}_E}\frac{T}{\mathcal{T}} \end{pmatrix}, \tag{10b}$$

with dimensionless densities [see Eq. (15a) below]

$$\tilde{n}=\mathrm{Li}_2\left(-e^{-x}\right)-\mathrm{Li}_2\left(-e^{x}\right), \quad \tilde{n}_I=x^2/2+\pi^2/6,$$

$$\tilde{n}_E=-\mathrm{Li}_3\left(-e^{x}\right)-\mathrm{Li}_3\left(-e^{-x}\right), \tag{10c}$$

$$x=\mu/T, \quad \mathcal{T}=2T\ln\left[2\cosh(x/2)\right],$$

and dimensionless scattering rates

$$\widehat{\mathfrak{T}}=\begin{pmatrix} t_{11}^{-1} & t_{12}^{-1} \\ t_{12}^{-1} & t_{22}^{-1} \end{pmatrix}, \qquad t_{ij}^{-1}=\frac{8\pi\mathcal{T}}{\alpha_g^2 N T^2}\tau_{ij}^{-1}, \tag{10d}$$

where τ_{ij}^{-1} are the scattering rates that can be obtained by solving the kinetic equation within the three-mode approximation [15, 16, 25, 44]. The zeros in the matrix (10d) are the manifestation of energy and momentum conservation, which is also responsible for the vanishing dissipative correction to the energy current in the absence of the magnetic field [16]. The three dimensionless elements of the matrix $\widehat{\mathfrak{T}}$ are shown in Fig. 2 as a function of $x=\mu/T$.

The resulting matrix elements of the conductivity $\widehat{\Sigma}$ are shown in Fig. 3 as functions of $x=\mu/T$. As discussed below, the numerical precision of the present calculation is insufficient to track the exponential corrections to the scattering rates in the degenerate regime. Hence, the decay shown in the inset in Fig. 3 might be an artifact.

2. *Dimensionless scattering rates*

In the degenerate regime all scattering rates (i.e., the matrix elements t_{ij}^{-1}) coincide (up to exponentially small corrections) approaching the limiting value

$$t_{ij}^{-1}(\mu\gg T)\to 8\pi^2/3. \tag{11}$$

At $\mu=0$, the off-diagonal elements $t_{12}^{-1}(0)=0$, while the diagonal elements $t_{ii}^{-1}(0)$ determine the diagonal elements of the conductivity matrix, σ_Q and σ_I, see below. For small $x\ll 1$ the dimensionless "scattering rates" t_{ij} have the form [45] (see Fig. 4 for illustration)

$$\frac{1}{t_{11}}=\frac{1}{t_{11}^{(0)}}+x^2\left(\frac{1}{t_{11}^{(2)}}-\frac{1}{8\ln 2}\frac{1}{t_{11}^{(0)}}\right)+\mathcal{O}(x^3), \tag{12a}$$

$$\frac{1}{t_{12}}=\frac{x}{t_{12}^{(1)}}+\mathcal{O}(x^3), \tag{12b}$$

$$\frac{1}{t_{22}}=\frac{1}{t_{22}^{(0)}}+x^2\left(\frac{1}{t_{22}^{(2)}}-\frac{1}{8\ln 2}\frac{1}{t_{22}^{(0)}}\right)+\mathcal{O}(x^3). \tag{12c}$$

For unscreened Coulomb interaction, the dimensionless quantities $t_{ij}^{(0,1,2)}$ are just numbers without any dependence on any physical parameter. Numerically, one finds the following values (neglecting the small [52] exchange contribution):

$$\left(t_{11}^{(0)}\right)^{-1}\approx 34.63, \quad \left(t_{11}^{(2)}\right)^{-1}\approx 5.45,$$

$$\left(t_{12}^{(1)}\right)^{-1}\approx 5.72, \quad \left(t_{22}^{(0)}\right)^{-1}\approx 19.73, \quad \left(t_{22}^{(2)}\right)^{-1}\approx 5.65.$$

Note that these values are slightly different from those listed in Ref. 45. The reason for this is the use of different numerical methods. In the case of screened interaction, the quantities $t_{ij}^{(0,1,2)}$ depend on the screening length.

3. *Conductivity matrix close to charge neutrality*

Close to charge neutrality we expand the matrix $\widehat{\mathfrak{M}}$

$$\widehat{\mathfrak{M}}=\widehat{\mathfrak{M}}(0)+\delta\widehat{\mathfrak{M}}+\mathcal{O}(x^3),$$

with

$$\widehat{\mathfrak{M}}(0)=\begin{pmatrix} 1 & 0 \\ 0 & \delta_I \end{pmatrix}, \tag{13}$$

where $\zeta(z)$ is the Riemann's zeta function and

$$\delta_I=1-\frac{\pi^4}{162\zeta(3)\ln 2}\approx 0.28.$$

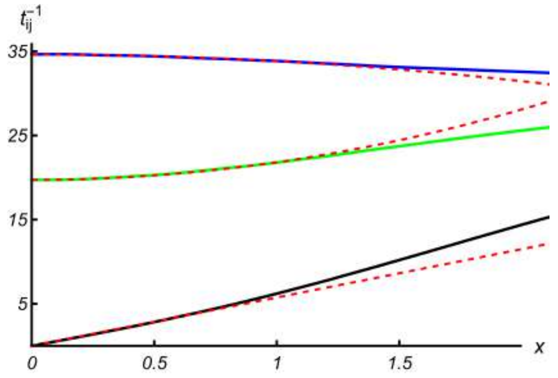


FIG. 4. Dimensionless scattering rates close to charge neutrality. The blue, black, and green curves correspond to t_{11}^{-1}, t_{12}^{-1}, t_{22}^{-1}, respectively. The red dashed lines indicate the leading behavior close to charge neutrality (11).

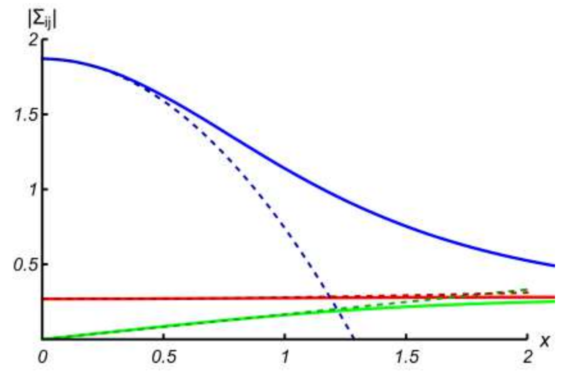


FIG. 5. Matrix elements of the dimensionless conductivity $\widehat{\Sigma}$ for small $x = \mu/T$. The blue, green, and red curves correspond to Σ_{11}, Σ_{21}, $|\Sigma_{22}|$, respectively. The dashed lines indicate the leading behavior close to charge neutrality.

The leading-order correction is given by

$$\delta\widehat{\mathfrak{M}} = \frac{x}{54\zeta(3)\ln 2} \times \begin{pmatrix} -16x\ln^2 2 & 27\zeta(3)-4\pi^2\ln 2 \\ 27\zeta(3)-4\pi^2\ln 2 & 2\pi^2 x\left[\frac{\pi^2}{48\ln 2}+\frac{\pi^2\ln 2}{9\zeta(3)}-1\right] \end{pmatrix}.$$

The matrix $\widehat{\mathfrak{S}}_{xx}$ can be expanded in the same way, using the expansion of the scattering rates (12):

$$\widehat{\mathfrak{S}}_{xx} = \widehat{\mathfrak{S}}_{xx}(0) + \delta\widehat{\mathfrak{S}}_{xx} + \mathcal{O}(x^3),$$

where

$$\widehat{\mathfrak{S}}_{xx}(0) = \frac{\pi}{2T\ln 2}\left[\begin{pmatrix} \tau_{11}^{-1} & 0 \\ 0 & \tau_{22}^{-1} \end{pmatrix} + \frac{1}{\tau_{\text{dis}}}\widehat{\mathfrak{M}}\right], \tag{14}$$

and

$$\delta\widehat{\mathfrak{S}}_{xx} = \frac{\alpha_g^2}{8\ln^2 2}\delta\widehat{\mathfrak{T}} + \frac{\pi}{2T\ln 2}\tau_{\text{dis}}^{-1}\delta\widehat{\mathfrak{M}},$$

with

$$\delta\widehat{\mathfrak{T}} = x\begin{pmatrix} \frac{x}{t_{11}^{(2)}} - \frac{1}{8\ln 2}\frac{x}{t_{11}^{(0)}} & 1/t_{12}^{(1)} \\ 1/t_{12}^{(1)} & \frac{x}{t_{22}^{(2)}} - \frac{1}{8\ln 2}\frac{x}{t_{22}^{(0)}} \end{pmatrix}.$$

Combining the above matrices, one finds the leading corrections to the conductivity matrix in the vicinity of the Dirac point, see Fig. 5.

Equations (6) and (7) reviewed in this Section represent a close set of hydrodynamic equations describing the electronic flows in graphene in the intermediate ("hydrodynamic") temperature window [1, 2]. So far, these equations were mostly studied within linear response (nonlinear phenomena were discussed, e.g., in Ref. 15). The hydrodynamic collective modes are also obtained by linearizing the hydrodynamic equations.

II. LINEARIZED HYDRODYNAMIC THEORY AT $B = 0$

In this Section, we discuss the linearization of the hydrodynamic theory in graphene suitable for a discussion of the bulk collective modes in the absence of the magnetic field, which is the primary focus of this paper.

Within linear response one considers small deviations of hydrodynamic quantities from their equilibrium values. At equilibrium, the stationary fluid is characterized by vanishing macroscopic currents and homogeneous thermodynamic quantities. Equilibrium quantities are most conveniently expressed in terms of the equilibrium values of temperature and chemical potential:

$$\mu = \bar{\mu}, \qquad T = \bar{T}, \qquad \mu_I = 0, \qquad x = \bar{\mu}/\bar{T}, \tag{15a}$$

$$n = \bar{n} = \frac{N\bar{T}^2}{2\pi v_g^2}\tilde{n}, \qquad n_I = \bar{n}_I = \frac{N\bar{T}^2}{2\pi v_g^2}\tilde{n}_I,$$

$$P = \bar{P} = \frac{N\bar{T}^3}{2\pi v_g^2}\tilde{n}_E, \quad W = 3\bar{P}, \quad s = \frac{3\bar{P} - \bar{\mu}\bar{n}}{\bar{T}}.$$

Finally, the electric potential is homogeneous as well

$$\varphi = \bar{\varphi}, \quad \boldsymbol{E} = -\boldsymbol{\nabla}\bar{\varphi} = 0, \quad \boldsymbol{j} = \boldsymbol{j}_I = \boldsymbol{j}_E = 0. \tag{15b}$$

The values $\bar{\mu}$, $\bar{\varphi}$, and $\bar{T}$ are determined by the environment in which the system is placed or, in other words, by the boundary conditions.

Once the system is subjected to a weak external voltage and temperature gradient, the hydrodynamic velocity $\boldsymbol{u}$ acquires a nonzero value and thermodynamic quantities become inhomogeneous. To the lowest (linear) order, one introduces small inhomogeneous fluctuations of the equilibrium quantities (not all being independent)

$$\mu = \bar{\mu} + \delta\mu, \qquad T = \bar{T} + \delta T, \qquad \varphi = \bar{\varphi} + \delta\varphi, \tag{16a}$$

$$n = \bar{n} + \delta n, \qquad n_I = \bar{n}_I + \delta n_I, \qquad P = \bar{P} + \delta P, \tag{16b}$$

as well as small values for those quantities that vanish in equilibrium

$$\boldsymbol{u}, \qquad \mu_I. \tag{16c}$$

The macroscopic currents have the form (9). Within linear response, the nonequilibrium corrections (9) [in general given in Eq. (7)] may be expressed as

$$\begin{pmatrix} \delta\boldsymbol{j} \\ \delta\boldsymbol{j}_I \end{pmatrix} = \widehat{\Sigma} \begin{pmatrix} -e\boldsymbol{\nabla}\delta\zeta + x\boldsymbol{\nabla}\delta T \\ -\boldsymbol{\nabla}\mu_I \end{pmatrix}, \tag{16d}$$

where $\widehat{\Sigma}$ is evaluated at equilibrium and

$$\delta\zeta = \delta\varphi + \frac{1}{e}\delta\mu, \tag{16e}$$

is the electrochemical potential. Here we used the fact that μ_I and $\boldsymbol{\nabla}\delta T$ are both assumed to be small, so that their products, e.g., $\mu_I\boldsymbol{\nabla}\delta T$, have to be neglected.

The same corrections can be expressed in terms of the density fluctuations rather than the chemical potentials [15]

$$\begin{pmatrix} \delta\boldsymbol{j} \\ \delta\boldsymbol{j}_I \end{pmatrix} = \widehat{\Sigma} \begin{pmatrix} e\boldsymbol{E} \\ 0 \end{pmatrix} - \frac{\bar{T}^2}{\mathcal{T}}\widehat{\Sigma}' \begin{pmatrix} \boldsymbol{\nabla}\delta\tilde{n} - \frac{2\tilde{n}}{3\tilde{n}_E}\boldsymbol{\nabla}\delta\tilde{n}_E \\ \boldsymbol{\nabla}\delta\tilde{n}_I - \frac{2\tilde{n}_I}{3\tilde{n}_E}\boldsymbol{\nabla}\delta\tilde{n}_E \end{pmatrix}, \tag{16f}$$

with dimensionless fluctuations of the densities and pressure [cf. Eqs. (16a) and (16b)] defined as

$$\delta n = \frac{N\bar{T}^2}{2\pi v_g^2}\delta\tilde{n}, \quad \delta n_I = \frac{N\bar{T}^2}{2\pi v_g^2}\delta\tilde{n}_I, \quad \delta P = \frac{N\bar{T}^3}{2\pi v_g^2}\delta\tilde{n}_E, \tag{16g}$$

the quantity $\mathcal{T}$ is related to the equilibrium compressibility [15, 16, 43, 45]

$$\frac{\partial\bar{n}}{\partial\bar{\mu}} = \frac{N\mathcal{T}}{2\pi v_g^2}, \quad \mathcal{T} = 2\bar{T}\ln 2\cosh\frac{\bar{\mu}}{2\bar{T}}, \tag{16h}$$

and finally

$$\widehat{\Sigma}' = \widehat{\boldsymbol{\mathfrak{M}}}\widehat{\boldsymbol{\mathfrak{S}}}_{xx}^{-1}, \qquad \widehat{\Sigma} = \widehat{\Sigma}'\widehat{\boldsymbol{\mathfrak{M}}}. \tag{16i}$$

The expressions (16d) and (16f) are completely equivalent, however one has to be careful with the electric field. Indeed, electrical conductivity is typically measured as a response to the "total" electric field and not to the "external electric field." The total electric field includes the so-called Vlasov self-consistency [1, 2, 15, 16, 43] taking into account the electric field induced by the density fluctuations. The latter can be obtained using Poisson's equation

$$\boldsymbol{E}_V = -e\boldsymbol{\nabla}\int d^2r'\frac{\delta n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{17a}$$

This relation simplifies in gated structures, where [60, 61]

$$\boldsymbol{E}_V = -\frac{e}{C}\boldsymbol{\nabla}\delta n(\boldsymbol{r}). \tag{17b}$$

Here $C = \varepsilon/(4\pi d)$ is the gate-to-channel capacitance per unit area, d is the distance to the gate, and ε is the dielectric constant. This approximation neglects the long-ranged (dipole-type) part of the Coulomb interaction (screened by the gate) and is valid as long as the charge density $n(\boldsymbol{r})$ varies on length scales much longer than d.

Linearizing the hydrodynamic equations (6) we find

$$\frac{3\bar{P}}{v_g^2}\partial_t\boldsymbol{u} + \boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} + e\bar{n}\boldsymbol{E} - \frac{3\bar{P}\boldsymbol{u}}{v_g^2\tau_{\text{dis}}}, \tag{18a}$$

$$\partial_t\delta n + \bar{n}\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{18b}$$

$$\partial_t\delta n_I + \bar{n}_I\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -\delta n_I/\tau_R, \tag{18c}$$

$$2\partial_t\delta P + 3\bar{P}\boldsymbol{\nabla}\cdot\boldsymbol{u} = -2\delta P/\tau_{RE}. \tag{18d}$$

Notice that the linearized "thermal transport" equation (18d) is completely equivalent (within linear response) to the continuity equation for the energy flow; see Refs. 1, 2, 15, 16, and 43. The energy relaxation term in Eq. (18d) was derived in Ref. [55].

At this point one has to choose the set of independent variables. Based on the form of the linearized equations (18), one can choose δn, δn_I, and δP. Together with the two components of $\boldsymbol{u}$ one has five variables for five differential equations (18). This set was used in Ref. 43 to discuss collective modes in the electronic fluid.

An alternative choice based on the form of dissipative corrections (16d) may include $\delta\zeta$, μ_I, and δT. These variables were chosen in Ref. 53 for the discussion of the role of the imbalance mode in thermoelectric effects. Indeed, using the thermodynamic relation [2, 16, 18, 53]

$$dP = nd\mu + n_I d\mu_I + sdT, \tag{19}$$

in the linearized Navier-Stokes equation (18a), one finds

$$\frac{3\bar{P}}{v_g^2}\left(\partial_t + \tau_{\text{dis}}^{-1}\right)\boldsymbol{u} = \eta\Delta\boldsymbol{u} - e\bar{n}\boldsymbol{\nabla}\zeta - \bar{n}_I\boldsymbol{\nabla}\mu_I - \frac{3\bar{P} - \bar{n}\bar{\mu}}{\bar{T}}\boldsymbol{\nabla}\delta T, \tag{20}$$

where we combine the electric and chemical potential into the electrochemical potential (16e). Given that the densities and pressure are given by known functions of the chemical potentials and temperature, see Eqs. (15a), it's a matter of simple algebra to express the rest of Eqs. (18) in terms of $\delta\zeta$, μ_I, and δT.

While the choice of the thermodynamic variables is a matter of taste, there is an important distinction between static and dynamic response [2]. Static linear response equations contain only the electrochemical potential ζ. However, the dynamic part of Eq. (18d) contains the chemical potential only. Consequently, one has to be careful considering response functions that depend on time and spatial coordinates at the same time. In this case, an additional equation (17) describing Vlasov self-consistency has to be taken into account [15].

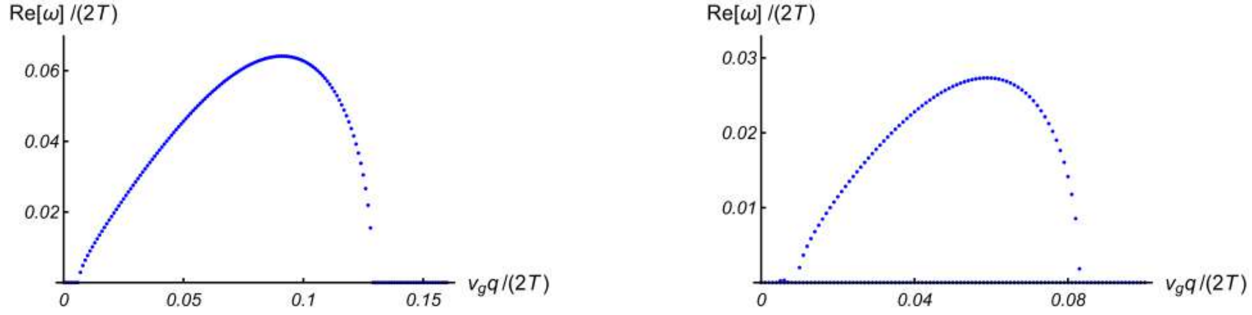


FIG. 6. Real part of the sound dispersion in moderately doped, gated graphene in the presence of both weak disorder and viscosity. Left pane: results for $n = 10^{12}\,\mathrm{cm}^{-2}$. Right panel: same for $n = 10^{11}\,\mathrm{cm}^{-2}$. The right panel also shows the zero mode Eq. (38).

III. COLLECTIVE MODES AT $\boldsymbol{B} = 0$

Collective modes in the electronic fluid were considered within the same approach in Ref. 15, see also Refs. 2, 44, and 62. These are the eigenmodes of the linearized equations (18). The most convenient choice of variables for this task is the density-pressure variables, δn, δn_I, and δP, and the velocity $\boldsymbol{u}$. The dissipative corrections to the currents are given by Eq. (16f) and the electric field in Eq. (18a) is the total electric field.

Now, it is convenient to solve linear differential equations with the help of the Fourier transform. Using the standard convention

$$\boldsymbol{u}(t,\boldsymbol{r}) = \int \frac{d\omega d\boldsymbol{q}}{(2\pi)^3} e^{-i\omega t + i\boldsymbol{q}\boldsymbol{r}} \boldsymbol{u}(\omega,\boldsymbol{q}),$$

we rewrite Eqs. (18) in the dimensionless form

$$\left(\tilde{\omega} + i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\mathrm{dis}}}\right)\tilde{n}_E \mathbf{v} - \frac{1}{3}\tilde{\boldsymbol{q}}\delta\tilde{n}_E - \tilde{\boldsymbol{q}}\tilde{V}_q\frac{e\tilde{n}}{6}\delta\tilde{n} = \frac{i}{6}e\tilde{n}\boldsymbol{\mathcal{E}}_0, \tag{21a}$$

$$\tilde{\omega}\delta\tilde{n} - \tilde{n}\tilde{\boldsymbol{q}}\cdot\mathbf{v} - \frac{2\pi}{N}\tilde{\boldsymbol{q}}\cdot\delta\tilde{\boldsymbol{j}} = 0, \tag{21b}$$

$$\left(\tilde{\omega} + \frac{i}{\tilde{\tau}_R}\right)\delta\tilde{n}_I - \tilde{n}_I\tilde{\boldsymbol{q}}\cdot\mathbf{v} - \frac{2\pi}{N}\tilde{\boldsymbol{q}}\cdot\delta\tilde{\boldsymbol{j}}_I = 0, \tag{21c}$$

$$\left(\tilde{\omega} + \frac{i}{\tilde{\tau}_{RE}}\right)\delta\tilde{n}_E - \frac{3}{2}\tilde{n}_E\tilde{\boldsymbol{q}}\cdot\mathbf{v} = 0, \tag{21d}$$

where

$$\tilde{\boldsymbol{q}} = v_g\boldsymbol{q}/(2\bar{T}), \quad \tilde{\omega} = \omega/(2\bar{T}), \quad \mathbf{v} = \boldsymbol{u}/v_g, \tag{22a}$$

$$\tilde{\tau}_j = 2\bar{T}\tau_j \ (j = \mathrm{dis},\, R,\, RE), \quad \boldsymbol{\mathcal{E}}_0 = v_g\boldsymbol{E}_0\bar{T}^{-2}, \tag{22b}$$

the dimensionless Gurzhi length is defined so that

$$\tilde{q}\tilde{\ell}_G = q\ell_G, \tag{22c}$$

and the self-consistent Vlasov potential is given by

$$\tilde{V}_q = \frac{eN\bar{T}}{\pi v_g^2}V_s, \qquad V_s(\boldsymbol{q}) = \begin{cases} e/C, & \text{gated}, \\ 2\pi e/q, & \text{Coulomb}, \end{cases} \tag{22d}$$

Finally, the dimensionless form of the dissipative corrections to the macroscopic currents is given by

$$\begin{pmatrix} \delta\tilde{\boldsymbol{j}} \\ \delta\tilde{\boldsymbol{j}}_I \end{pmatrix} = \widehat{\Sigma}\begin{pmatrix} e\boldsymbol{\mathcal{E}}_0 - i\tilde{\boldsymbol{q}}\tilde{V}_q\delta\tilde{n} \\ 0 \end{pmatrix} - \frac{i\tilde{\boldsymbol{q}}\,\widehat{\Sigma}'}{\ln 2\cosh\frac{x}{2}}\begin{pmatrix} \delta\tilde{n} - \frac{2\tilde{n}}{3\tilde{n}_E}\delta\tilde{n}_E \\ \delta\tilde{n}_I - \frac{2\tilde{n}_I}{3\tilde{n}_E}\delta\tilde{n}_E \end{pmatrix}. \tag{23}$$

The collective modes can now be found by analyzing the system of Eqs. (21). For convenience, it can be written in the matrix form

$$\begin{pmatrix} \tilde{\omega} + \frac{i2\pi\tilde{q}^2}{N}\left[\Sigma_{11}\tilde{V}_q + \frac{\Sigma'_{11}}{\ln 2\cosh\frac{x}{2}}\right] & \frac{i2\pi\tilde{q}^2\Sigma'_{12}}{N\ln 2\cosh\frac{x}{2}} & -\frac{i4\pi\tilde{q}^2}{3N}\frac{\Sigma'_{11}\tilde{n}+\Sigma'_{12}\tilde{n}_I}{\tilde{n}_E\ln 2\cosh\frac{x}{2}} & -\tilde{n}\tilde{\boldsymbol{q}} \\ \frac{i2\pi\tilde{q}^2}{N}\left[\Sigma_{21}\tilde{V}_q + \frac{\Sigma'_{21}}{\ln 2\cosh\frac{x}{2}}\right] & \tilde{\omega} + \frac{i}{\tilde{\tau}_R} + \frac{i2\pi\tilde{q}^2\Sigma'_{22}}{N\ln 2\cosh\frac{x}{2}} & -\frac{i4\pi\tilde{q}^2}{3N}\frac{\Sigma'_{21}\tilde{n}+\Sigma'_{22}\tilde{n}_I}{\tilde{n}_E\ln 2\cosh\frac{x}{2}} & -\tilde{n}_I\tilde{\boldsymbol{q}} \\ 0 & 0 & \tilde{\omega} + \frac{i}{\tilde{\tau}_{RE}} & -\frac{3}{2}\tilde{n}_E\tilde{\boldsymbol{q}} \\ -\tilde{\boldsymbol{q}}\tilde{V}_q\frac{\tilde{n}}{6} & 0 & -\frac{\tilde{\boldsymbol{q}}}{3} & \left[\tilde{\omega} + i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\mathrm{dis}}}\right]\tilde{n}_E \end{pmatrix}\begin{pmatrix} \delta\tilde{n} \\ \delta\tilde{n}_I \\ \delta\tilde{n}_E \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \frac{2\pi e}{N}\Sigma_{11}\tilde{q}\mathcal{E}_0 \\ \frac{2\pi e}{N}\Sigma_{21}\tilde{q}\mathcal{E}_0 \\ 0 \\ \frac{i}{6}e\tilde{n}\boldsymbol{\mathcal{E}}_0 \end{pmatrix}. \tag{24}$$

Dispersion relations of the collective modes are given by the zeros of the determinant of the matrix in the left-hand side of (24)

$$\begin{aligned}
&\left[\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right] \qquad (25)\\
&\times\Bigg\{\left[\tilde{\omega}+\frac{i}{\tilde{\tau}_R}+\frac{i2\pi\tilde{q}^2\Sigma'_{22}}{N\ln 2\cosh\frac{x}{2}}\right]\left[\tilde{\omega}+\frac{i2\pi\tilde{q}^2}{N}\left(\Sigma_{11}\tilde{V}_{\tilde{\boldsymbol{q}}}+\frac{\Sigma'_{11}}{\ln 2\cosh\frac{x}{2}}\right)\right]\left[\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)-\frac{\tilde{q}^2}{2}\right]\\
&\quad-\frac{2\pi\tilde{q}^2\Sigma'_{12}}{N\ln 2\cosh\frac{x}{2}}\Bigg[\frac{\tilde{q}^2\tilde{V}_{\tilde{\boldsymbol{q}}}\tilde{n}}{6\tilde{n}_E}\left(\tilde{n}\frac{2\pi\tilde{q}^2\Sigma'_{21}}{N\ln 2\cosh\frac{x}{2}}+\frac{\tilde{n}_I}{\tilde{\tau}_{RE}}-\frac{\tilde{n}_I}{\tilde{\tau}_R}\right)\\
&\qquad\qquad-\frac{2\pi\tilde{q}^2}{N}\left(\Sigma_{21}\tilde{V}_{\tilde{\boldsymbol{q}}}+\frac{\Sigma'_{21}}{\ln 2\cosh\frac{x}{2}}\right)\left[\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)-\frac{\tilde{q}^2}{2}\right]\Bigg]\\
&\quad-\frac{\tilde{q}^2\tilde{V}_q\tilde{n}^2}{6\tilde{n}_E}\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}+\frac{i2\pi\tilde{q}^2\Sigma'_{11}}{N\ln 2\cosh\frac{x}{2}}\right)\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_R}+\frac{i2\pi\tilde{q}^2\Sigma'_{22}}{N\ln 2\cosh\frac{x}{2}}\right)\Bigg\}=0.
\end{aligned}$$

The first line in Eq. (25) is the factor determining the dispersion of the transverse fluctuations of the velocity field. Under our assumptions this mode is completely decoupled from the rest of the system and remains diffusive for all values of the carrier density. This might change if one considers long-range disorder [63], where it was argued to induce vortical flow near charge neutrality.

The rest of the equation is best solved numerically. In Fig. 6 we present the results of a numerical calculation of the real part of the dispersion for the two values of the carrier density, $n = 10^{12}\,\mathrm{cm}^{-2}$ and $n = 10^{11}\,\mathrm{cm}^{-2}$. Equation (25) was solved using the typical values of the effective coupling constant [12, 64] $\alpha_g = 0.23$, disorder scattering time [12] $\tau_{\rm dis}^{-1} = 1\,\mathrm{THz}$, kinematic viscosity [3, 46] $\nu = 0.2\,\mathrm{m}^2/\mathrm{s}$, and temperature $T = 300\,\mathrm{K}$. The result is qualitatively similar to that shown in Fig. 1, therefore we postpone the discussion until after we have considered the two limiting cases where the dispersion can be obtained analytically, see Eq. (2).

1. *Collective modes in neutral graphene*

At charge neutrality, the linearized equations (21) can be simplified using the fact that the "conductivity matrices" $\widehat{\Sigma}$ and $\widehat{\Sigma}'$ are block-diagonal (here we take into account weak disorder)

$$\widehat{\Sigma}=\frac{1}{e^2}\begin{pmatrix}\sigma_0 & 0\\ 0 & \sigma_I\delta_I\end{pmatrix},\quad \widehat{\Sigma}'=\frac{1}{e^2}\begin{pmatrix}\sigma_0 & 0\\ 0 & \sigma_I\end{pmatrix}. \qquad (26)$$

As a result, the dissipative corrections (23) simplify

$$\delta\tilde{\boldsymbol{j}}=\frac{1}{e}\sigma_0\boldsymbol{\mathcal{E}}-\frac{i\tilde{\boldsymbol{q}}\sigma_0}{e^2\ln 2}\delta\tilde{n}, \qquad (27a)$$

$$\delta\tilde{\boldsymbol{j}}_I=-\frac{i\tilde{\boldsymbol{q}}\sigma_I}{e^2\ln 2}\left(\delta\tilde{n}_I-\frac{2\pi^2}{27\zeta(3)}\delta\tilde{n}_E\right). \qquad (27b)$$

Using the explicit form of the equilibrium quantities [16], we rewrite Eqs. (21) in the form

$$\tilde{\boldsymbol{q}}\delta\tilde{n}_E-\frac{9\zeta(3)}{2}\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)\mathbf{v}=0, \qquad (28a)$$

$$\left(\tilde{\omega}+i\frac{2\pi\tilde{q}^2\sigma_0}{e^2N\ln 2}\right)\delta\tilde{n}=\frac{2\pi\sigma_0}{eN}\tilde{\boldsymbol{q}}\cdot\boldsymbol{\mathcal{E}}, \qquad (28b)$$

$$\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_R}+\frac{i2\pi\tilde{q}^2\sigma_I}{e^2N\ln 2}\right)\delta\tilde{n}_I-\frac{\pi^2}{6}\tilde{\boldsymbol{q}}\cdot\mathbf{v}-\frac{i4\pi^3\tilde{q}^2\sigma_I\delta\tilde{n}_E}{27\zeta(3)Ne^2\ln 2}=0, \qquad (28c)$$

$$2\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\delta\tilde{n}_E-\frac{9\zeta(3)}{2}\tilde{\boldsymbol{q}}\cdot\mathbf{v}=0. \qquad (28d)$$

Combining Eqs. (28a) and (28d) to exclude the velocity, one finds

$$\tilde{q}^2\delta\tilde{n}_E=2\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\delta\tilde{n}_E,$$

yielding the spectrum (2) [in dimensionless units; in Eq. (2) we have neglected weak energy relaxation]

$$\tilde{\omega}=\sqrt{\frac{\tilde{q}^2}{2}-\frac{1}{4}\left[\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}-\frac{1}{\tilde{\tau}_{RE}}\right]^2}-i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{2\tilde{\tau}_{\rm dis}}-\frac{i}{2\tilde{\tau}_{RE}}. \qquad (29)$$

In the absence of dissipation this is the so-called "cosmic sound" wave [2, 15, 20] with the linear dispersion (1).

Same conclusions can be reached using the general form Eq. (25). At charge neutrality, Eq. (25) factorizes

$$\begin{aligned}
&\left[\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)-\frac{\tilde{q}^2}{2}\right]\left[\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right] \qquad (30)\\
&\times\left[\tilde{\omega}+\frac{i}{\tilde{\tau}_R}+\frac{2\pi i\tilde{q}^2\sigma_I}{Ne^2\ln 2}\right]\left[\tilde{\omega}+\frac{2\pi i\tilde{q}^2\sigma_0}{Ne^2}\left(\tilde{V}_{\tilde{\boldsymbol{q}}}+\frac{1}{\ln 2}\right)\right]=0.
\end{aligned}$$

Here the first factor yields the spectrum (29), the last factor describes the transverse fluctuations of the velocity field, while the remaining two correspond to the charge and imbalance modes.

The sound mode (29) is the energy wave not involving charge density fluctuations [since neither Eq. (28a) nor Eq. (28d) contains $\delta\tilde{n}$]. Consequently, the sound spectrum is not affected by the Vlasov self-consistency (17).

Other modes are diffusive. Since Eqs. (28a) and (28d) are independent of the density fluctuations $\delta\tilde{n}$ and $\delta\tilde{n}_I$, the diffusive modes can be read off Eqs. (28b) and (28c).

The electric charge density fluctuations are decoupled from the rest of the variables. Restoring the dimensionfull units and using the explicit form (5) of the conductivity at charge neutrality [1, 2, 16, 19, 43, 44, 52, 62, 65] we can write the corresponding dispersion as

$$\omega = -iD_0 q^2\left[1+eV_s(q)\frac{\partial n}{\partial \mu}\right], \quad D_0 = \frac{1}{2}\frac{v_g^2\tau_{11}\tau_{\rm dis}}{\tau_{11}+\tau_{\rm dis}}. \tag{31}$$

In a gated structure the mode is diffusive with the diffusive coefficient containing a correction due to the Vlasov self-consistency. In the case of the long-range Coulomb interaction the dispersion is still purely imaginary, with $\omega \sim iq$ at small q.

Similarly, the imbalance mode is characterized by the diffusive spectrum

$$\omega = -iD_I q^2 - \frac{i}{\tau_R}, \quad D_I = \frac{1}{2}\frac{v_g^2\tau_{22}\tau_{\rm dis}\delta_I}{\tau_{22}\delta_I+\tau_{\rm dis}}, \tag{32}$$

which is gapped by the recombination processes.

The hydrodynamic theory outlined in Section I is justified by the gradient expansion and hence for momenta smaller than a certain scale defined by the electron-electron interaction

$$q\ell_{\rm hydro} \ll 1, \quad \ell_{\rm hydro} \sim \frac{v_g}{\alpha_g^2 T}.$$

Assuming an ultra-clean sample with $\tau_{\rm dis} \to \infty$ (where energy relaxation due to supercollisions [55] may be neglected, $\tau_{RE} \gg \tau_{\rm dis}$), the expression under the square root in Eq. (29) yields

$$\frac{v_g^2 q^2}{2} - \frac{\left(1+q^2\ell_G^2\right)^2}{4\tau_{\rm dis}^2} \to \frac{v_g^2 q^2}{2}\left[1 - Aq^2\ell_{\rm hydro}^2 - \mathcal{O}(\tau_{\rm dis}^{-1})\right],$$

where A is a numerical coefficient. As a result, within the region of applicability of the hydrodynamic theory the viscous term should be neglected. The resulting dispersion acquires a simple form [15]

$$\omega = \sqrt{\frac{v_g^2 q^2}{2} - \frac{1}{4\tau_{\rm dis}^2}} - \frac{i}{2\tau_{\rm dis}}, \tag{33}$$

illustrated in Fig. 7. Now, keeping the viscous term to the leading order, but neglecting disorder scattering [23] yields an expansion

$$\omega = \frac{v_g q}{\sqrt{2}}\left(1 - \frac{\nu^2 q^2}{4v_g^2}\right) - \frac{i\nu q^2}{2}. \tag{34}$$

Similar expression was obtained in Ref. 23 based on the phenomenological collision integral (which did not take into account graphene-specific collinear scattering singularity). However, the viscosity-induced correction to the real part was positive indicating a tendency towards an indefinite growth of the dispersion instead of the decrease towards zero implied in Eqs. (1) and (29) and illustrated in Figs. 1 and 6. The sign of the correction in Eq. (34) is, in fact, dictated by the dissipative nature of viscosity, which represents an additional decay mechanism and hence affects the dispersion similarly to weak disorder; see Eq. (33). Indeed, both terms, $\tau_{\rm dis}^{-1}$ and νq^2, enter the dispersion equation [following from the first term in Eq. (30)] on equal footing.

As shown in Refs. 15, 43, 45, and 66 the linearized theory (18) has a wider applicability range due to the kinematic peculiarity of the Dirac fermions in graphene known as the "collinear scattering singularity" [1, 2, 15, 44]. In the weak coupling limit, the linear response theory is valid at much shorter length scales

$$q\ell_{\rm coll} \ll 1, \quad \ell_{\rm coll} \sim \frac{v_g}{\alpha_g^2 T|\ln\alpha_g|} \ll \ell_{\rm hydro}. \tag{35}$$

At the same time, the viscous term is the result of the gradient expansion that is justified at smaller momenta

$$q\ell_{\rm hydro} \ll 1,$$

which formally restricts us to small values of $\nu q/v_g$, such that the result (29) should be expressed in terms of the expansion (34). Moreover, the imaginary part of the sound dispersion becomes comparable to the real part at $q\ell_{\rm hydro} \sim 1$, such that the decline of the dispersion at larger q shown in Figs. 1 and 6 is unlikely to be observable anyway. Nevertheless in Figs. 1, 6, and 10 we show the sound dispersion in the whole range of momenta to illustrate the analytic structure of our results.

For realistic model parameters, the dispersion (29) shown in Figs. 1 and 6 is overdamped practically over the whole range of momenta. In the limit of large $\tau_{\rm dis}$ and small viscosity, the dispersion (29) approaches the ideal sound dispersion (1) if

$$(v_g\tau_{\rm dis})^{-1} \ll q \ll \ell_G^{-1}.$$

However, taking into account the numerical prefactors and realistic parameter values leads to Figs. 1 and 6, where the dispersion strongly deviates from Eq. (1).

2. *Collective modes in the degenerate regime*

In the opposite limit of the degenerate regime, $\mu \gg T$, the matrix in the left-hand side of Eq. (24) simplifies to

$$\begin{pmatrix} \tilde{\omega} & 0 & 0 & -\tilde{n}\boldsymbol{q} \\ 0 & \tilde{\omega}+\frac{i}{\tilde{\tau}_R} & 0 & -\tilde{n}_I\boldsymbol{q} \\ 0 & 0 & \tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}} & -\frac{3}{2}\tilde{n}_E\boldsymbol{q} \\ -\boldsymbol{q}\tilde{V}_q\frac{e\tilde{n}}{6} & 0 & -\frac{\boldsymbol{q}}{3} & \left[\tilde{\omega}+i\frac{1+q^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right]\tilde{n}_E \end{pmatrix}, \tag{36}$$

such that Eq. (25) factorizes again

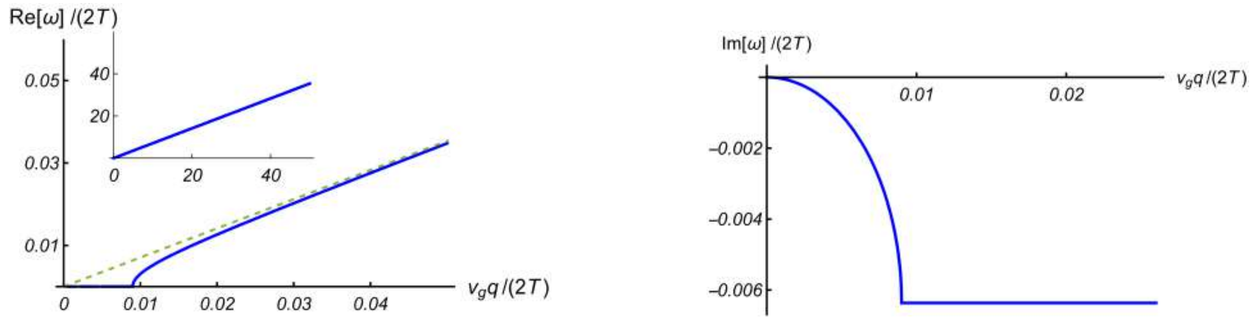


FIG. 7. Real (left panel) and imaginary (right panel) parts of the sound dispersion in neutral graphene neglecting viscosity. The dashed line represents the ideal “cosmic sound” dispersion (1).

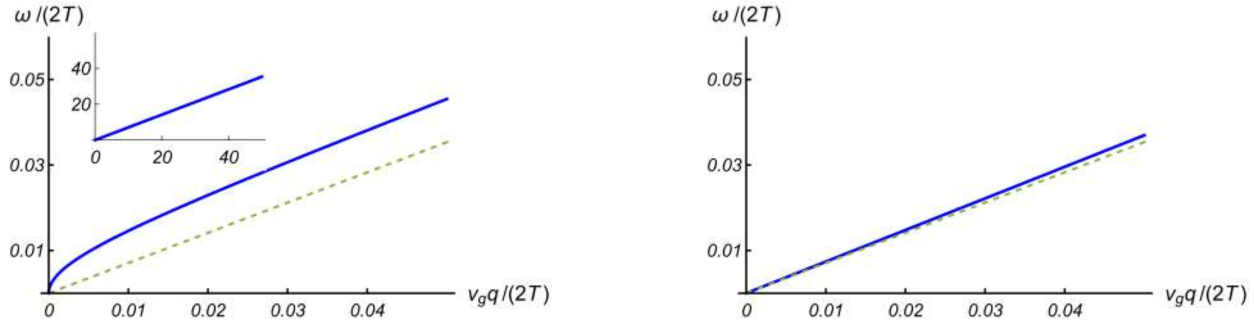


FIG. 8. Sound dispersion in strongly doped graphene neglecting both weak disorder and viscosity. Left panel: the result for the Coulomb screening, resembling the 2D plasmon for very low q. Right panel: same for a gated structure. The dashed line represents the ideal “cosmic sound” dispersion (1).

$$\left\{\tilde{\omega}\left[\left(\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right)\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)-\frac{\tilde{q}^2}{2}\right]-\frac{\tilde{q}^2\tilde{V}_q e\tilde{n}^2}{6\tilde{n}_E}\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\right\}\left[\tilde{\omega}+\frac{i}{\tilde{\tau}_R}\right]\left[\tilde{\omega}+i\frac{1+\tilde{q}^2\tilde{\ell}_G^2}{\tilde{\tau}_{\rm dis}}\right]=0. \tag{37}$$

The transverse velocity fluctuations remain decoupled with the same diffusive dispersion. The imbalance mode is no longer diffusive: if created, any imbalance density fluctuations decay exponentially in agreement with physical intuition.

The charge and energy densities are now coupled by the self-consistent Vlasov field. The corresponding dispersion can be found by equating the expression in curly brackets in Eq. (37) to zero. This leads to a cubic equation that can be solved exactly, but the analytic solution is cumbersome and not physically transparent. Instead, we focus on the limit $\tau_{RE}\gg\tau_{\rm dis}$ solving the equation perturbatively. Neglecting energy relaxation yields two modes, one being a flat zero mode and another the “sound mode” (29) renormalized by the Vlasov self-consistency. To the leading order in energy relaxation, the zero mode in a gated structure acquires the diffusive dispersion

$$\omega=-\frac{i}{\tau_{RE}}\frac{\varkappa v_g^2 q^2}{(\varkappa+2\pi C)v_g^2q^2+4\pi C\tau_{RE}^{-1}\tau_{\rm dis}^{-1}}, \tag{38}$$

where the Thomas-Fermi screening length is given by

$$\varkappa=N\alpha_g k_F=Ne^2\mu/v_g^2. \tag{39}$$

In the case of the long-range Coulomb interaction, the factor $2\pi C$ should be replaced with the momentum q. Physically, Eq. (38) describes energy diffusion appearing due to Vlasov self-consistency that couples charge and energy fluctuations.

Similarly to the above limit of neutral graphene, these results can be obtained from a direct analysis of the linearized hydrodynamic equations (21). In the degenerate regime ($\mu\gg T$ or $x\gg 1$), Eqs. (21) can be simplified by noticing that only one band contributes. For electron doping, $n\approx n_I$, while the dissipative corrections to the currents vanish [16]

$$\delta\boldsymbol{j}(T\ll\mu)=\delta\boldsymbol{j}_I(T\ll\mu)=0.$$

As a result, one of Eqs. (21) is redundant.

Assuming a gated structure and substituting the explicit form of equilibrium densities, we find

$$\left[\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{\rm dis}}\right)\frac{x^3}{2}+i\tilde{q}^2\tilde{\eta}\right]\mathbf{v}-\tilde{\boldsymbol{q}}\delta\tilde{n}_E = ie\left[\boldsymbol{\mathcal{E}}_0-\frac{i\tilde{\boldsymbol{q}}e}{\tilde{C}}\delta\tilde{n}\right]\frac{x^2}{4}, \tag{40a}$$

$$\tilde{\omega}\delta\tilde{n}-(x^2/2)\tilde{\boldsymbol{q}}\cdot\mathbf{v}=0, \tag{40b}$$

$$2\left(\tilde{\omega}+\frac{i}{\tilde{\tau}_{RE}}\right)\delta\tilde{n}_E-(x^3/2)\tilde{\boldsymbol{q}}\cdot\mathbf{v}=0. \tag{40c}$$

Combining Eqs. (40a) and (40c) one finds the cosmic sound mode [2, 15, 20] damped by disorder and viscosity (back to dimensionful units and for $\tau_{RE}\gg\tau_{\rm dis}$)

$$\omega=\sqrt{\frac{v_g^2q^2}{2}\left[1+\frac{\varkappa}{2\pi C}\right]-\frac{\left(1+\ell_G^2q^2\right)^2}{4\tau_{\rm dis}^2}}-\frac{i\left(1+\ell_G^2q^2\right)^2}{2\tau_{\rm dis}}. \tag{41}$$

This is clearly the same mode as Eq. (29), albeit with the velocity renormalized by the capacitive screening.

Long-range Coulomb interaction modifies the screening contribution to the sound mode (41)

$$\omega=\sqrt{\frac{v_g^2q^2}{2}\left[1+\frac{\varkappa}{q}\right]-\frac{\left(1+\ell_G^2q^2\right)^2}{4\tau_{\rm dis}^2}}-\frac{i\left(1+\ell_G^2q^2\right)^2}{2\tau_{\rm dis}}. \tag{42}$$

Taking the naive limit $q\to 0$ (and $x\to\infty$) in Eq. (42), one arrives at the spectrum similar to the usual two-dimensional plasmon [15, 67]

$$\omega(q\ll\varkappa)=-\frac{i}{2\tau_{\rm dis}}+\sqrt{\frac{1}{2}v_g^2q\varkappa-\frac{1}{4\tau_{\rm dis}^2}}. \tag{43}$$

The dispersion (43) is meaningful if the following conditions are met

$$q\ell_G\ll 1, \quad q\ll\varkappa, \quad v_g^2\varkappa q\tau_{\rm dis}^2\gg 1.$$

At the same time, for the hydrodynamic approach to be valid at all, the gradients are supposed to be small on the scale that is defined by the electron-electron interaction

$$q\ell_{\rm hydro}\ll 1, \quad \ell_{\rm hydro}\sim\frac{v_g}{\alpha_g^2\bar{T}}.$$

These conditions to be consistent if (using the explicit form of physical quantities in the degenerate regime)

$$v_g\varkappa\tau_{\rm dis}\gg 1 \quad\Rightarrow\quad N\alpha_g\mu\tau_{\rm dis}\gg 1,$$

$$\ell_G\ll v_g^2\varkappa\tau_{\rm dis}^2 \quad\Rightarrow\quad N^2\alpha_g^4\mu\tau_{\rm dis}(\bar{T}\tau_{\rm dis})^2\gg 1,$$

providing a possibility to observe the dispersion (43) in a parametrically defined range of wavevectors.

The eigenvectors of the "flat zero mode" and the sound mode mix the charge, energy density, and velocity fluctuations. In that sense, the mode (43) is not a true plasmon, even though its dispersion is identical with that of the usual plasmon in two dimensions. Moreover, the dispersion (43) resembles the plasmon dispersion only in an intermediate interval of rather small q, while the true plasmon exists at large values of q.

The above dispersion can be illustrated numerically as follows. Using the same typical values $\tau_{\rm dis}^{-1}=1\,{\rm THz}$, $\nu=0.2\,{\rm m^2/s}$ (the kinematic viscosity varies only weakly with the carrier density [46]), and $T=300\,{\rm K}$, as well as the typical value of the coupling constant [12, 64] $\alpha_g=0.23$ and the parameters characterizing the external gate in a typical graphene-on-boron nitride structure [3], the dielectric constant of the hexagonal boron nitride $\epsilon=4.4$ and the graphene to gate distance $d=80\,{\rm nm}$, we plot the two dispersions (41) and (42) in Figs. 8-10. In Fig. 8, we show the two dispersions (41) and (42) in the absence of both weak disorder and viscosity. The effect of the screening can be summarized as follows. In a gated structure screening leads to a slight (for the realistic parameter values chosen above) change of slope of the sound mode dispersion. In contrast, Coulomb screening leads to a plasmon-like square-root dispersion for the smallest values of momentum, which soon turns into a linear dispersion with the same slope as the "cosmic sound" of the ideal fluid, but slightly (again, for the realistic parameter values) shifted upwards. Taking into account dissipative processes washes out qualitative differences between different types of screening. The results are also qualitatively the same for strongly doped and neutral graphene. In Fig. 9 we show the results for the dispersion in the presence of weak disorder, but still neglecting viscosity. Qualitatively, the results for both types of screening are similar with the only difference being that the real part of the dispersion in the case of the Coulomb screening is shifted upwards relative to the ideal sound dispersion, similarly to the left panel in Fig. 8, while in the case of the gated structure the resulting straight line at large enough q has a slightly larger slope than $1/\sqrt{2}$.

Once viscosity is taken into account, the curves in Fig. 10 strongly resemble the results in neutral graphene, cf. Fig. 1. The results for gated graphene show only insignificant numerical differences from the curves in Fig. 1, while in the case of the Coulomb screening the real part of the dispersion appears at a smaller value of q and exceeds the ideal spectrum (represented in all figures by the dotted line) in a small intermediate range of q.

IV. HYDRODYNAMIC COLLECTIVE MODES AND PLASMONS

The hydrodynamic approach is applicable in the long-time and long-wavelength limit [1, 2, 45, 68], i.e., at momenta that are small compared to the typical "equilibration" length scale $\ell_{\rm hydro}$. At higher momenta (and frequencies), the system is not in equilibrium. In this regime (sometimes referred to [27] as "collisionless"), the electronic fluid exhibits well-known collective excitations, the plasmons. In two dimensions and in the absence of

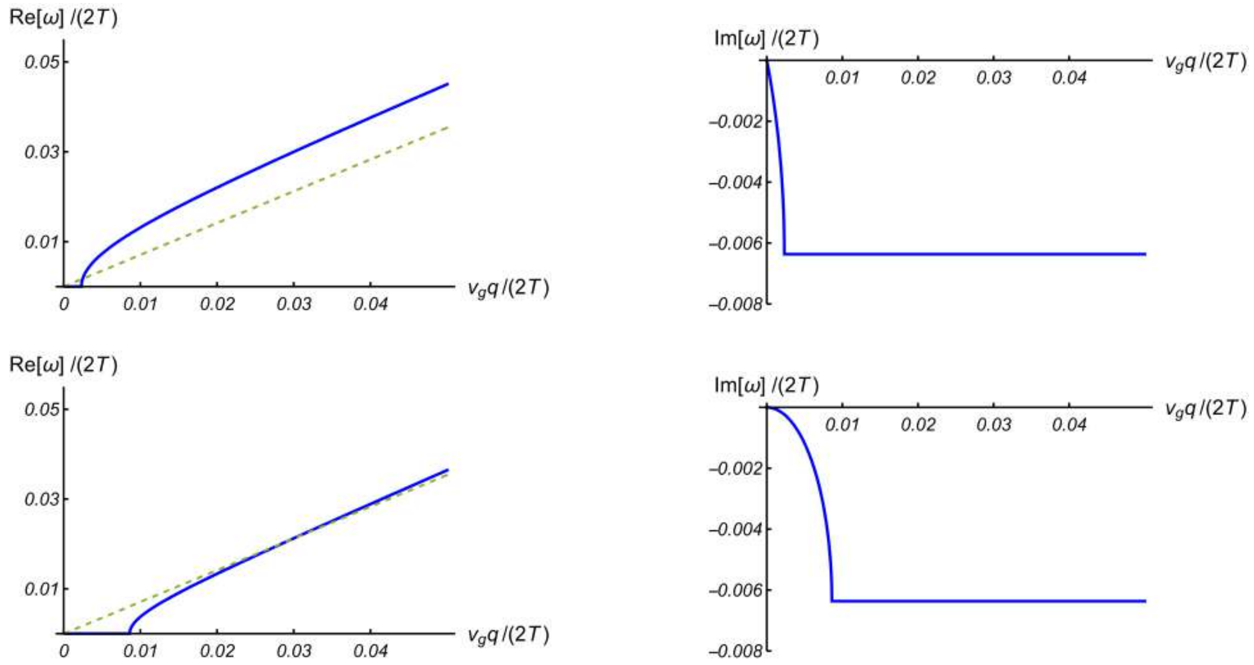


FIG. 9. Real and imaginary parts of the sound dispersion in strongly doped graphene in the presence of weak disorder, but neglecting viscosity. Top panels: the result for the Coulomb screening. Bottom panels: same for a gated structure. Dashed lines represents the ideal "cosmic sound" dispersion (1).

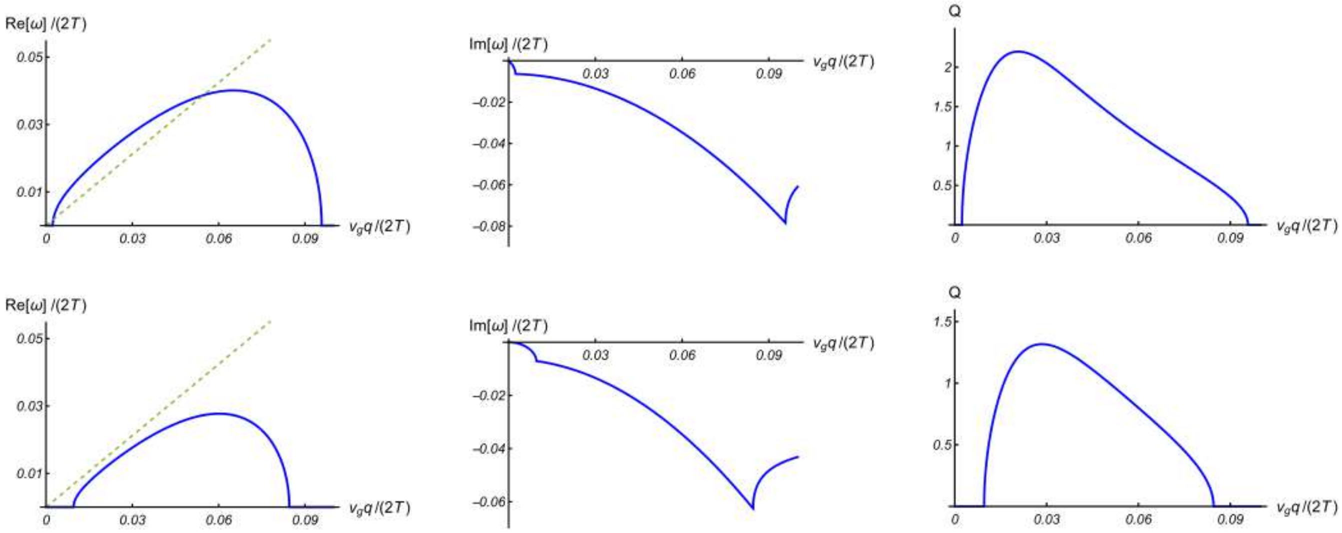


FIG. 10. Real and imaginary parts and the quality factor $Q = \mathrm{Re}\,\omega/\mathrm{Im}\,\omega$ of the sound dispersion in strongly doped graphene in the presence of both weak disorder and viscosity. Top panels: the result for the Coulomb screening. Bottom panels: same for a gated structure. Dashed lines represent the ideal "cosmic sound" dispersion (1).

impurity scattering ($\tau_{\mathrm{dis}} \to \infty$) the plasmon dispersion in the degenerate electron gas has the form [27]

$$\omega = \sqrt{2e^2\mu q}\left(1+\gamma\frac{q}{\varkappa}\right), \tag{44}$$

where γ is a numerical coefficient (see below). The "proper" way to derive Eq. (44) is to evaluate the Lindhard function within the random phase approximation (RPA), which would lead [27] to the coefficient $\gamma = 3/4$. An attempt to derive the plasmon dispersion from a macroscopic (hydrodynamic-like) theory leads to the same form (44), but with a different value for γ. This discrepancy is well known and can be attributed to the failure of the hydrodynamic description at high frequencies and momenta [27]. As a result, one concludes that the hydrodynamic collective modes have nothing to do with plasmons simply because they belong to a different parameter regime. In this Section we extend these arguments to Dirac fermions in graphene and establish the relation between the above hydrodynamic modes and plasmons.

A. Degenerate regime

The case of graphene is special because of the kinematic peculiarity known as the "collinear scattering singularity" [1, 2, 15, 16, 19, 43, 44, 62, 66] leading to the existence of the two parametrically (in the weak coupling limit) different length scales associated with electron-electron interactions, $\ell_{\rm coll} \ll \ell_{\rm hydro}$. In an intermediate momentum range, $\ell_{\rm hydro}^{-1} \ll q \ll \ell_{\rm coll}^{-1}$, the hydrodynamic theory of Section I breaks down, while the linear response theory of Ref. 43 is still valid. Remarkably, the macroscopic equations of the latter theory are identical with the linearized hydrodynamic equations, so that the collective modes in the two parameter regimes coincide.

In the degenerate regime and in the absence of magnetic field, the linear response theory [43] reduces to the single macroscopic equation describing the dynamics of the electric current $\boldsymbol{J}$ (here ρ is the charge density)

$$\frac{\partial \boldsymbol{J}}{\partial t} + \frac{v_g^2}{2}\boldsymbol{\nabla}\rho - \nu\Delta\boldsymbol{J} - \frac{v_g^2}{2}\frac{\partial n}{\partial \mu}e^2\boldsymbol{E} = -\frac{\boldsymbol{J}}{\tau_{\rm dis}}, \tag{45}$$

which is essentially the generalized Ohm's law. To obtain the plasmon dispersion, we introduce the Vlasov field [cf. Eq. (17)] and use the continuity equation. In the case of Coulomb interaction, the standard algebra [27] leads to the following equation

$$\omega\left(1 + q^2\ell_G^2 - i\omega\tau_{\rm dis}\right) = -iDq^2 - i2\pi\sigma q,$$

where $D = v_g^2\tau_{\rm dis}/2$ and $\sigma = v_g^2(\partial n/\partial\mu)\tau_{\rm dis}/2$ are the diffusion coefficient and the Drude conductivity. The resulting spectrum has the form

$$\omega = \sqrt{2e^2\mu q\left(1+\frac{q}{\varkappa}\right) - \frac{(1+q^2\ell_G^2)^2}{4\tau_{\rm dis}^2}} - \frac{i(1+q^2\ell_G^2)}{2\tau_{\rm dis}}. \tag{46}$$

The spectrum (46) is exactly the same as Eq. (42). For a clean system ($\tau_{\rm dis} \to \infty$), the expansion for small $q \to 0$ yields the form (44) with the "wrong" coefficient, $\gamma = 1/2$. At the same time, the leading term (neglecting the correction for $q \ll \varkappa$) agrees with the standard Fermi liquid result even in the presence of disorder [67] (neglecting viscosity).

The expression (46) is valid for momenta up to $\ell_{\rm coll}^{-1}$, but in fact it becomes overdamped already at momenta of order $\ell_{\rm hydro}^{-1}$. At larger momenta, $q \gg \ell_{\rm coll}^{-1}$, the quasi-equilibrium description breaks down and the true plasmons emerge with the dispersion (44). By that time the spectrum (46) becomes purely imaginary (see Fig. 10), and hence the two modes are not connected. Similar conclusions have been reached in Ref. 24, where it was argued that Coulomb interaction precludes the appearance of hydrodynamic sound in Fermi liquids.

B. Two-fluid hydrodynamics

Let us slightly digress and consider the curious case of the two-fluid hydrodynamics [49, 50, 69] in compensated semimetals. Following Ref. 49 we assume that the full electronic systems comprises two weakly coupled fluids, one consisting of electrons and another of holes. This means that the length scales ℓ_{ee} and ℓ_{hh} describing intraband electron-electron scattering are much smaller than the interband scattering length ℓ_{eh}. In that case, the system is described by two equations similar to Eq. (45) with an extra interband scattering term

$$\frac{\partial \boldsymbol{j}_\alpha}{\partial t} + \frac{\langle v^2\rangle}{2}\boldsymbol{\nabla}n_\alpha - \nu\Delta\boldsymbol{j}_\alpha - \frac{\langle v^2\rangle}{2}\frac{\partial n_\alpha}{\partial\mu}e_\alpha\boldsymbol{E} = \tag{47}$$

$$= -\frac{\boldsymbol{j}_\alpha}{\tau_{\rm dis}} - \frac{\boldsymbol{j}_\alpha - \boldsymbol{j}_{\alpha'}}{2\tau_{eh}},$$

where $e_h = -e > 0$, $e_e = e < 0$, $\boldsymbol{j}_\alpha$ denotes the quasiparticle currents, and α' denotes the other constituent. For simplicity we assume the system to be electron-hole symmetric ($\ell_{ee} = \ell_{hh}$).

Combining the two currents into the linear combinations, $\boldsymbol{j} = \boldsymbol{j}_e - \boldsymbol{j}_h$ and $\boldsymbol{j}_I = \boldsymbol{j}_e + \boldsymbol{j}_h$, we find the decoupled (in the absence of the magnetic field) equations

$$\frac{\partial \boldsymbol{j}}{\partial t} - \nu\Delta\boldsymbol{j} - \frac{\langle v^2\rangle}{2}\frac{\partial n_I}{\partial\mu}e\boldsymbol{E} = -\frac{\boldsymbol{j}}{\tau_{\rm dis}} - \frac{\boldsymbol{j}}{\tau_{eh}}, \tag{48a}$$

$$\frac{\partial \boldsymbol{j}_I}{\partial t} + \frac{\langle v^2\rangle}{2}\boldsymbol{\nabla}n_I - \nu\Delta\boldsymbol{j}_I = -\frac{\boldsymbol{j}_I}{\tau_{\rm dis}}. \tag{48b}$$

Combining these equations with the two continuity equations (6b) and (6c), we find a sound-like mode

$$\omega = \sqrt{\frac{\langle v^2\rangle q^2}{2} - \left(\frac{1+q^2\ell_G^2}{2\tau_{\rm dis}} - \frac{1}{2\tau_R}\right)^2} - \frac{i(1+q^2\ell_G^2)}{2\tau_{\rm dis}} - \frac{i}{2\tau_R}, \tag{49a}$$

and a plasmon-like mode

$$\omega = \sqrt{\frac{\langle v^2\rangle\varkappa q}{2} - \frac{(1+q^2\ell_{G*}^2)^2}{4\tau_*^2}} - \frac{i(1+q^2\ell_{G*}^2)}{2\tau_*}, \tag{49b}$$

where

$$\varkappa = 2\pi e^2\frac{\partial n_I}{\partial\mu}, \qquad \tau_* = \frac{\tau_{\rm dis}\tau_{eh}}{\tau_{\rm dis}+\tau_{eh}}, \qquad \ell_{G*} = \sqrt{\nu\tau_*}.$$

In the hydrodynamic parameter range, both modes (49) are well defined. The expression under the square root in Eq. (49a) can be rewritten as

$$\frac{\langle v^2\rangle}{2}\left[q^2\left(1-\frac{\tau_{ee}}{\tau_{\rm dis}}+\frac{\tau_{ee}}{\tau_R}-\frac{q^2\ell_{ee}^2}{2}\right)-\frac{1}{2}\left(\frac{1}{\ell_{\rm dis}}-\frac{1}{\ell_R}\right)^2\right].$$

Here $\tau_{ee} \ll \tau_{\rm dis}$, $\tau_{ee} \ll \tau_R$ by the assumptions of the hydrodynamic regime and $q\ell_{ee} \ll 1$ under the assumption of the gradient expansion in the hydrodynamic theory (here we consider a generic semimetal and hence do not have the aforementioned scale separation specific to graphene, hence we cannot extend the argument beyond the validity region of the gradient expansion). Therefore apart from the small gap due to the interplay between disorder scattering and recombination processes, the sound mode is well defined within the hydrodynamic range of momenta.

Similar arguments can be extended to the plasmon-like mode (49b). Assuming a clean system, $\tau_{eh} \ll \tau_{\rm dis}$, $\tau_* \to \tau_{eh}$, one finds under the square root in Eq. (49b)

$$\frac{\langle v^2 \rangle}{2}\left[q\varkappa - q^2\frac{\tau_{ee}}{\tau_{eh}} - \frac{q^4\ell_{ee}^2}{2} - \frac{1}{2\ell_{eh}^2}\right].$$

Typically, the Thomas-Fermi screening radius is smaller then the electron-electron scattering length, $\varkappa\ell_{ee} \gg 1$. Hence, the mode (49b) is also well defined. Here the electron-hole scattering yields the (small) gap in the dispersion similarly to the disorder scattering in Eq. (46).

C. Graphene at charge neutrality

Utilizing the scale separation in graphene (see above), we can approach the question of the collective modes from the standpoint of the linear response theory of Ref. 43. Here, instead of formulating the hydrodynamic equations (6), we turn to the macroscopic equations describing the behavior of the three inequivalent currents in the system, $\boldsymbol{j}$, $\boldsymbol{j}_I$, and $\boldsymbol{j}_E$

$$\frac{\partial \boldsymbol{j}}{\partial t} + \frac{v_g^2}{2}\boldsymbol{\nabla} n - \frac{2\ln 2}{\pi}e^2 T\boldsymbol{E} = -\frac{\boldsymbol{j}}{\tau_{\rm dis}} - \frac{\boldsymbol{j}}{\tau_{11}}, \tag{50a}$$

$$\frac{\partial \boldsymbol{j}_I}{\partial t} + \frac{v_g^2}{2}\boldsymbol{\nabla} n_I - \frac{\gamma_1\nu}{T}\Delta\boldsymbol{j}_E = -\frac{\boldsymbol{j}_I}{\tau_{\rm dis}} - \frac{\boldsymbol{j}_I - \frac{\pi^2 \boldsymbol{j}_E}{27\zeta(3)T}}{\tau_{22}\delta_I}, \tag{50b}$$

$$\frac{\partial \boldsymbol{j}_E}{\partial t} + \frac{v_g^2}{2}\boldsymbol{\nabla} n_E - \nu\Delta\boldsymbol{j}_E = -\frac{\boldsymbol{j}_E}{\tau_{\rm dis}}, \tag{50c}$$

where γ_1 is a numerical prefactor. At charge neutrality, the viscous term vanishes from Eq. (50a) in contrast to the two-fluid model, see Eq. (49a). In graphene, the electron and hole subsystems are strongly coupled ($\ell_{ee} = \ell_{hh} \sim \ell_{eh}$) forming a single fluid, where the electric current is not affected by viscous effects because of electron-hole symmetry. Viscosity still affects neutral quasiparticle and energy flows in agreement with the hydrodynamic approach, where the hydrodynamic velocity in neutral graphene describes the flow of energy.

Similarly to the hydrodynamic regime (Section III 1), the energy and charge decouple completely. Combining Eq. (50c) with the continuity equation for the energy density (18d) – that is equivalent to the linearized heat transport equation (6d) – we recover the sound mode (2).

On the other hand, combining Eq. (50a) with the continuity equation (6b) we find

$$\omega^2 + i\omega\left(\frac{1}{\tau_{\rm dis}} + \frac{1}{\tau_{11}}\right) = \frac{v_g^2}{2}q^2 + (4\ln 2)e^2 Tq, \tag{51}$$

leading to the plasmon-like spectrum. For large enough frequencies, $\omega \gg \tau_{11}^{-1} \gg \tau_{\rm dis}^{-1}$, and small momenta, $q \to 0$, the resulting dispersion coincides with the leading behavior of the true plasmon dispersion established in Ref. 19

$$\omega = \sqrt{(4\ln 2)e^2 Tq} \quad \Rightarrow \quad \tilde{\omega} = \sqrt{2(\ln 2)\alpha_g \tilde{q}}, \tag{52}$$

where the last equality is expressed in terms of the dimensionless variables (22a), also used in Ref. 19. Note, that at large momenta, where the first term in the left-hand side of Eq. (51) dominates, the resulting dispersion resembles the cosmic sound (1), contradicting the result of Ref. 19, where the dispersion in the large-q limit also becomes linear, but without the extra $\sqrt{2}$.

Considering the limit $\tau_* \to \infty$ in Eq. (49b), we arrive at the same result [in graphene at the charge neutrality point, $v_g^2\varkappa/2 = (4\ln 2)e^2 T$, while viscosity does not affect charge transport]. In the absence of disorder, the two-fluid model considered in Section IV B describes the electron and hole subsystems as being weakly coupled (similarly to the effect of Coulomb drag [70], but without spatial separation). Charge density fluctuations are correspond to the out-of-phase motion of electrons and holes. In the absence of the electron-hole scattering ($\tau_* \to \infty$), charge transport is effectively decoupled from the in-phase (imbalance) mode and hence Eq. (48a) becomes equivalent to Eq. (50a) yielding the same plasmonic mode.

Rewriting Eq. (51) in the form

$$i\omega\left[-i\omega + \frac{1}{\tau_{\rm dis}} + \frac{1}{\tau_{11}}\right] = \frac{v_g^2}{2}q^2 + (4\ln 2)e^2 Tq,$$

we express the plasmon dispersion in the form closely resembling Eq. (31)

$$\omega = -i\frac{\sigma(\omega)q^2}{e^2\partial n/\partial\mu}\left[1 + eV_s(q)\frac{\partial n}{\partial\mu}\right],$$

where instead of the static conductivity (5) we find the optical conductivity [45]

$$\sigma(\omega) = \frac{2e^2 T\ln 2}{\pi}\frac{1}{-i\omega + \tau_{11}^{-1} + \tau_{\rm dis}^{-1}}.$$

In the hydrodynamic regime $\sigma(\omega \to 0) \to \sigma_0$ and we recover the diffusive mode (31).

Resolving Eq. (51) we find the full plasmon dispersion

$$\omega = -i\frac{\tau_{\rm dis} + \tau_{11}}{2\tau_{\rm dis}\tau_{11}} + \sqrt{\frac{v_g^2}{2}q^2 + (4\ln 2)e^2 Tq - \frac{(\tau_{\rm dis} + \tau_{11})^2}{4\tau_{\rm dis}^2\tau_{11}^2}}\,. \tag{53}$$

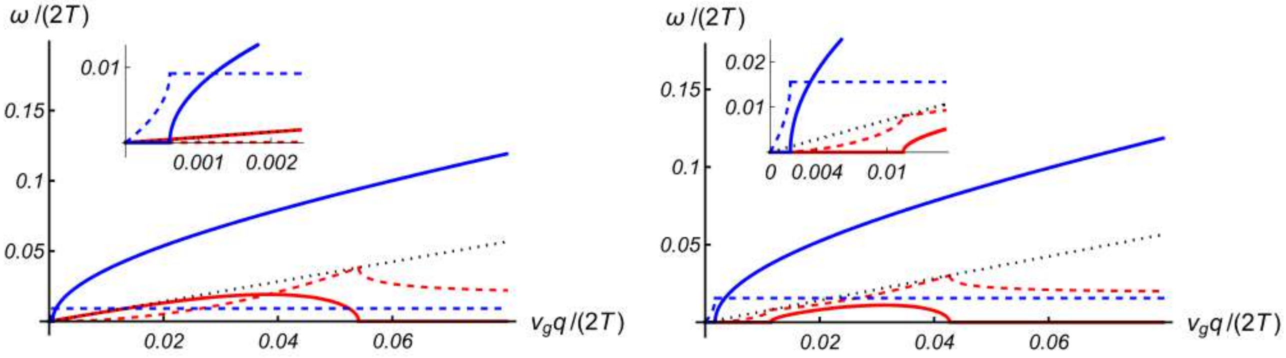


FIG. 11. Comparison between the plasmon mode (54) and the sound mode (55) within the linear response theory. Solid curves show the real part of the dispersion, dashed curves the absolute value of the imaginary part. The dotted line shows the ideal "cosmic sound" dispersion (1). The plasmon dispersion is shown in blue, the sound in red. The distinction between the two modes is clearly defined by their frequencies that are much higher for the plasmon mode. Left panel shows the dispersion for a clean sample; right panel the same for the typical value $\tau_{\rm dis}^{-1} = 1\,$THz. The coupling constant is taken at a model value $\alpha_g = 0.1$, hence, no renormalization of the velocity v_g is taken into account strongly underestimating viscosity. The real part of the sound dispersion vanishes at $\tilde{q} \approx 0.54$, which is similar to the applicability limit of the linear response theory, $\ell_{\rm coll}^{-1}$. The imaginary part exceeds the real part at a lower value of $\tilde{q}$, such that the mode becomes overdamped and disappears still within the applicability region of the theory. In the presence of disorder (right panel) the sound model is completely overdamped, see Fig. 1 for more realistic values.

To analyze the two modes – the plasmon and sound – together, we rewrite the above dispersion in dimensionless units (22a). The plasmon dispersion takes the form

$$\tilde{\omega}_p = \sqrt{2(\ln 2)\alpha_g\tilde{q}\left[1+\frac{\tilde{q}}{4(\ln 2)\alpha_g}\right] - \left[\frac{1}{2\tilde{\tau}_{\rm dis}}+\frac{\alpha_g^2\ln 2}{2\pi\mathcal{A}}\right]^2} \\ -\frac{i}{2}\left(\frac{1}{\tilde{\tau}_{\rm dis}}+\frac{\alpha_g^2\ln 2}{\pi\mathcal{A}}\right), \tag{54}$$

where the constant $\mathcal{A} \approx 0.12$ determines the quantum conductivity at charge neutrality [1, 2, 16, 52]

$$\sigma_Q = \mathcal{A}e^2/\alpha_g^2.$$

At the same time, the sound dispersion (1) is given by

$$\tilde{\omega}_s = \sqrt{\frac{\tilde{q}^2}{2} - \left[\frac{1}{2\tilde{\tau}_{\rm dis}}+\frac{\pi\mathcal{B}\tilde{q}^2}{9\zeta(3)\alpha_g^2}\right]^2} - \frac{i}{2}\left(\frac{1}{\tilde{\tau}_{\rm dis}}+\frac{2\pi\mathcal{B}\tilde{q}^2}{9\zeta(3)\alpha_g^2}\right), \tag{55}$$

where the constant $\mathcal{B} \approx 0.45$ determines the shear viscosity in neutral graphene [1, 2, 16, 17]

$$\eta(\mu{=}0) = \mathcal{B}T^2/(\alpha_g^2 v_g^2).$$

In pure graphene ($\tilde{\tau}_{\rm dis} \to \infty$) in the weak coupling limit ($\alpha_g \to 0$), the regions where the two dispersions are real overlap: the plasmon dispersion (54) is real for $\tilde{q} \gg \alpha_g^3$, while the sound dispersion (55) is real for $\tilde{q} \ll \alpha_g^2$. Weak disorder does not yield any qualitative changes.

The linear response theory, Eqs. (50), is applicable at length scales larger than $\ell_{\rm coll}$, the graphene-specific scale [see Eq. (35)], reflecting the collinear scattering singularity. In dimensionless units, $\ell_{\rm coll}^{-1} \sim \alpha_g^2|\ln\alpha_g|$, which in the weak coupling limit greatly exceeds $\ell_{\rm hydro}^{-1} \sim \alpha_g^2$, which determines the applicability of the hydrodynamic theory of Section I. In the limit $\tilde{\tau}_{\rm dis} \to \infty$, the real part of the sound dispersion (55) vanishes when

$$\tilde{q} = \tilde{q}_0 = \frac{9\zeta(3)}{\sqrt{2\pi}\mathcal{B}}\alpha_g^2 \approx 5.41\alpha_g^2.$$

Here the large numerical coefficient may mask the difference between the two length scales $\ell_{\rm hydro}$ and $\ell_{\rm coll}$ for all but the lowest values of α_g. We illustrate the resulting dispersions in Fig. 11, where we use a model value $\alpha_g = 0.1$ to keep the two length scales well separated. Even though $\tilde{q}_0$ is of the same order of magnitude as $\ell_{\rm coll}$, the imaginary part of the dispersion becomes comparable to the real part at a significantly lower value of $\tilde{q}$. At that point the mode becomes overdamped and essentially disappears. Adding realistic disorder renders the mode completely overdamped, see the right panel in Fig. 11.

V. SUMMARY

In this paper we described electronic collective modes in graphene based on the hydrodynamic approach and compared the results with the more general linear response theory. Our results generalize the discussion of these issues reported in Ref. 15 within the small momentum expansion. Given the universality of hydrodynamics, the results for the collective modes in the hydrodynamic regime are applicable to other semimetals (where the momentum density represented by $\boldsymbol{u}$ is effectively decoupled from the charge transport unless the system is doped far away from charge neutrality), while the three-mode approximation used to derive the linear response theory discussed in Section IV is specific to graphene.

Our main results are illustrated in Figs. 1 and 11. The former shows the dispersion of the sound mode in the hydrodynamic regime with the viscous damping and weak disorder taken into account. Using the typical experimental values of the viscosity and disorder scattering time, we find that the sound mode in real graphene is strongly damped, making it difficult to observe the ideal "cosmic sound" dispersion (1) experimentally.

In Fig. 11 we illustrate the sound and plasmon modes in neutral graphene obtained within the linear response theory of Ref. 43 (extended beyond the stationary and uniform fields). Both modes are evaluated with the "bare" parameter values (ignoring, e.g., the renormalization of quasiparticle spectrum in graphene [46, 71]) for clarity. Effectively, this approach strongly underestimates the kinematic viscosity and hence the sound mode in Fig. 11 is much more pronounced than in Fig. 1.

The plasmon mode (54) is characterized by higher frequencies that the sound mode (55) and hence is not accessible within the standard hydrodynamic approach of Section I. The connection between the two calculations can be made by allowing for the frequency-dependent (optical) conductivity in Eqs. (25) and (31). Reducing the dissipative coefficients in the hydrodynamic theory to frequency-independent constants (following the standard approach of Ref. 18) leads to the diffusive behavior of the collective charge fluctuations, see Eq. (31). Similarly, all other hydrodynamic collective modes (except for the sound mode) are characterized by purely imaginary spectra. This should be contrasted with the linear response theory, Eqs. (50), that allows for the frequency-dependent conductivities leading to the real plasmon dispersion (54), as well as a third (neutral) collective mode following from Eqs. (50b) and (6c). The fact that these additional (to the sound) modes can be reached within the linear response theory and connected to the hydrodynamic description should be attributed to the scale separation in graphene (due to the kinematic peculiarity of Dirac fermions [1, 2, 16, 44, 62]), see Eq. (35). All other qualitative conclusions of the paper are valid in a wider class of semimetals. The obtained collective modes can be observed using the by now standard plasmonics experiments, see Refs. 30–38.

ACKNOWLEDGMENTS

The authors are grateful to U. Briskot, A.D. Mirlin, J. Schmalian, M. Schütt, and A. Shnirman for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 program (Project 873028 HYDROTRONICS), and by the Russian Science Foundation, Grant No. 17-12-01182 c (IG). BNN acknowledges the support by the MEPhI Academic Excellence Project, Contract No. 02.a03.21.0005.

---

[1] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[2] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).

[3] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[4] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, et al., Science **351**, 1058 (2016).

[5] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[6] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).

[7] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[8] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. Krishna Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

[9] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[10] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, et al., Phys. Rev. B **98**, 241304(R) (2018).

[11] A. Jaoui, B. Fauqué, C. W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D. L. Maslov, C. Felser, et al., npj Quantum Materials **3**, 64 (2018).

[12] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[13] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).

[14] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[15] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).

[16] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[17] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[18] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1959).

[19] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).

[20] T. V. Phan, J. C. W. Song, and L. S. Levitov (2013), arXiv:1306.4972.

[21] L. S. Levitov, A. V. Shtyk, and M. V. Feigelman, Phys. Rev. B **88**, 235403 (2013).

[22] Z. Sun, D. N. Basov, and M. M. Fogler, Proceedings of the National Academy of Sciences **115**, 3285 (2018).

[23] D. Svintsov, Phys. Rev. B **97**, 121405(R) (2018).

[24] A. Lucas and S. Das Sarma, Phys. Rev. B **97**, 115449 (2018).

[25] E. I. Kiselev and J. Schmalian, Phys. Rev. B **102**, 245434 (2020).

[26] D. V. Fateev and V. V. Popov, Semiconductors **54**, 941 (2020).

[27] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[28] A. Hill, S. A. Mikhailov, and K. Ziegler, EPL (Europhysics Letters) **87**, 27005 (2009).

[29] A. Principi, R. Asgari, and M. Polini, Solid State Communications **151**, 1627 (2011).

[30] Z. Fei, A. S. Rodin, G. O. Andreev, W. Bao, A. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, et al., Nature **487**, 82 (2012).

[31] J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, et al., Nature **487**, 77 (2012).

[32] G. X. Ni, L. Wang, M. D. Goldflam, M. Wagner, Z. Fei, A. S. McLeod, M. K. Liu, F. Keilmann, B. Özyilmaz, A. H. Castro Neto, et al., Nature Photonics **10**, 244 (2016).

[33] M. B. Lundeberg, Y. Gao, R. Asgari, C. Tan, B. Van Duppen, M. Autore, P. Alonso-González, A. Woessner, K. Watanabe, T. Taniguchi, et al., Science **357**, 187 (2017).

[34] D. Alcaraz Iranzo, S. Nanot, E. J. C. Dias, I. Epstein, C. Peng, D. K. Efetov, M. B. Lundeberg, R. Parret, J. Osmond, J.-Y. Hong, et al., Science **360**, 291 (2018).

[35] P. Novelli, I. Torre, F. H. L. Koppens, F. Taddei, and M. Polini, Phys. Rev. B **102**, 125403 (2020).

[36] T. Giovannini, L. Bonatti, M. Polini, and C. Cappelli, The Journal of Physical Chemistry Letters **11**, 7595 (2020).

[37] N. C. H. Hesp, I. Torre, D. Rodan-Legrain, P. Novelli, Y. Cao, S. Carr, S. Fang, P. Stepanov, D. Barcons-Ruiz, H. Herzig-Sheinfux, et al. (2019), arXiv:1910.07893.

[38] A. T. Costa, P. A. D. Gonçalves, D. N. Basov, F. H. L. Koppens, N. A. Mortensen, and N. M. R. Peres, PNAS **118**, e2012847118 (2021).

[39] A. Klein, D. L. Maslov, L. P. Pitaevskii, and A. V. Chubukov, Phys. Rev. Research **1**, 033134 (2019).

[40] B. A. Ferreira, B. Amorim, A. J. Chaves, and N. M. R. Peres, Phys. Rev. A **101**, 033817 (2020).

[41] A. Klein, D. L. Maslov, and A. V. Chubukov, npj Quantum Mater. **5**, 55 (2020).

[42] Z. M. Raines, V. I. Fal'ko, and L. I. Glazman, Phys. Rev. B **103**, 075422 (2021).

[43] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

[44] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[45] B. N. Narozhny, Phys. Rev. B **100**, 115434 (2019).

[46] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).

[47] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[48] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[49] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).

[50] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).

[51] S. Danz and B. N. Narozhny, 2D Materials **7**, 035001 (2020).

[52] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).

[53] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[54] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[55] B. N. Narozhny and I. V. Gornyi (2021), arXiv:2102.00207.

[56] O. Kashuba, B. Trauzettel, and L. W. Molenkamp, Phys. Rev. B **97**, 205129 (2018).

[57] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).

[58] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[59] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).

[60] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[61] I. L. Aleiner and B. I. Shklovskii, Phys. Rev. B **49**, 13721 (1994).

[62] M. Müller, L. Fritz, and S. Sachdev, Phys. Rev. B **78**, 115406 (2008).

[63] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **102**, 075305 (2020).

[64] A. A. Kozikov, A. K. Savchenko, B. N. Narozhny, and A. V. Shytov, Phys. Rev. B **82**, 075424 (2010).

[65] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Rev. Mod. Phys. **83**, 407 (2011).

[66] M. Schütt, P. M. Ostrovsky, M. Titov, I. V. Gornyi, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. Lett. **110**, 026601 (2013).

[67] G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).

[68] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).

[69] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).

[70] B. N. Narozhny and A. Levchenko, Rev. Mod. Phys. **88**, 025003 (2016).

[71] D. E. Sheehy and J. Schmalian, Phys. Rev. Lett. **99**, 226803 (2007).

arXiv:2110.15369v2 [cond-mat.mes-hall] 20 Nov 2022

Spread and erase – How electron hydrodynamics can eliminate the Landauer-Sharvin resistance

Ady Stern,[1] Thomas Scaffidi,[2, 3] Oren Reuven,[1] Chandan Kumar,[1] John Birkbeck,[1] and Shahal Ilani[1]

[1]*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel*
[2]*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*
[3]*Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, M5S 1A7, Canada*

It has long been realized that even a perfectly clean electronic system harbors a Landauer-Sharvin resistance, inversely proportional to the number of its conduction channels. This resistance is usually associated with voltage drops on the system's contacts to an external circuit. Recent theories have shown that hydrodynamic effects can reduce this resistance, raising the question of the lower bound of resistance of hydrodynamic electrons. Here we show that by a proper choice of device geometry, it is possible to spread the Landauer-Sharvin resistance throughout the bulk of the system, allowing its complete elimination by electron hydrodynamics. We trace the effect to the dynamics of electrons flowing in channels that terminate within the sample. For ballistic systems this termination leads to back-reflection of the electrons and creates resistance. Hydrodynamically, the scattering of these electrons off other electrons allows them to transfer to transmitted channels and avoid the resistance. Counter-intuitively, we find that in contrast to the ohmic regime, for hydrodynamic electrons the resistance of a device with a given width can decrease with its length, suggesting that a long enough device may have an arbitrarily small total resistance.

Introduction The electronic resistivity to the flow of current is a fundamental quantity in condensed matter physics. Frequently, its minimization is desired. The Drude model, dating back to 1900, suggests that the resistivity originates mostly from momentum loss to impurities. However, it was realized that even in the ballistic limit, in which impurities and phonons are absent, the interface between the electronic system and the metallic contacts to which it is coupled carries another fundamental source of resistance - the Landauer-Sharvin (LS) resistance[1–4]. This resistance is inversely proportional to the number of quantum mechanical channels that are transmitted through the system.

More recently, another regime of transport was discovered, in which electrons behave like a viscous fluid due to strong momentum-conserving electron-electron scattering. [5–50]. Somewhat counter-intuitively, it was shown that the resistance in this hydrodynamical regime may be lower than the ballistic one, suggesting the term "super-ballistic"[29, 30, 51–53]. Furthermore, conditions in which field-free current flow may locally exist were suggested[42].

In this work, using a combination of Landauer and Boltzmann analyses we demonstrate a mechanism by which electron hydrodynamics can eliminate the LS resistance, and find the minimal value that this resistance may attain. Our study is semi-classical and focuses on two dimensional systems. We describe an electronic system in terms of its conduction channels, and show that when the number of channels varies along the direction of the current flow, the Landauer-Sharvin resistance detaches from the contacts and spreads over the bulk of the electronic system. When the length scale of this spreading is larger than the electron-electron scattering mean free path, ℓ_{ee}, the resistance is dramatically suppressed.

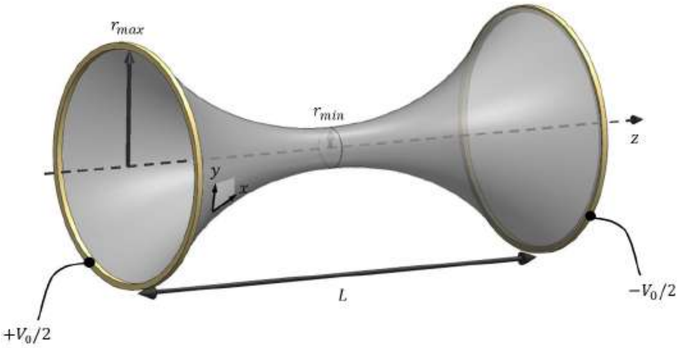


FIG. 1. The wormhole geometry is a two dimensional azimuthally-symmetric electronic system embedded in three dimensional space, described by the equation $r = r(z)$, where the radius is maximal (r_{max}) at the interface to the contacts and minimal (r_{min}) at the center. A current I is driven from negative to positive z, and the potentials at the two contacts are $V(\mp L/2) = \pm\frac{V_0}{2}$.

Microscopically, this suppression results from the scattering of electrons whose channels are being terminated due to a narrowing of the system's cross-section or a decrease of its carrier density. In a ballistic system, these electrons are reflected back and do not contribute to the current, thereby generating LS resistance in the sample's bulk. In contrast, in the hydrodynamic regime electron-electron scattering transfers these electrons into transmitted channels, thus avoiding their reflection and the corresponding resistance.

Equipped with this analysis, we can ask the question of the minimal resistance of hydrodynamic electrons flowing through a constriction. In the ballistic case, for a sample of length L and a minimal cross section $2\pi r_{min}$, the LS resistance is given by $\frac{h}{2e^2 k_F r_{min}}$($k_F$ is the Fermi momentum, and we consider a single spin species). In the hydrodynamic case, previous works[29, 30, 42] reported a reduction of the LS resistance by a factor of ℓ_{ee}/r_{min} due to electron hydrodynamics. We find a further reduction of the resistance by an additional factor of r_{min}/L if the constriction's width varies over a scale $L \gg r_{min}$. In contrast to the familiar ohmic regime, in which resistance increases with L, the resistance in the hydrodynamic regime decreases with L. This implies that a system with a given r_{min} and a large enough L may have an arbitrarily small total resistance.

Wormhole geometry In order to study the resistance of hydrodynamic electrons in a generic expanding geometry while avoiding boundary effects, we use a "wormhole" geometry (Fig. (1)). This geometry is a two-dimensional surface of revolution embedded in three dimensions, with azimuthal symmetry (toward the end of the paper we consider also a Corbino disk and a bar with varying electronic density). In cylindrical coordinates the wormhole is defined by $r = r(z)$, with its minimum radius, $r_{\min}$, occurring at $z = 0$, and maximum radius, r_{max}, occurring at the contacts positioned at $z = \pm L/2$. For simplicity we assume $r(z) = r(-z)$. A current I driven through the wormhole in the z-direction leads to a potential $V(z = \mp L/2) = \pm\frac{V_0}{2}$ at its contacts. On the manifold, we define a local Cartesian system of coordinates tangent to the manifold, in which $\hat{y}$ is the unit vector in the azimuthal direction, and $\hat{x} = \frac{1}{\sqrt{1+r'^2}}(r', 0, 1)$ is the unit vector in the direction along the manifold. For brevity, we set $\hbar = e = 1$.

Boltzmann description Time-independent transport in a wormhole geometry may be described by a Boltzmann equation. In the absence of a driving force, the magnitude of the electron's momentum is constant, but its direction varies. Consequently, the equation reads:

$$\cos\theta\partial_z f - \frac{r'}{r}\sin\theta\partial_\theta f = \sqrt{1+r'^2}I[f] \tag{1}$$

where $f(\mathbf{r}, \mathbf{p})$ is the deviation of the number of electrons in a position $\mathbf{r}$ with momentum $\mathbf{p}$ from its equilibrium value, θ is the angle of the momentum with respect to the locally defined x-direction, $r' \equiv dr/dz$, and $I[f]$ is the scattering integral, elaborated below. As explained in the Appendix, this equation is derived in two steps. First, we solve for the trajectories of free particles constrained to the manifold. Second, we equate the variation of f along these trajectories with the scattering integral $I[f]$.

It is common to substitute the ansatz

$$f(\mathbf{p}, \mathbf{r}) = \delta(\epsilon_F - \epsilon(\mathbf{p}))h(\mathbf{p}, \mathbf{r}) \tag{2}$$

in (1), and integrate both sides over the magnitude of the momentum $\int \frac{pdp}{4\pi^2}$, with $p = |\mathbf{p}|$. This integration fixes $|\mathbf{p}| = p_F$ such that h becomes a function of $\mathbf{r}$ and θ, which describes the non-equilibrium angular shape of the Fermi surface. The integration replaces the δ-function in (2) by a density of states at the Fermi energy and angle θ, $\nu(E_F, \theta) = \nu_F/2\pi$ (here ν_F is the density of states at the Fermi energy). The Boltzmann equation becomes an equation for $\nu_F h(\theta, \mathbf{r})/2\pi$. The azimuthal symmetry reduces the dependence on $\mathbf{r}$ to a dependence on z only.

Landauer description In the Landauer picture, the system is composed of $2j_{max} + 1$ channels, enumerated by their angular momentum $j = -j_{max}..0..j_{max}$, with $j_{max} = k_F r_{\max}$. The angular momentum $j = p_y(z)r(z)$, with p_y the momentum in the azimuthal direction. Each channel is characterized by transmission and reflection probabilities T_j, R_j satisfying $T_j + R_j = 1$. We assume $r(z)$ to vary slowly on the scale of the Fermi wavelength, such that in the absence of interactions, channels with $|j| < k_F r_{\min}$ are fully transmitted, and all other channels are fully reflected. The reflection takes place at the classical turning point $r(z) = |j|/k_F$. The current flowing through the wormhole is $I = \frac{k_F r_{min}}{\pi}V_0$, leading to a dimensionless LS resistance $R_{ballistic} = \pi/k_F r_{min}$.

"Landauerizing" Boltzmann We reformulate the Boltzmann equation to elucidate its relation to the Landauer picture. To that end, we express the shape of the Fermi surface in terms of different variables - the channel angular momentum j, the direction of motion, right (R) or left (L), and the position, z. Semi-classically the angular momentum is a real number, which is quantized to an integer in Landauer's quantum mechanical analysis. Here, we think about it semi-classically.

Two steps are needed to transform the Boltzmann equation from an equation for $h(\theta, z)$ to an equation for the occupation in terms of j, z and direction of motion, which we denote by $h^j_{R,L}(z)$. First, we change the variables in Eq. (1). Second, the integral $\int \frac{pdp}{4\pi^2}$ should be replaced by an integral over the x-component of the momentum, namely $p_F \int \frac{dp_x}{2\pi}$, where the limits are given by $p_x = 0$ and $p_x = \pm\infty$, for R,L respectively. The δ-function in (2) is then replaced by a density of states at the Fermi level at a fixed $j = p_y r(z)$,

$$\nu^j = \frac{\nu_F}{\sqrt{1 - \left(\frac{j}{k_F r(z)}\right)^2}}\Theta\left(k_F r(z) - |j|\right) \tag{3}$$

This density of states is inversely proportional to the x–component of the velocity, as familiar from Landauer's analysis. The details of the transformation are given in the Appendix, but the outcome is quite expected from the conservation of angular momentum:

$$\pm\,\nu_F \partial_z h^j_{R,L}(z) = \sqrt{1+r'^2}\tilde{I}[h^j_{R,L}(z)] \tag{4}$$

where the $\pm$ refers to right and left moving electrons, and $\tilde{I}$ is the scattering term expressed as a functional of $h^j_{R,L}(z)$, derived below.

The electronic density $\rho(z)$, current density $J_x(z)$ and potential $V(z)$ are,

$$\rho(z) = \int \frac{d\mathbf{p}}{4\pi^2} f(\mathbf{p},z) = \frac{1}{2\pi k_F r(z)} \int dj \nu^j(z) \left[h_R^j + h_L^j\right]$$
$$J_x(z) = \int \frac{d\mathbf{p}}{4\pi^2}\frac{p_x}{m} f(\mathbf{p},z) = \frac{1}{4\pi^2 r(z)} \int dj \left[h_R^j - h_L^j\right]$$
$$V(z) = \rho(z)/\nu_F \tag{5}$$

where the limits of integration are over all angular momenta for which $\nu^j \neq 0$, i.e. from $-j_{max}$ to $j_{max} = k_F r_{max}$.

Ballistic regime. In the absence of collisions ($\tilde{I} = 0$), Eq. (4) implies that $h^j_{R,L}$ is independent of z and is such that $h^j_R = h^j_L$ at the classical turning point, where $j = k_F r(z)$. The solution states that there is no inter-channel scattering along the wormhole, which is a consequence of angular momentum conservation. As for intra-channel back-scattering, two situations may exist: fully transmitted channels are those with $j < k_F r_{min}$. For these channels, each of the two non-equilibrium occupations $h^j_{R,L}(z)$ is determined by the contact from which it emanates, and is independent of z. In contrast, if there is a point z_0 for which $j = k_F r(z_0)$, at this point the momentum has no x-component, $h_R = h_L$ and the channel is fully reflected. Then, on one side of z_0 where the channel exists we have $h_R = h_L$, with the value being determined by the contact from which the channel emanates and to which it is back-reflected. Both occupations vanish at the other side of z_0, in which the channel does not exist. Figure (2a,b) presents $h^j_R \mp h^j_L$ for a particular example of a ballistic wormhole, showing the non-equilibrium channel-dependent contributions to the local potential and current density.

Each contact feeds into the wormhole all channels below its potential, $\pm V_0/2$ for the left and right contacts respectively, thus specifying the boundary conditions. By Landauer's formula, $V_0 = \pi I / k_F r_{min}$. With these boundary conditions, we can solve for $h^j_{R,L}(z)$ and use the solution to calculate the potential as a function of z. We find the potential to be,

$$V_{\text{ballistic}}(z) = -\text{sgn}(z)\frac{V_0}{\pi}\int_{k_F r_{min}}^{k_F r(z)} \frac{dj}{\sqrt{(k_F r(z))^2 - j^2}}$$
$$= -\text{sgn} z \frac{V_0}{2}\left[1 - \frac{2}{\pi}\arcsin\frac{r_{min}}{r(z)}\right] \tag{6}$$

Interestingly, although there are no collisions, we see that there is a potential drop, and thus resistance, in the bulk of the wormhole. Eq. (6) shows that the potential close to the edges of the wormhole ($z = \pm L/2$) is smaller than that in the contacts by $\frac{V_0}{\pi}\arcsin\frac{r_{min}}{r_{max}}$. This difference is the LS contact resistance. In the limit $r_{max} \gg r_{min}$ this contact resistance becomes negligible, and practically all the LS resistance drops in the bulk. From Eq. (6) we see that voltage drops in the bulk when the upper limit of the integral varies with z. Hence, the bulk LS resistance appears whenever the number of conduction channels varies in the bulk. As we show below, electron-electron scattering can dramatically suppress the bulk potential drop, allowing the system to conduct much better than the fundamental LS limit.

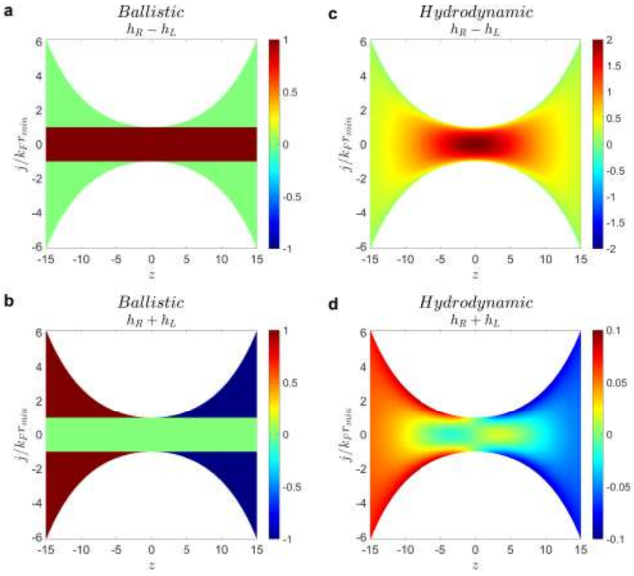


FIG. 2. Non-equilibrium distribution functions $h_R - h_L$ and $h_R + h_L$ for ballistic (a,b) and hydrodynamic (c,d) cases. These distribution functions contribute to the current density and voltage respectively (see Eq. (5)). They are plotted for the wormhole defined in (11) with $a/r_0 = 6$, as a function of the spatial coordinate z, and the normalized channel index, $j/k_F r_{min}$. In panels (c,d) $\ell_{ee}/r_0 = 0.3$. Green color corresponds to zero population, while white reflects states above the Fermi energy.

Electron-electron scattering and the hydrodynamic regime. We now turn to consider the effect of momentum conserving electron-electron interactions on the wormhole resistance. Within the relaxation time approximation, taking conservation laws into account [11, 54], we have

$$I[f(\mathbf{p},\mathbf{r})] = -\frac{1}{\ell_{ee}}\left[f - \frac{\delta(\epsilon_F - \epsilon(\mathbf{p}))}{\nu_F}\left(\rho(\mathbf{r}) + \frac{2J_x(\mathbf{r})\cos\theta}{v_F}\right)\right] \tag{7}$$

The second and third terms on the right hand side guarantee charge and momentum conservation, respectively. We obtain $\tilde{I}[h^j]$ using the same ansatz we used before,

$$\tilde{I}[h^j_{R,L}(z)] =$$
$$-\frac{\nu^j}{\ell_{ee}}\left[h^j_{R,L}(z) - \frac{\rho(z)}{\nu_F} \mp \frac{4\pi J_x(z)}{k_F}\sqrt{1 - \left(\frac{j}{k_F r(z)}\right)^2}\right] \tag{8}$$

The ν^j/ℓ_{ee} factor makes the mean free path j-dependent and shortens it from ℓ_{ee} to $\ell_{ee}\sqrt{1 - \left(\frac{j}{k_F r(z)}\right)^2}$. This may be understood by noting that for $j/r(z)$ large, p_x is small and a shorter distance is traversed in the z-direction between two scattering events. In particular, the scattering length vanishes when the channel is about to be terminated, opening a way for the electrons to avoid back-scattering by being scattered to a transmitted channel. Furthermore, in contrast to the case of impurity scattering, in which in Eq. (8) ℓ_{ee} is replaced by a momentum-relaxing mean free path and the third term is absent, here the presence of the third term allows for a Galilean boost of the Fermi surface $h^j{}_{R,L}(z) = \pm\frac{4\pi J_x(z)}{k_F}\sqrt{1-(j/k_F r(z))^2}$ to be carried out without developing a resistance.

We find the solution to a leading order in ℓ_{ee} (the calculation is given in the Appendix),

$$\begin{aligned} h^j_{R,L}(z) = &\pm \frac{2I}{k_F r(z)}\sqrt{1-\left(\frac{j}{k_F r(z)}\right)^2} \\ &+ \frac{2I\ell_{ee}\sin\xi(z)}{k_F r^2(z)}[1-2(\frac{j}{k_F r(z)})^2] \\ &- \int_0^z \frac{I\ell_{ee}}{k_F r^2(z')}\cos\xi(z')\frac{d\xi}{dz'}(z')dz' \end{aligned} \tag{9}$$

where $\xi(z)$ is the local angle between z-axis and the manifold, i.e. $r(z)' = \tan\xi(z)$. This solution is valid in the bulk, away from the contacts. We comment on the role of the contacts below, with details in the Appendix.

The first term in Eq. (9) is a rigidly shifted Fermi surface. It is the solution expected for $r' = 0$ far away from the contacts, after all deformations of the Fermi surface are suppressed by the scattering term. The second and third terms are smaller than the first by a factor of $\ell_{ee}/r(z)$, and originate from the breaking of Galilean invariance. The second term makes the shifted Fermi surface acquire an elongated shape, with more electrons in the head-on direction (small j), and less in the $j \approx k_F r(z)$ channels. The third term is independent of j. It carries an electronic density, and leads to a potential drop and resistivity. Note that while the second term exists when $\sin\xi \neq 0$, the third term requires $\frac{d\sin\xi}{dz} \neq 0$. Stated differently, in contrast to ballistic electrons for which local resistance appears when the number of conduction channels varies with z, i.e., when $r' \neq 0$, for hydrodynamic electrons resistance is generated only when this function has a non-zero curvature, $r'' \neq 0$.

The potential originating from the third term of Eq. (9) may be written also as:

$$V_{\text{hydro}}(z) = I\int_0^{\xi(z)} \frac{\ell_{ee}}{4\pi k_F r^2(\xi)}\cos\xi d\xi \tag{10}$$

The resistance scale may be estimated from Eq. (10). The r^2 in the denominator suggests that the wormhole resistance is characterized by a “super-ballistic” scale [29, 30, 42, 50], $\frac{2\pi\ell_{ee}}{k_F r^2_{min}}$, smaller by $2\ell_{ee}/r_{min}$ than the ballistic LS resistance. However, Eq. (10) opens the way for a much smaller scale, $\frac{\ell_{ee}}{4\pi k_F r^2_{min}}\sin\xi_0$, where ξ_0 is the angle at which r becomes much larger than r_{min}. If r grows slowly, $\sin\xi_0$ may be much smaller than one, with the resistance becoming much smaller than the super-ballistic scale. Consequently, for a fixed $r_{max} \gg r_{min}$ the resistance generally decreases with increasing L, opposite to the familiar Ohmic dependence.

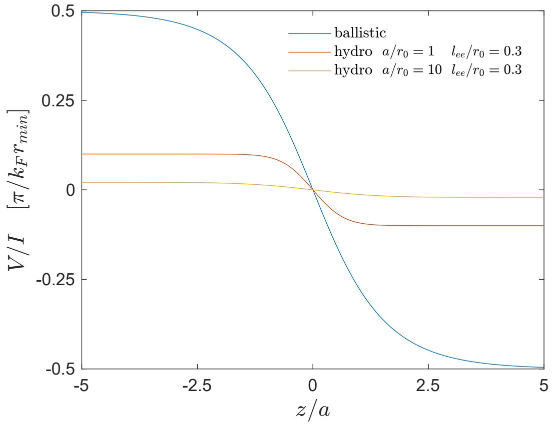


FIG. 3. The potential along the wormhole defined in (11), divided by the current, V/I, in units of the LS resistance, plotted for a ballistic flow ($\ell_{ee} = \infty$) and hydrodynamic flows ($\ell_{ee}/r_{min} = 0.3$) with varying values of a/r_0 (see legend).

To illustrate the two hydrodynamic scales, consider an example where

$$r(z) = r_0\cosh z/a \tag{11}$$

In this wormhole $r_{min} = r_0$ and $r_{max} \gg r_{min}$ for $L \gg a$. Under the latter condition, the contribution to the resistance decays fast with $|z| \gg a$, and we can take $L \to \infty$. Then, using Eq. (10),

$$R_{\cosh} = \frac{\ell_{ee}}{2\pi k_F}\left[\frac{1}{r_0^2 - a^2} - \frac{a^2 \operatorname{arctanh}\frac{\sqrt{r_0^2-a^2}}{r_0}}{r_0(r_0^2-a^2)^{3/2}}\right] \tag{12}$$

In the limit $a \to 0$ the resistance tends to $\frac{\ell_{ee}}{2\pi k_F r_0^2}$, but when $a \gg r_0$ it decreases to become of order $\frac{\ell_{ee}}{4k_F r_0 a}$. As can be seen in Eq. (10), most of the resistance originates from the product of the minimum radius r_{min} and the change in angle $\Delta\xi$ over which the radius becomes significantly larger than r_0. When $a \gg r_0$ the change in angle is $\Delta\xi \sim r_0/a$ and hence the decrease in resistance. Fig. (2c,d) show the calculated $h_L - h_R$ and $h_L + h_R$ for hydrodynamic flow in the wormhole in Eq. (11). These quantities contribute to the current density and potential, respectively (Eq. (5)). Fig. (3) shows potential drop in this wormhole as a function of z, in the ballistic case and in the hydrodynamic cases for two values of a/r_0. The hydrodynamic suppression of the resistance with increasing a is evident. Note that when $r(z)$ is constant the resistance in the bulk vanishes, since the bulk is

Galilean invariant. However, the LS voltage drop occurs then sharply at the contacts, and is not suppressed by electron-electron scattering. To suppress the resistance by electron-electron scattering $r(z)$ should vary slowly from r_{min} to $r_{max} \gg r_{min}$.

Our analysis elucidates this suppression of the resistance: a potential drop results from reflection of electrons. In the ballistic regime the contact sends into the sample electrons in channels for which j is too large to be transmitted. Those electrons are reflected, leading to a voltage drop (Eq. (6)). In contrast, in the hydrodynamic regime electrons of high j are scattered to channels of smaller j, and largely end up being transmitted, without generating a potential drop. Note that our entire analysis assumes $\ell_{ee} \ll a$ and $\ell_{ee} \ll r_0$, in contrast to the sharp constriction case, studied, e.g., in 29, leading to a rather different evolution of R with ℓ_{ee}.

Eqs. (10) and (11) show that the bulk resistance of a Corbino disk vanishes, as a consequence of the lack of variation of ξ. With the limitation of z to a proper range, and with the limit $a \to 0$, Eq. (11) may be used to describe a Corbino disk. The resulting bulk resistance vanishes in that limit. Indeed, in a Corbino disk the number of channels grows linearly with the radial coordinate, its second derivative vanishes, and so does the hydrodynamic resistance. Importantly, this vanishing bulk resistance is in series with a contact resistance which in this case is $\pi/(2k_F r_{min})$, where r_{min} is the inner radius of the disk.

The elimination of the LS resistance in a Corbino disk was experimentally confirmed, as reported in a companion article[55]. In that article, we generalized the present calculations to include momentum relaxation due to phonon and impurity scattering, and showed that it leads to a simple additive contribution to the resistance.

Finally, although the wormhole is illuminating theoretically, it is a rather exotic geometry for real-life transistor devices. Those typically have a long rectangular bar geometry, in which the density varies along the x-axis and is maximal near the contacts. In a bar geometry, previous work (e.g. [27]) has focused on a viscous contribution arising from the no-slip boundary condition. Here, we neglect this contribution by assuming specular boundary scattering, or a wide bar. By carrying out an analysis similar to that of the wormhole (see Appendix), we find the resistance

$$R_{bar} = \int_{-\infty}^{\infty} dx \frac{(k_F' \ell_{ee})'}{2k_F^2 r}. \qquad (13)$$

Here we account also for the possibility that ℓ_{ee} varies with the variation of k_F. Assuming that the change in k_F is much larger than its minimnal value, we can estimate $R \sim \ell_{ee}/k_F r a$, where a is the scale over which k_F and ℓ_{ee} become much larger than their minimal value.

In summary, we showed here that when the LS resistance of an electronic system is spread into its bulk, rather than being localized at the interface with the contacts, it may be significantly reduced by electron-electron scattering, in principle all the way down to zero.

We thank L. Ella, G. Falkovich, L. Levitov, M. Polini, M. Shavit, A. Rozen, A. V. Shytov and U. Zondiner for useful discussions. Work was supported by the Leona M. and Harry B. Helmsley Charitable Trust grant, ISF grant (No. 1182/21), Minerva grant (No. 713237), Hydrotronics (No. 873028) and the ERC-Cog (See-1D-Qmatter, No. 647413). T. S. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), in particular the Discovery Grant (No. RGPIN-2020-05842), the Accelerator Supplement (No. RGPAS-2020-00060) and the Discovery Launch Supplement (No. DGECR-2020-00222). This research was enabled in part by support provided by Compute Canada. A. S. was supported by the ERC under the Horizon 2020 Research and Innovation programme (LEGOTOP No. 788715), the DFG (CRC/Transregio 183, EI 519/7-1), and the ISF Quantum Science and Technology (2074/19).

---

[1] R. Landauer, IBM Journal of Research and Development **1**, 223 (1957).

[2] Y. V. Sharvin, Soviet Journal of Experimental and Theoretical Physics **21**, 655 (1965).

[3] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge Studies in Semiconductor Physics and Microelectronic Engineering (Cambridge University Press, 1995).

[4] Y. Imry, *Introduction to mesoscopic physics* (Oxford University Press on Demand, 2002).

[5] R. Gurzhi, Sov. Phys. JETP **44**, 771 (1963).

[6] R. Gurzhi, Physics-Uspekhi **11**, 255 (1968).

[7] J. E. Black, Phys. Rev. B **21**, 3279 (1980).

[8] Z. Z. Yu, M. Haerle, J. W. Zwart, J. Bass, W. P. Pratt, and P. A. Schroeder, Phys. Rev. Lett. **52**, 368 (1984).

[9] R. Gurzhi, A. Kalinenko, and A. Kopeliovich, Physical review letters **74**, 3872 (1995).

[10] L. W. Molenkamp and M. J. M. de Jong, Phys. Rev. B **49**, 5038 (1994).

[11] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[12] M. Dyakonov and M. Shur, Phys. Rev. Lett. **71**, 2465 (1993).

[13] E. Chow, H. P. Wei, S. M. Girvin, and M. Shayegan, Phys. Rev. Lett. **77**, 1143 (1996).

[14] B. Spivak and S. A. Kivelson, Annals of Physics **321**, 2071 (2006).

[15] A. V. Andreev, S. A. Kivelson, and B. Spivak, Phys. Rev. Lett. **106**, 256804 (2011).

[16] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys. Rev. B **92**, 165433 (2015).

[17] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).

[18] A. Tomadin, G. Vignale, and M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).

[19] L. Levitov and G. Falkovich, Nat Phys **12**, 672 (2016).

[20] A. Principi, G. Vignale, M. Carrega, and M. Polini, Phys. Rev. B **93**, 125410 (2016).

[21] M. Sherafati, A. Principi, and G. Vignale, Phys. Rev. B **94**, 125427 (2016).

[22] H. Guo, E. Ilseven, G. Falkovich, and L. Levitov, ArXiv e-prints (2016), arXiv:1612.09239 [cond-mat.mes-hall].

[23] A. Lucas, Phys. Rev. B **95**, 115425 (2017).

[24] A. Levchenko, H.-Y. Xie, and A. V. Andreev, Phys. Rev. B **95**, 121301 (2017).

[25] D. A. Bandurin, I. Torre, R. K. Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, L. A. Ponomarenko, A. K. Geim, and M. Polini, Science **351**, 1055 (2016).

[26] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, Science **351**, 1058 (2016).

[27] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[28] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[29] H. Guo, E. Ilseven, G. Falkovich, and L. S. Levitov, Proceedings of the National Academy of Sciences **114**, 3068 (2017).

[30] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, and A. K. Geim, Nature Physics **13**, 1182 (2017).

[31] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[32] A. Lucas and S. A. Hartnoll, Phys. Rev. B **97**, 045105 (2018).

[33] D. Y. H. Ho, I. Yudhistira, N. Chakraborty, and S. Adam, Phys. Rev. B **97**, 121404 (2018).

[34] A. Shytov, J. F. Kong, G. Falkovich, and L. Levitov, Phys. Rev. Lett. **121**, 176805 (2018).

[35] P. S. Alekseev and M. A. Semina, Phys. Rev. B **98**, 165412 (2018).

[36] D. Svintsov, Phys. Rev. B **97**, 121405 (2018).

[37] J. Gooth, F. Menges, N. Kumar, V. Süss, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[38] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, M. Polini, T. Ihn, and K. Ensslin, Phys. Rev. B **98**, 241304 (2018).

[39] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin, Science **364**, 162 (2019), arXiv:1806.01606 [cond-mat.mes-hall].

[40] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A. K. Geim, and S. Ilani, Nature **576**, 75 (2019).

[41] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[42] M. Shavit, A. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).

[43] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M. M. Fogler, P. Kim, A. Yacoby, and R. L. Walsworth, Nature **583**, 537 (2020).

[44] A. Levchenko and J. Schmalian, Annals of Physics **419**, 168218 (2020).

[45] T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, Phys. Rev. B **100**, 245305 (2019).

[46] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, K. Watanabe, T. Taniguchi, A. Lucas, A. F. Young, and A. C. Bleszynski Jayich, arXiv e-prints , arXiv:2002.05065 (2020), arXiv:2002.05065 [cond-mat.mes-hall].

[47] A. C. Keser, D. Q. Wang, O. Klochan, D. Y. H. Ho, O. A. Tkachenko, V. A. Tkachenko, D. Culcer, S. Adam, I. Farrer, D. A. Ritchie, O. P. Sushkov, and A. R. Hamilton, Phys. Rev. X **11**, 031030 (2021).

[48] A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021).

[49] Z. J. Krebs, W. A. Behn, S. Li, K. J. Smith, K. Watanabe, T. Taniguchi, A. Levchenko, and V. W. Brar, arXiv e-prints , arXiv:2106.07212 (2021), arXiv:2106.07212 [cond-mat.mes-hall].

[50] Q. Hong, M. Davydova, P. J. Ledwith, and L. Levitov, arXiv e-prints , arXiv:2012.03840 (2020), arXiv:2012.03840 [cond-mat.mes-hall].

[51] K. E. Nagaev and O. S. Ayvazyan, Phys. Rev. Lett. **101**, 216807 (2008).

[52] K. E. Nagaev and T. V. Kostyuchenko, Phys. Rev. B **81**, 125316 (2010).

[53] M. Y. Melnikov, J. P. Kotthaus, V. Pellegrini, L. Sorba, G. Biasiol, and V. S. Khrapai, Phys. Rev. B **86**, 075425 (2012).

[54] J. Callaway, Phys. Rev. **113**, 1046 (1959).

[55] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, and S. Ilani, Nature **609**, 276 (2022).

Supplemental material to "*Spread and erase – How electron hydrodynamics can eliminate the Landauer-Sharvin resistance*"

In this supplemental material, we give (A) the derivation of the Boltzmann equation in the wormhole geometry, (B) the solution of the Boltzmann equation in the hydrodynamic regime for a wormhole, (C) the solution of the Boltzmann equation in the hydrodynamic regime for a bar with a density variation, and (D) a discussion of the contact resistance.

Appendix A: Derivation of Boltzmann equation for the wormhole

The wormhole is a surface of revolution defined, in a cylindrical system of coordinates, by $r(z)$. In Cartesian components it is defined $(r(z)\cos(\phi), r(z)\sin(\phi), z)$. The kinetic energy of a particle (which is also the Lagrangian) is given by

$$|v_F|^2 = (r\dot{\phi})^2 + \dot{z}^2(1+r'^2) \tag{A1}$$

and is conserved. Since the norm of the velocity is conserved, we only need to keep track of its angle θ, measured with respect to an arbitrary axis. We define θ by :

$$\begin{aligned} r\dot{\phi} &= v_F\sin(\theta) \\ \sqrt{1+r'^2}\dot{z} &= v_F\cos(\theta) \end{aligned} \tag{A2}$$

The equations of motion are given by:

$$\begin{aligned} \ddot{\phi} &= -2\frac{r'}{r}\dot{z}\dot{\phi} \\ \ddot{z} &= \frac{rr'}{1+r'^2}\dot{\phi}^2 - \frac{r'r''}{1+r'^2}\dot{z}^2 \end{aligned} \tag{A3}$$

We start from the initial 2D problem with 2 coordinates and 2 velocities. Time independent distribution of electrons will necessarily be rotational invariant, and hence independent of ϕ. Furthermore, in the absence of a driving force the magnitudes of the momentum and the velocity are conserved, such that their dynamical variable is the he angle θ, measured with respect to the x-axis. The Boltzmann equation is then an equation for two coordinates - the radial coordinate r and the direction of the velocity θ. The equations of motion for r, θ:

$$\begin{aligned} \dot{r} &= r'\dot{z} = \frac{r'}{\sqrt{1+r'^2}}v_F\cos(\theta) \\ \dot{\theta} &= -\frac{r'}{r}\frac{1}{\sqrt{1+r'^2}}v_F\sin(\theta) \end{aligned} \tag{A4}$$

We can now write the Boltzmann equation for $f(r,\theta)$ as:

$$\frac{df}{dt} = \partial_r(f)\dot{r} + \partial_\theta(f)\dot{\theta} = v_F I[f] \tag{A5}$$

with $I[f]$ the scattering integral given in the main text. This leads to

$$\partial_r(f)\frac{r'}{\sqrt{1+r'^2}}\cos(\theta) - \partial_\theta(f)\frac{r'}{r}\frac{1}{\sqrt{1+r'^2}}\sin(\theta) = I[f] \tag{A6}$$

This can be rewritten as

$$\partial_r(f)\cos(\theta) - \partial_\theta(f)\frac{1}{r}\sin(\theta) = \frac{\sqrt{1+r'^2}}{r'}I[f] \tag{A7}$$

The left hand side is independent of z, which means in the ballistic regime, all surfaces are equivalent once expressed in terms of r. In the non-ballistic case, the only difference between surfaces is that the scattering rate acquires a dependence on r' through the factor $\sqrt{1+r'^2}/r'$. Multiplying Eq. (A7) by r', and using $r'\partial_r = \partial_z$ we obtain Eq. (1).

Appendix B: Wormhole and Corbino disk - solution of the Boltzmann equation in the hydrodynamic regime

The equation describing the non-equilibrium current distribution is obtained from Eq. (A7) by setting $f(\mathbf{p},z) = \delta(\epsilon_F - \epsilon(\mathbf{p}))h^y(p_y)$ and integrating $p_F \int \frac{dp_x}{2\pi}$. We get,

$$\pm\left[\nu_F \partial_z {h^y}_{R,L} - \nu_F \frac{r'}{r} p_y \partial_{p_y} {h^y}_{R,L}\right] = \sqrt{1+r'^2}\tilde{I}[h^y] \tag{B1}$$

with,

$$\tilde{I}[h^y_{R,L}(\mathbf{p},\mathbf{r})] = -\frac{\nu^y}{\ell_{ee}}\left[{h^y}_{R,L} - \frac{\rho(\mathbf{r})}{2\nu_F} \mp \frac{4\pi j_x(\mathbf{r})}{k_F}\sqrt{1-(p_y/k_F)^2}\right] \tag{B2}$$

The transformation of this equation to an equation for h^j is explained below.

1. Deriving the Boltzmann equation for h^j

To get the Boltzmann equation that appears in Eq. (4) of the main text, we start from Eq. (B1) and change the variables $z, p_y \to z, j$. Since $j = p_y r(z)$ we must treat the dependence between the coordinates carefully. It is easy to confirm that the second term in the LHS of (B1) becomes $-j\frac{r'}{r}\partial_j h_{R,L}$, however, the first term also changes because it is a total derivative, $\partial_z h_{R,L} \to \partial_z h_{R,L} + \frac{\partial j}{\partial z}\partial_j h_{R,L}$. Using $\frac{\partial j}{\partial z} = p_y r'$ we get that

$$\frac{\partial j}{\partial z}\partial_j h_{R,L} - j\frac{r'}{r}\partial_j h_{R,L} = 0,$$

thus obtaining Eq. (4).

2. Solving the Boltzmann equation to linear order in ℓ_{ee}

In this subsection we solve the equation for h^j up to first order in $\ell_{ee}r'/r$.

Assuming that $r(z)$ varies slowly on a scale of ℓ_{ee} we try as a first attempt the z-independent solution of a uniform r, adjusted at each point to the local $r(z)$, i.e., $\pm\frac{2I}{k_F r(z)}\sqrt{1-\left(\frac{j}{k_F r(z)}\right)^2}$. This attempted solution conserves current. Substituting it into Eq. (4) we find on the LHS a remainder term $-\frac{2I}{k_F}\frac{r'\left(1-2\left(\frac{j}{k_F r}\right)^2\right)}{r^2\sqrt{1-\left(\frac{j}{k_F r(z)}\right)^2}}$. It is linear in r', as expected. To compensate for this term, we modify our solution, making it

$$\begin{aligned} h^j_{R,L}(z) &= \pm\frac{2I}{k_F r(z)}\sqrt{1-\left(\frac{j}{k_F r(z)}\right)^2} + \frac{2I\ell_{ee}r'}{k_F r^2(z)\sqrt{1+r'^2}}\left[1-2\left(\frac{j}{k_F r(z)}\right)^2\right] \\ &= \pm\frac{2I}{k_F r(z)}\sqrt{1-\left(\frac{j}{k_F r(z)}\right)^2} + \frac{2I\ell_{ee}\sin\xi(z)}{k_F r^2(z)}\left[1-2\left(\frac{j}{k_F r(z)}\right)^2\right] \end{aligned} \tag{B3}$$

The term we added to h is linear in ℓ_{ee}, such that when substituted into the RHS, it balances the remainder term on the LHS, which is ℓ_{ee}-independent. However, it generates a new remainder term on the LHS, which is linear in ℓ_{ee}. Specifically, this term is,

$$\pm\frac{2I\ell_{ee}}{k_F}\frac{r\xi'\cos\xi\left(1-2\left(\frac{j}{k_F r}\right)^2\right) + 2r'\sin\xi\left(4\left(\frac{j}{k_F r}\right)^2 - 1\right)}{r^3} \tag{B4}$$

Naively, we should balance this term by adding a term $\delta h \propto \ell_{ee}^2$ to our solution, thereby generating a term $\delta h \nu^j/\ell_{ee}$ on the RHS to cancel the contribution (B4) on the LHS. We should note, however, that the RHS cannot cancel parts of δh that carry a current or a density. Furthermore, we cannot add to h a term that carries current, because we assume a fixed driven current. As it turns out, the second term in (B4) can be cancelled by the RHS, but the first term requires more care. We write it as

$$\pm\frac{I\ell_{ee}}{k_F r^2}\xi'\cos\xi\pm\frac{I\ell_{ee}}{k_F r^2}\xi'\cos\xi\left(1-4\left(\frac{j}{k_F r}\right)^2\right) \tag{B5}$$

and cancel the first, j-independent, term by subtracting a j-independent, density carrying, term,

$$\int dz'\frac{I\ell_{ee}}{k_F r^2}\xi'\cos\xi \tag{B6}$$

When divided by ν_F, this term gives the local electrochemical potential.

Appendix C: Bar with a density variation - solution of the Boltzmann equation for the hydrodynamic regime

We consider an infinite bar parallel to the z-axis, in which the equilibrium density, and hence k_F vary with z. Furthermore, with a variation of density comes also a variation of ℓ_{ee}. When the walls of the bar are specular, we can view it as a cylinder, and we denote the circumference by $2\pi r$. With these assumptions, Boltzmann equation (4) and the collision term (8) remain the same as they were for a wormhole, with $r'=0$:

$$\pm\frac{\nu_F}{\nu_j}\partial_z h_{R,L}=-\frac{1}{\ell_{ee}}\left[h_{R,L}-\frac{\rho}{2\nu_F}\mp\frac{4\pi j_x}{k_F}\sqrt{1-\left(\frac{j}{k_F r}\right)^2}\right] \tag{C1}$$

For the limit of small ℓ_{ee}, we first try as a naive ℓ_{ee}-independent solution the locally Galilean boosted Fermi sphere: $h_{R,L}=\pm\frac{2I}{k_F r}\sqrt{1-\left(\frac{j}{k_F r}\right)^2}$, and we aim to find all corrections of order ℓ_{ee} to this term. Any amendment we do to the naive solution should not carry current, since the current is fixed to I. When the naive solution is substituted in (C1) the right hand side vanishes since $I=2\pi r j_x$ and $\rho=0$. However we get an extra term of $-\frac{2Ik'_F}{k_F^2 r}\left(1-2\left(\frac{j}{k_F r}\right)^2\right)$ on the LHS. We can balance this term by adding to our solution $\frac{2I\ell_{ee}k'_F}{k_F^2 r}\left(1-2\left(\frac{j}{k_F r}\right)^2\right)$.

This amended solution solves (C1), up to a remainder term on the LHS:

$$\sqrt{1-\left(\frac{j}{k_F r}\right)^2}\,\partial_z\left[\frac{2I\ell_{ee}k'_F}{k_F^2 r}\left(1-2\left(\frac{j}{k_F r}\right)^2\right)\right] \tag{C2}$$

This remainder term is of the order ℓ_{ee} and we need to amend our solution further to eliminate it. In principle, there are two ways to do that. The part of (C2) that does not carry density or current can be eliminated by an addition of a term of order ℓ_{ee}^2 to h. Such a term will yield an order ℓ_{ee} term on the RHS. However, being a contribution to h that is of order ℓ_{ee}^2, it is beyond our scope. The part of (C2) that carries density of current, on the other hand, should be canceled by adding a term of order ℓ_{ee} to h, that is purely a density term. Such a term will not affect the RHS, and its substitution in the LHS will cancel the terms in (C2). It is this term we are after. An inspection of (C1) and (C2) allows us to find it and write the full expression of h to linear order in ℓ_{ee} as

$$\begin{aligned}h^j_{R,L}=&\pm\frac{2I}{k_F r}\sqrt{1-\left(\frac{j}{k_F r}\right)^2}\\&+\frac{2I\ell_{ee}}{k_F^2 r}k'_F\left(1-2\left(\frac{j}{k_F r}\right)^2\right)\\&-I\int^z d\tilde{z}\frac{(k'_F\ell_{ee})'}{2k_F^2 r}\end{aligned} \tag{C3}$$

where in the integrand in the last term k_F,ℓ_{ee} are both functions of $\tilde{z}$.

Appendix D: Contact resistance

At the contact the density variation is fast. As a consequence the full solution of the Boltzmann equation becomes hard to obtain, but we can still estimate the voltage drop on the contact region. For the simplest case of a ballistic cylinder ($\xi = 0$), or a ballistic rectangular-shaped conductor with specular walls, the entire potential drop is on the two contacts. By symmetry, at the center of the wormhole $h^j_R = -h^j_L = \frac{I}{2k_F r}$ and consequently $V = 0$. In fact, these values of h^j_R, h^j_L hold anywhere within the sample, at $|z| < L/2$. In the contacts ($z = \pm(L/2 + \epsilon)$) themselves $h^j_R = h^j_L$, and their value is determined by the local potential. Thus, there is a jump in the value of h^j_L at $z = -L/2$ and of h^j_R at $z = L/2$, and this jump leads to the expected jump of the potentials at the contacts.

Next, we think of the cylindrical geometry with electron-electron scattering. Far from the contacts (a distance much larger than ℓ_{ee}) our main-text analysis holds, leading to $h^j_R = -h^j_L = \frac{4\pi J_x}{k_F}\sqrt{1-(j/k_F r)^2}$ and $V = 0$. At the two contacts $h^j_R = h^j_L = \pm V(I)/2$, and it is $V(I)$ that we estimate now. For clarity we focus on the left contact, at $z = -L/2$.

The equations satisfied by $h^j_{R,L}$ are of the form

$$\pm\nu_F \partial_z h^j_{R,L}(j,z) = -\frac{\nu^j}{\ell_{ee}}\left[h^j{}_{R,L} - \frac{\rho(z)}{\nu_F} \mp \frac{4\pi J_x}{k_F}\sqrt{1-\left(\frac{j}{k_F r}\right)^2}\right] \tag{D1}$$

These equations are equivalent to,

$$h^j_{R,L} = h^j_{R,L}(\pm L/2)e^{\frac{\pm(z\pm L/2)\nu_j}{\nu_F \ell_{ee}}} + \int_{z_i}^{z} dz' \frac{\nu_j}{\nu_F \ell_{ee}}\left[-\frac{\rho(z')}{\nu_F} \mp \frac{4\pi J_x}{k_F}\right] e^{\frac{\pm(z\pm z')\nu_j}{\nu_F \ell_{ee}}} \tag{D2}$$

Here, $h^j_{R,L}(\pm L/2)$ are the boundary conditions for the right- and left- moving electrons at the points where they enter the sample. The dependence of the distribution functions $h^j_{R,L}$ on z near the left contact is very different for the left- and right- moving electrons. For the left-moving electrons the entry point is very far from the contact we look at, and therefore the initial condition is long forgotten. In the bulk, the distribution of the left moving electrons does not vary in space, and since the second term in Eq. (D2) averages over a scale of ℓ_{ee}, we expect the variation of h^j_L to be slow even close to the contact.

For the right-moving electrons, in contrast, near the contact the solution is dominated by initial conditions. The distribution function, that starts as a constant $h^j_R = \pi^2 J_x/k_F$, decays into $h^j_R = \frac{4\pi J_x}{k_F}\sqrt{1-(j/k_F r)^2}$, at a distance ℓ_{ee}/ν_j from the contact.

Motivated by these considerations, we make the ansatz,

$$\begin{aligned} h^j_L &\approx -\frac{4\pi J_x}{k_F}\sqrt{1-(j/k_F r)^2} \\ h^j_R &= \frac{\pi^2 J_x}{k_F} e^{-(z+L/2)\nu^j/\nu_F \ell_{ee}} + (1 - e^{-(z+L/2)\nu^j/\ell_{ee}\nu_F})\frac{4\pi J_x}{k_F}\sqrt{1-(j/k_F r)^2} \end{aligned} \tag{D3}$$

Within this ansatz, the voltage difference between the contact itself and the bulk is the same as it is in the ballistic case, namely half of the Landauer-Sharvin voltage drop falls on each contact. The effect of the scattering term is limited to distributing this voltage drop from being at the interface itself to being spread on a scale of ℓ_{ee}.

arXiv:2007.00365v2 [cond-mat.str-el] 18 Sep 2020

Non-local hydrodynamic transport and collective excitations in Dirac fluids

Egor I. Kiselev[1] and Jörg Schmalian[1, 2]
[1]*Institut für Theorie der Kondensierten Materie,*
Karlsruher Institut für Technologie, 76131 Karlsruhe, Germany
[2]*Institut für Quantenmaterialien und -technologien,*
Karlsruher Institut für Technologie, 76131 Karlsruhe, Germany

We study the response of a Dirac fluid to electric fields and thermal gradients at finite wavenumbers and frequencies in the hydrodynamic regime. We find that non-local transport in the hydrodynamic regime is governed by infinite set of kinetic modes that describe non-collinear scattering events in different angular harmonic channels. The scattering rates of these modes τ_m^{-1} increase as $|m|$, where m labels the angular harmonics. In an earlier publication, we pointed out that this dependence leads to anomalous, Lévy-flight-like phase space diffusion [1]. Here, we show how this surprisingly simple, non-analytic dependence allows us to obtain exact expressions for the non-local charge and electronic thermal conductivities. The peculiar dependence of the scattering rates on m also leads to a non-trivial structure of collective excitations: Besides the well known plasmon, second sound and diffusive modes, we find non-degenerate damped modes corresponding to excitations of higher angular harmonics. We use these results to investigate the transport of a Dirac fluid through Poiseuille-type geometries of different widths, and to study the response to surface acoustic waves in graphene-piezoelectric devices.

I. INTRODUCTION

In many instances transport properties can be described in terms of a local relationship between forces and currents. Examples are Fourier's law of heat conduction $\boldsymbol{j}_\varepsilon = -\kappa\nabla T$, Fick's law of diffusion $\boldsymbol{j}_c = -D\nabla\mu$, or Ohm's law of electrical conduction $\boldsymbol{j}_c = \sigma\boldsymbol{E}$. Here the thermal conductivity κ, the diffusion coefficient D, or the electrical conductivity σ establish a relationship between the value of the forces, such as a temperature gradient or electric field, and the corresponding current density at the same location. Such local relations break down when the electron propagation is almost ballistic. Important examples worked out in particular by Brian Pippard are the nonlocal current-field relations to describe the Meissner effect in clean superconductors or the anomalous skin effect in clean metals[2–4]. However, non-local transport relations are not limited to the ballistic transport regime. Another example for non-local transport occurs when hydrodynamic flow of charge or heat sets in. Indeed, hydrodynamic flow patterns are frequently identified by complex "non-local" flow lines. It is therefore necessary to find closed expressions for the nonlocal heat conductivity $\kappa_{\alpha\beta}\left(\mathbf{r}-\mathbf{r}', t-t'\right)$, electrical conductivity $\sigma_{\alpha\beta}\left(\mathbf{r}-\mathbf{r}', t-t'\right)$ or even non-local shear viscosities $\eta(\mathbf{r}-\mathbf{r}', t-t')$ of many-body systems in the hydrodynamic regime. In this regime collisions between partices are not weak, it merely holds that momentum relaxing collisions are weak while momentum-conserving collisions are not. A formulation in terms of non-local transport coefficients allow for a microscopic description of hydrodynamic flow pattern and goes beyond the usual description in terms of the linear Navier-Stokes equation. The latter corresponds to the leading gradient expansion of the theory. In addition, the inclusion of dynamical phenomena - here expressed in terms of the dependency on the time difference $t-t'$ between force and current – allows to determine the system's collective modes.

In this paper we develop the theory of non-local transport in Dirac systems at charge neutrality in the collision-dominated hydrodynamic regime and find closed expressions for the frequency and wave-vector-dependent, charge and electronic thermal conductivities as well as the non-local viscosity. Remarkably, the calculations of this paper are exact in the limit of a small graphene fine structure constant α in the regime of linear response. This is made possible by the peculiar $\propto |m|$ dependence of the scattering rates of collinear zero modes in higher angular momentum channels $m > 2$ - a behavior that was shown to lead to a super-diffusive Lévy-flight-like phase space dynamics in an earlier work [1]. Collinear zero modes do not decay due to the strong collinear scattering that give rise to rapid equilibration and therefore dominate the long-time dynamics. We make specific predictions for measurements such as the velocity shift of surface acoustic waves, determine the flow of charge and heat in finite geometries, and determine the collective mode spectrum of the system including plasma waves and second-sound-ike thermal waves. The dispersion relations of collective modes can be derived from the poles of transport coefficients, or found from the solutions of the homogeneous quantum Boltzmann equation. Here, focusing on the charge neutrality point, we go beyond the phenomenological treatment of electron-electron interactions of Refs. [5, 6]. Our detailed analysis reveals a complex structure of damped collective excitations. These excitations are similar to the so-called "non-hydrodynamic" modes that were shown to be relevant for the equilibration of unitary fermi gases [7] and QCD plasmas [8–10]. In fact, the term non-hydrodynamic is somewhat misleading. What is meant is that these modes correspond to excitations of high angular momentum components of the kinetic distribution function, which are not captured by the Navier-Stokes equations.

Transport in a Dirac fluid is in many respects different from the archetypical example of the Fermi liquid. One important difference is that electric currents in a Dirac fluid are not protected by momentum conservation, and therefore decay even in a perfectly clean system. Negatively charged electrons and positive holes flowing in opposite directions sum up to a finite electric current with zero momentum. Thus, even in the absence of impurities, pristine graphene – the prime example of a Dirac fluid – has a finite conductivity that is induced by electron-electron interactions [11, 12]. On the other hand, the energy current is proportional to the momentum density, and therefore propagates ballistically [13, 14]. Both phenomena, the interaction induced conductivity and the ballistic transport of energy, are relevant in the broader context of quantum criticality [15–17]. Several experiments addressed the unique transport properties of graphene at the charge neutrality point. A violation of the Wiedemann-Franz law was observed in Ref. [18], indicating the ballistic transport of energy. The interaction induced resistivity was recently measured at finite frequencies [19] and showen to be in good agreement with the theoretical prediction of Ref. [11]. Graphene has become one of the most important host systems for electron hydrodynamics in general, extensively studied in both experiment [20–24] and theory [25–40].

An important experimental prerequisite for the realization of hydrodynamic electron flow is the dominance of electron-electron scattering over any momentum relaxing scattering mechanism. Besides graphene, materials such as delafossite metals [41, 42] and Weyl semimetals [43] show non-local transport patterns and have been identified as potential candidates for the realization of hydrodynamic electron flows - a development that boosted experimental and theoretical work on the subject [44–64].

In a clean system, hydrodynamics prevails when the electron-electron scattering rate l_{ee} is much smaller than the system size l_{geo}. The ratio between these two lengths is the Knudsen number $\text{Kn} = l_{ee}/l_{\text{geo}}$. In a Poiseuille-like geometry l_{geo} corresponds to the width of the sample. The geometry of the system then sets a finite wavenumber $q \sim 2\pi/l_{\text{geo}}$. Therefore, for finite Knudsen numbers, the wave-vector dependence of transport coefficients determines the behavior of the fluid. Thinking in real space, this means that higher-order spatial derivatives have to be included into the equations of motion of the fluid, and the flow becomes highly non-local. A very similar situation occurs when the system is subjected to spatially modulated force fields, e.g. an electric field of the form $E_{\mathbf{q}} = E_0 e^{i\mathbf{q}\cdot\mathbf{x}} e^{-i\omega t}$ (see Fig. 1). The response of the fluid is then determined by a non-local conductivity tensor $\sigma_{\alpha\beta}(\mathbf{q}, \omega)$. An important example that is treated in Sec VI are surface acoustic waves (SAWs) in piezoelectric materials, which produce spatially modulated electric fields and can be used to study the longitudinal part of the non-local charge conductivity.

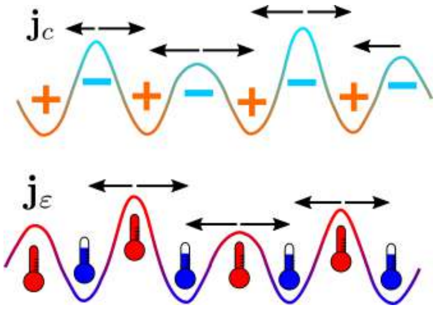


Figure 1: Charge (upper row) and energy currents (lower row) excited by wavelike longitudinal electric fields and temperature differences.

II. MAIN RESULTS

In this paper, we focus on the non-local transport properties and collective excitations of graphene electrons at the charge neutrality point - prime example for a Dirac fluid. The quantum Boltzmann method developed in Ref. [11] is used. This method relies on the fact, that at low temperatures the graphene fine structure constant α is renormalized to small values. Thermally excited electrons and holes therefore appear as sharply-defined quasiparticles, whose transport properties can be studied by means of a kinetic equation. The solution of this equation is facilitated by the presence of so-called collinear modes, whose scattering rates are enhanced by a large factor of $\log(1/\alpha)$. Here, the velocities of the interacting particles are parallel to each other. Due to the linear graphene spectrum, all particles travel at the same speed, regardless of their momentum. Particles traveling in parallel have a particularly long time to interact with each other, hence the strong enhancement. Transport in the hydrodynamic regime, however, is dominated by processes, which have the smallest scattering rates (for details see Eq. (27) and below). Such "slow" processes are represented by collinear zero modes – functions that set the collinear part of the collision operator to zero [11, 12].

We solve the kinetic equation by reducing it to a matrix equation in the space of collinear zero modes $\chi^{(m,s)}_{\mathbf{k},\lambda} = \lambda^m e^{im\theta} \{1, \lambda, \lambda\beta v\hbar k\}$ (see Sec. III B). Here, θ is the polar angle and k the modulus of the momentum variable $\mathbf{k}$, $\lambda = \pm 1$ is the band index, m labels the angular harmonics $\exp(im\theta)$, and $s \in \{1, 2, 3\}$ labels the three basis functions written in curly brackets. To an excellent approximation, it is sufficient to retain only the $s = 1$ and $s = 3$ modes. These modes describe charge (c) and energy (ε) excitations, respectively. A numerical evaluation of the collision integral's matrix elements with respect to the modes $\chi^{(m,s)}_{\mathbf{k},\lambda}$ (see Fig. 2) shows, that the relaxation rates of these modes grow linearly with increasing m:

$$\tau^{-1}_{\varepsilon/c,m} \sim |m|, \tag{1}$$

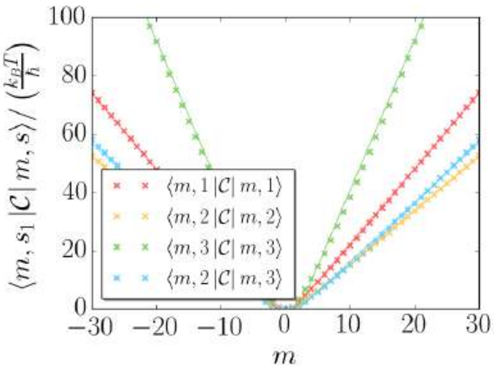


Figure 2: The matrix elements of the collision operator $\mathcal{C}$ of Eq. (24) with respect to the collinear zero modes $\chi^{(m,s)}_{\mathbf{k},\lambda} = \lambda^m e^{im\theta} \{1, \lambda, \lambda\beta v\hbar k\}$ of Eq. (32) grow linearly with increasing angular harmonic numbers m. The linear fits of Eqs. (45), (47) are plotted as solid red and green lines. The linear behavior of the matrix elements and scattering rates allows to solve the quantum Boltzmann equation exactly.

for large m (see sections III B and IV B). This unusual behavior allows us to solve the (linearized) Boltzmann equation exactly in the limit of a small α. The details of this solution are given in Sec. IV C.

It is an important feature of graphene at the neutrality point, that the hydrodynamic modes excited by electric and thermal fields decouple in linear response, and in the absence of magnetic fields [13, 65]. The modes are characterized by the distinct scattering, with all of them following Eq. (1). Using our full solution of the Boltzmann equation, the non-local, i.e. wave-vector-dependent, charge and thermal conductivities as well as the non-local viscosity were calculated. The longitudinal and transverse non-local charge conductivities as functions of wave-vector q and frequency ω are given by

$$\begin{aligned} \sigma_\parallel &= \frac{\sigma_0}{1 - i\tau_{c,1}\omega + \frac{1}{4}v^2\tau_{c,1}q^2\left(\frac{2i}{\omega} + \frac{1}{M_c(q,\omega) - i\omega}\right)}, \\ \sigma_\perp &= \frac{\sigma_0}{1 - i\tau_{c,1}\omega + \frac{\frac{1}{4}v^2\tau_{c,1}q^2}{M_c(q,\omega) - i\omega}}, \end{aligned} \tag{2}$$

where $\sigma_0 = \frac{2e^2 \log(2) k_B T \tau_{c,1}}{\pi\hbar^2}$ is the conductivity at vanishing wave-numbers and frequencies [11]. M_c is a memory function containing information on scattering in high angular momentum channels $m \geq 2$:

$$M_c(q,\omega) = \tau_{c,2}^{-1} + \frac{1}{2} vq \frac{\mathrm{I}_{3+\frac{\eta_c}{\gamma_c} - i\omega\tau_c}(\tau_c v q)}{\mathrm{I}_{2+\frac{\eta_c}{\gamma_c} - i\omega\tau_c}(\tau_c v q)}. \tag{3}$$

This result is a direct consequence of the depence of the scattering rate $\tau_{c,m}^{-1} \sim \alpha^2 k_B T |m|$ on the angular momentum state of the Dirac electron. A similar $\tau_{c,m}^{-1} \sim |m|$ behavior was found in Ref.[66] for scattering off a random magnetic field and gives rise to similar expressions for the nonlocal conductivities, caused by rather different microscopic mechanism. In Eq. (3), τ_c, γ_c and η_c determine the slopes and the offset in Eq. (1) (see Sec. IV B). The results for the non-local thermal conductivity and viscosity are given in Eqs. (65), and (70). The transport coefficients show pronounced resonance features at $vq \approx \omega$ where q and ω are the wavenumber and frequency of the applied electric field or thermal gradient (see Figs. 3, 4) and v is the electron group velocity. The longitudinal charge conductivity can be measured in experiments with surface acoustic waves (SAWs) [67–72]. The transverse conductivity determines the skin effect, which is however not a feasible measurement for a two-dimensional graphene sheet. In section VI we consider a simple device consisting of a graphene sheet laid on top of a piezoelectric crystal. We calculate the velocity shift and damping of SAWs induced by the graphene sheet and find that, while damping effects are small, a substantial velocity shift can be expected. The damping and the velocity shift measured as functions of temperature can give important insights into the nature interaction effects in a Dirac fluid.

Non-local transport coefficients also determine in confined geometries. The latter case is illustrated in Sec. VII for the electric conductivity, using the Poiseuille geometry as an example. The constitutive relation linking the electric current to the electric field along the channel is interpreted as a differential equation (Eq. (87)) and solved with the appropriate boundary conditions (Eq. (88)). We find, that the flow profiles strongly depend on the channel width w as compared to the electron-electron scattering lengths in the $m = 1$ and $m = 2$ channels: $l_{c,1} = v\tau_{c,1}$, $l_{c,2} = v\tau_{c,2}$. While $l_{c,1}$ governs the decay of charge currents, $l_{c,2}$ determines the effectiveness of current transfer from regions with high current density to regions with low current density. This latter mechanism is analogous to viscous momentum transfer. The flow profiles in dependence on w can be separated into three regimes. For $w \gg l_{c,1} > l_{c,2}$, the samples are in the Ohmic regime, where the current is dissipated uniformly across the sample. The flow profile is flat. For $l_{c,1} < w < l_{c,2}$, the profile curvature is maximal, since on the one hand the current decay due to electron-electron scattering in the $m = 1$ channel becomes inefficient, on the other hand the current transfer to the boundaries of the sample, where the flow is slowed down, is sufficiently strong. For even smaller widths $w < l_{c,2}$, the profile turns flat again, because the current transfer mechanism associated with $l_{c,2}$ ceases to be efficient. This characteristic pattern is shown in Fig. 12. Current profiles are accessible experimentally, e.g. through the scanning single electron transistor technique of Refs. [20, 73].

Finally, we calculated the dispersions of the collective modes of a Dirac fluid. As do the transport coefficients, the collective modes separate into a sector of charge excitations and a sector of energy and imbalance excitations ($s = 2$). These two sectors are decoupled and can be studied separately. We find, that while the plasmon mode is gapped out at small wave-numbers due to the interaction induced resistivity (see Fig. 6), a so-called second sound mode, corresponding to a wavelike propagation of energy, appears (Fig. 9). Diffusive modes, corresponding to the diffusion of charge, heat and quasiparticles were found (see Figs. 5, 7). Their dispersion relations were calculated and showed to agree with known results [6, 26, 74]. Besides these well studied modes, an infinite set of damped modes connected to excitations in higher angular harmonic channels was found (see Figs. 5, 8). The dispersions of these modes are purely imaginary at vanishing wave-numbers and approach in the long wavelength limit the values $\omega_m(q=0) = -i/\tau_{\varepsilon/c,m}$ for the m-th angular harmonic in the energy (ε) or the charge (c) channels. At finite wave-numbers, these modes show a complex structure of merging branches. Similar modes play an important role in the equilibration of unitary fermi gases [7] and the QCD plasma [8–10]. They also determine the unusual phase space dynamics of graphene electrons which was the subject of an earlier work [1].

Regime of validity

Transport in graphene is of interest to researchers with diverse backgrounds. Here we want to discuss the validity of our results in the context of other graphene related research. Our paper is concerned with the hydrodynmic regime, where electron tranport is governed by momentum conserving electron-electron collisions and the electron-electron mean free path is the smallest length scale [75]. In particular, momentum relaxing scattering off impurities and phonons must be weak. This demand sets serious limitations on sample sizes and on the temperature range.

1. *The Dirac fluid of graphene at the charge neutrality point*

Throughout the paper we are interested in the low energy effective behavior of graphene electrons near the Dirac point. Here, to a very good approximation, the electron dispersion is given by the massless two-dimensional Dirac Hamiltonian of Eq. (9) [76]. At $T=0$, the lower Dirac cone is fully occupied and the upper Dirac cone is empty. At finite temperatures, electrons and holes in a region of size k_BT around the Dirac cone are created. These quasiparticles are carriers of electric and thermal currents. Since their density is determined by temperature, k_BT is the only energy scale in the system. We call this regime the Dirac fluid regime. The chemical potential is vanishingly small: $\mu \ll k_BT$. For the opposite case of a large chemical potential $\mu \gg k_BT$, the system enters the Fermi liquid regime. Here, the scattering rate is given by $\tau^{-1}\hbar \sim T^2/\mu$ [77–79] (up to logarithmic corrections in 2D[80]). For the quantum critical Dirac fluid, on the other hand, the electron-electron scattering rate is determined by the temperature alone:

$$\tau^{-1} \sim \alpha^2 k_BT/\hbar \tag{4}$$

where $\alpha = e^2/(\varepsilon v\hbar)$ is the graphene fine structure constant. v is the electron group velocity and ε the dielectric constant. Higher order interaction effects can be treated in terms of the renormalization group. Integrating out high energy states above the thermal cut-off k_BT results in a logarithmic increase of the electron's group velocity [11, 17]:

$$v = v_0\left(1 + \frac{\alpha_0}{4}\log\left(\frac{\Lambda}{k_BT}\right)\right). \tag{5}$$

Here, $v_0 \approx 10^6$ m/s and α_0 are the unrenormalized, bare electron velocity and the fine structure constant. Λ is an energy on the eV scale at which the electronic bands begin to deviate from the linear Dirac-like shape. It is essential to our theory, that the fine structure constant $\alpha(T)$ is renormalized to small values when the temperature is lowered. The system is gradually approaching the free Dirac fermion fixed point, thus ensuring the validity of the quasiparticle picture and the Boltzmann approach chosen here to study the transport of electrons. Eq. (5) is a perturbative result valid to lowest order in α. However, experiments show that the logarithmic increase of the Fermi velocity at low energies is quite robust and holds even in the case of suspended graphene where $\alpha_0 \approx 2$ as well as at intermediate temperatures [81]. Thus, there is good reason to believe that even suspended graphene is located sufficiently near the free Dirac fermion fixed point, such that weak coupling results are physically meaningful; much more so for graphene grown on substrates with larger dielectric constants.

2. *The quantum Boltzmann method*

The quantum Boltzmann method is well established for systems with sharply defined quasiparticles [82, 83], the prime example being the Fermi liquid [79]. Here, thermally excited quasiparticles have energies of the order of $\varepsilon_{\text{qp}} = k_BT$, such that the ratio $\varepsilon_{\text{qp}}/\left(\hbar\tau^{-1}\right) \sim \mu/k_BT \gg 1$ is large at temperatures below the Fermi temperature. This condition, which is based on phase-space arguments rather than the interaction strength, ensures the validity of the quasiparticle picture and the Boltzmann equation.

In the case of the Dirac fluid, the ratio of the characteristic qusiparticle energy and the scattering rate is

$$\frac{\varepsilon_{\text{qp}}}{\tau^{-1}} \sim \alpha^2(T). \tag{6}$$

Thus, the quasiparticle picture is valid only at small coupling strengths. However, as discussed in the preceding section, for small temperatures $\alpha(T)$ decreases, and the Dirac fluid asymptotically approaches the free Dirac fermion limit. In this regime, the Boltzmann equation provides a powerfull tool for the study of transport phenomena. Coulomb interactions between electrons enter through a long-range Vlasov term which describes electrostatic forces due to an inhomogeneous charge distribution, as well as through the collision operator describing short-range electron-electron collisions. We use the collision operator derived in Ref. [11], which includes all scattering processes to second order in the fine sctructure constant (Born approximation). While this approach is formally exact in the small $\alpha(T)$, low temperature limit, we believe, as argued above, that it should also provide reasonable results for larger values of the fine structure constant.

In this paper, we consider the linear response of the Dirac fluid to electric fields and thermal gradients at finite frequencies. The Boltzmann approach limits our discussion to small frequencies:

$$\omega \ll \frac{k_B T}{\hbar}. \tag{7}$$

At small frequencies, the system's response is governed by intra-band processes which take place within one of the two Dirac cones. Inter-band processes, on the other hand, involve the creation of electron-hole pairs and therefore can only be excited at energies comparable to $k_B T$ [84]. This means that the off-diagonal elements of the density matrix $\left\langle \psi^{\dagger}_{\lambda,\mathbf{k}} \psi_{\lambda',\mathbf{k}} \right\rangle$,where $\psi^{\dagger}_{\lambda,\mathbf{k}}$, $\psi_{\lambda,\mathbf{k}}$ are electron creation and anihilation operators and the band index λ labes the two Dirac cones, are strongly suppressed. Allowing us to interpret the diagonal components as a distribution function

$$f_{\mathbf{k}\lambda} = \left\langle \psi^{\dagger}_{\lambda,\mathbf{k}} \psi_{\lambda,\mathbf{k}} \right\rangle$$

which can be found by solving the Boltzmann equation [82, 83]. For further details on the quantum Boltzmann approach we refer to Sec. III, Appendix A and Ref. [11].

3. *Impurities and Phonons*

At the temperature of $\sim 50\,\mathrm{K}$, and assuming $\varepsilon \approx 5$, we estimate the electron-electron mean free path as $l_{ee} = v\tau \sim 2\,\mu\mathrm{m}$. In clean graphene samples, impurity mean free paths of more than 10µm can be achieved [85], such that tranport indeed will be dominated by electron-electron scattering. A major concern in experiments with graphene near the charge neutrality point are small variations of the local chemical potential $\mu(\mathbf{x})$ which have been dubbed electron-hole puddles [86, 87]. While the origins and properties of electron puddles and their influence on transport are the subject of many studies (see e.g. [88–91]), we choose not to include them in the present theory, which is concerned with interaction effects in a clean Dirac fluid. Our results are relevant for experiments with graphene sheets in the hydrodynamic regime. Here, the dominance of electron-electron scattering over any impurity induced effects was clearly demonstrated in Ref. [19] by showing that the electron scattering rates grow linearly in accordance with Eq. (4) above a trashhold temperature.

Electron-phonon scattering is a significant disturbance for hydrodynamic electron flows at high temperatures, unless one is in a regime governed by phonon drag, see e.g. [92]. In graphene, the scattering of electrons by 2D graphene lattice phonons is limited by the small size of the Fermi-surface [93], as well as by the high Debye temperature which loweres the phonon density of states [93]. These limitations are even more pronounced at the Dirac point, where due to momentum conservation only phonons with momenta $k_{\mathrm{ph}} < k_B T/v$ participate in scattering events. However, scattering with surface optical phonons of the substrate can lead to a significant increase of the sheet resistance at higher temperatures. In Ref. [94] this mechanism was reported to set in above $150\,\mathrm{K}$ for graphene grown on $\mathrm{SiO_2}$. To a large extend, scattering on surface acoustic photons determines the decay rates of graphene plasmons at finite charge densities [95, 96]. Experiments on the hydrodynamics of Dirac fluids have been carried out with graphene sheets encapsuled in hexagonal boron nitride [18, 19]. Here electron-phonon scattering is also reported to set in at the relatively high temperatures of $70\,\mathrm{K}$ [18], or even to be insignificant up to room temperatures [19].

4. *Sample sizes*

Currently, high quality graphene sheets have sizes on the order of tenth of micrometers. On the one hand side this means that the effects of boundary scattering can be important [97]. On the other hand, it has been demonstrated that such samples are sufficiently large to go well beyond the ballistic regime and to observe hydrodynamic behavior [18–24].

In graphene nanoribbons, gaps opening at the Dirac point can significantly influence the behavior of collective modes [96, 98]. These gaps can be estimated as $\Delta \approx t/N$, where t is a characteristic tight-binding hopping amplitude on the 1eV scale and N is the number of unit cells over which the ribbon extends. For hydrodynamic samples $N \approx 10^5$, and therefore the gaps are much smaller than quasiparticle energies at experimental temperatures.

Boundary effects on collective mode propagation will give a larger correction of order l_{ee}/w, where w is the sample size (see e.g. [99]).

III. THEORETICAL FRAMEWORK

A. Kinetic equation

In order to clarify our notation, in this section we sketch the derivation of the quantum Boltzmann formalism for the Dirac fluid, which was developed in Ref. [11]. We begin with the Hamiltonian of graphene electrons at the charge neutrality point:

$$H = H_0 + H_{\mathrm{int}}, \tag{8}$$

where the free part is given by

$$H_0 = v\hbar \int_{\mathbf{k}} \sum_{a,b,i} \psi^{\dagger}_{a,i}(\mathbf{k}) (\mathbf{k}\cdot\sigma)_{ab}\, \psi_{b,i}(\mathbf{k}), \tag{9}$$

and the interaction part reads

$$H_{\mathrm{int}} = \frac{1}{2}\int_{\mathbf{k},\mathbf{k}',\mathbf{q}} \sum_{a,b,i,j} V(\mathbf{q})\, \psi^{\dagger}_{\mathbf{k}+\mathbf{q},a,i} \psi^{\dagger}_{\mathbf{k}'-\mathbf{q},b,j} \psi_{\mathbf{k}',b,j} \psi_{\mathbf{k},a,i}. \tag{10}$$

$V(\mathbf{q}) = \frac{2\pi e^2}{\varepsilon|\mathbf{q}|}$ is the 2D Coulomb potential. The indices $i, j = 1, 2 ..., N = 4$ refer to the spin and valley quantum numbers of an electron, whereas the two sub-lattices are labelled by the indices a, b. The free particle Hamiltonian H_0 is diagonalized by the unitary transformation

$$U_{\mathbf{k}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & o^*_{\mathbf{k}} \\ 1 & -o^*_{\mathbf{k}} \end{bmatrix}, \tag{11}$$

where $o_{\mathbf{k}} = (k_x + ik_y)/\sqrt{k_x^2 + k_y^2}$.

For the derivation of the quantum Boltzmann equation, it is convenient to use the band representation of Dirac spinors $\psi_{\lambda,\mathbf{k}} = U_{\mathbf{k},\lambda a}\psi_{\mathbf{k},a}$ with $\lambda = \pm 1$ labeling the upper and lower Dirac cones. In this way, one can easily distinguish between processes that involve the creation of particle-hole pairs and those which do not. The thermally excited electron-hole pairs occupy states in a window of $k_B T$ around the Dirac point. Thus, if the applied fields have frequencies $\omega < 2k_B T/\hbar$, which is true in the hydrodynamic regime, processes that create electron-hole pairs are unlikely and can be neglected. This translates to neglecting the off-diagonal components of the distribution function in the band representation, which is then given by its diagonal elements:

$$f_{\mathbf{k}\lambda} = \left\langle \psi^{\dagger}_{\lambda,\mathbf{k}} \psi_{\lambda,\mathbf{k}} \right\rangle.$$

The quantum Boltzmann equation then reads

$$\left(\partial_t + \mathbf{v}_{\mathbf{k}\lambda}\cdot\nabla_{\mathbf{r}} - (e\nabla\varphi_{\mathrm{tot}})\cdot\nabla_{\mathbf{k}} + C\right) f_{\mathbf{k}\lambda}(\mathbf{r},t) = 0. \tag{12}$$

Here, $\mathbf{v}_{\mathbf{k}\lambda} = \partial\varepsilon_{\mathbf{k}\lambda}/\partial\mathbf{k}$ is the group velocity and

$$\varphi_{\mathrm{tot}}(\mathbf{r},t) = \varphi_{\mathrm{ext}}(\mathbf{r},t) + \varphi_{\mathrm{ind}}(\mathbf{r},t) \tag{13}$$

is the sum of the external electrostatic potential and the induced potential which is the result of an inhomogeneous distribution of charges. The term associated with φ_{tot} was first introduced by Vlasov[100]. It will be dealt with at the end of this section. C represents the central part of the kinetic theory - the Boltzmann collision operator describingelectron-electron Coulomb scattering. Details on the derivation of C are summarized in Appendix A, based on Refs. [1, 11].

Studying the linear response to φ_{tot}, we expand the distribution function around the local equilibrium distribution $f^{(0)}_{k\lambda}$

$$f_{\mathbf{k}\lambda}(\mathbf{r},t) = f^{(0)}_{k\lambda} + w_k \psi_{\mathbf{k}\lambda}(\mathbf{r},t). \tag{14}$$

where $f^{0}_{k,\lambda}$ is given by

$$f^{(0)}_{k\lambda} = \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}\lambda} - \mathbf{u}\cdot\mathbf{k})} + 1}. \tag{15}$$

The product $w_k \equiv f^{(0)}_k \left(1 - f^{(0)}_k\right)$, that will soon play the role of a weight function in the scalar product, does not depend on λ, and the corresponding index is dropped in Eq. (14) and in the following.

Performing a Fourier transformation $\psi_{\mathbf{k}\lambda}(\mathbf{r},t) \to \psi_{\mathbf{k}\lambda}(\mathbf{q},\omega)$ to frequency and momentum space, we obtain the linearized Boltzmann equation

$$(\mathcal{L} + \mathcal{C})\, \psi_{\mathbf{k}\lambda}(\mathbf{q},\omega) = S_{\mathbf{k}\lambda}(\mathbf{q},\omega). \tag{16}$$

$\mathcal{L}$ is the Liouville operator and given by

$$\mathcal{L} = -i\omega + i\mathbf{q}\cdot\mathbf{v}_{\mathbf{k}\lambda} \tag{17}$$

The linearization of the collision operator can be expressed in the form

$$\mathcal{C}\psi_{\mathbf{k}\lambda} \approx \frac{1}{w_k}\sum_{\lambda'}\int_{\mathbf{k}'} \frac{\delta(C\psi)_{\mathbf{k}\lambda}}{\delta\psi_{\mathbf{k}'\lambda'}} \psi_{\mathbf{k}'\lambda'}, \tag{18}$$

where the weight function w_k was introduced above.

Let the $\psi_{\boldsymbol{k}}$ be element of a function space with inner product

$$\langle \phi \mid \psi \rangle = \sum_{\lambda}\int_{\mathbf{k}} w_k \phi^*_{\mathbf{k}\lambda}\psi_{\mathbf{k}\lambda}, \tag{19}$$

such that

$$\begin{aligned} \langle \phi |\mathcal{C}| \psi \rangle &= \sum_{\lambda}\int_{\mathbf{k}} w_k \phi^*_{\mathbf{k}\lambda}\mathcal{C}\psi_{\mathbf{k}\lambda} \\ &= \sum_{\lambda\lambda'}\int_{\mathbf{k}\mathbf{k}'} \phi^*_{\mathbf{k}\lambda} \frac{\delta(C\psi)_{\mathbf{k}\lambda}}{\delta\psi_{\mathbf{k}'\lambda'}} \psi_{\mathbf{k}'\lambda'}. \end{aligned} \tag{20}$$

One can show that the entropy production in the absence of external driving terms is $\frac{\partial S}{\partial t} = k_B \langle \psi |\mathcal{C}| \psi \rangle$ which ensures that the collision operator is positive definite. In fact, $\mathcal{C}$ is Hermitian under the above scalar product. Therefore its eigenvalues are real and its eigenfunction form an orthonormal basis of the function space.

The right hand side of Eq. (16) is determined by the forces acting on the system. The three force terms studied here are due to electric fields, thermal gradients and viscous forces. For an electric field oriented along the x-axis, $\mathbf{E} = E_0\hat{\mathbf{e}}_x$, the force term reads

$$S_E = -eE_0\cos\theta\,(\lambda v\beta), \tag{21}$$

where θ is the polar angle of the momentum $\mathbf{k}$. It is important to notice, that

$$\mathbf{E} = -\nabla\varphi_{\mathrm{tot}}.$$

The corresponding term for a thermal gradient ∇T is given by

$$S_T = -k\,|\nabla T|\cos\theta k_B\,(v\beta)^2\,. \tag{22}$$

A viscous force is present if the drift velocity $\mathbf{u}$ in the local equilibrium distribution function (15) is a function of the coordinate $\mathbf{x}$. Then the drift term of the Boltzmann equation (12) can be thought of as a force term

$$\begin{aligned} S_S &= -vkX_{0,\alpha\beta}\left(\frac{k_\alpha k_\beta}{k^2} - \frac{1}{2}\delta_{\alpha\beta}\right)\lambda\beta \\ &= -\frac{1}{2}kX_0\sin(2\theta)\,(\lambda v\beta)\,, \end{aligned} \tag{23}$$

where the stress tensor is given by

$$X_{0,\alpha\beta} = \frac{1}{2}\left(\frac{\partial u_\alpha}{\partial x_\beta} + \frac{\partial u_\beta}{\partial x_\alpha} - 2\delta_{\alpha\beta}\nabla\cdot\mathbf{u}\right).$$

In the following, we consider a flow with $\mathbf{u}(y) = u(y)\,\hat{\mathbf{e}}_x$ and therefore only include the component $X_{0,xy}$, which is relevant for the calculation of the shear viscosity.

The collision operator is given by

$$\begin{aligned} (C\psi)_{\mathbf{k}\lambda} \;=\;& \frac{2\pi}{\hbar}\int_{k'q}\delta\left(k+k'-|\mathbf{k}+\mathbf{q}|-|\mathbf{k}'-\mathbf{q}|\right) \\ \times\;& \left(1-f_k^{(0)}\right)\left(1-f_{k'}^{(0)}\right)f^{(0)}_{|\mathbf{k}+\mathbf{q}|}f^{(0)}_{|\mathbf{k}'-\mathbf{q}|} \\ \times\;& \Big\{\gamma^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}\left(\psi_{\mathbf{k}+\mathbf{q},\lambda}+\psi_{\mathbf{k}'-\mathbf{q},\lambda}-\psi_{\mathbf{k}',\lambda}-\psi_{\mathbf{k},\lambda}\right) \\ +\;& \gamma^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}\left(\psi_{\mathbf{k}+\mathbf{q},\lambda}-\psi_{-\mathbf{k}'+\mathbf{q},\bar{\lambda}}+\psi_{-\mathbf{k}',\bar{\lambda}}-\psi_{\mathbf{k},\lambda}\right)\Big\}\,. \end{aligned} \tag{24}$$

The matrix elements $\gamma^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}$, $\gamma^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}$ can be found in Appendix A.

Another important term in the kinetic equation describes the electrostatic forces that arise due to an inhomogeneous distribution of charges. These forces are mediated by a self consistent potential φ_{ind}, first introduced by Vlasov [100]. It reads

$$e\varphi_{\mathrm{ind}}(\mathbf{r},t) = \alpha vN\int d^2r'\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\,\frac{\delta f_{\mathbf{k}\lambda}(\mathbf{r}',\omega)}{|\mathbf{r}-\mathbf{r}'|}, \tag{25}$$

where we have used the abbreviation $\delta f_{\mathbf{k}\lambda}(\mathbf{r},\omega) = w_k\psi_{\mathbf{k},\lambda}(\mathbf{r},\omega)$ and multiplied the potential by e for notational convenience. A derivation of the term can be found in Ref. [82] (Eqs. (7-3) and (9-16)). Applying a Fourier transform to Eq. (25) one finds

$$e\varphi_{\mathrm{ind}}(\mathbf{q},t) \;=\; \alpha vN\sum_\lambda\int\frac{d^2k}{(2\pi)^2}\,\frac{2\pi\delta f_{\mathbf{k}\lambda}(\mathbf{q},\omega)}{q}.$$

In Sec. V we will not be interested in the response to the total electric field $\mathbf{E} = -i\mathbf{q}\varphi_{\mathrm{tot}}$, but rather in solutions of the homogeneous Boltzmann equation

$$(\mathcal{L}+\mathcal{V}+\mathcal{C})\,\psi_{\mathbf{k}\lambda}(\mathbf{q},\omega) = 0.$$

Here, the Vlasov term

$$\mathcal{V}\psi_{\mathbf{k}\lambda} = -i\mathbf{q}\cdot\mathbf{v}_{\mathbf{k}\lambda}e\varphi_{\mathrm{ind}}. \tag{26}$$

has to be included explicitely.

B. Collinear zero modes

In this section, we summarize how Eq. (16) is solved in the limit of a small fine structure constant. A standard way to deal with an integral equation like (16) is to expand the function $\psi_{\mathbf{k},\lambda}$ into a set of suitable basis functions. The choice of this basis is facilitated by the fact that for small values of the graphene fine structure constant α, the collision operator (24) logarithmically diverges if the velocities of involved particles are parallel to each other. This is a consequence of the linear single particle spectrum, and the resulting momentum independent velocity of massless Dirac particles. Intuitively speaking, the scattering is enhanced, because particles traveling in the same direction interact with each other over a particularly long period of time. A more mathematical picture of this so-called collinear scattering anomaly is presented in Appendix (B). It is convenient to write the collision operator as a sum of the collinear part $\mathcal{C}_c$ and the non-collinear part $\mathcal{C}_{nc}$:

$$\mathcal{C} = \log(1/\alpha)\mathcal{C}_c + \mathcal{C}_{nc}. \tag{27}$$

The factor $\log(1/\alpha)$ is large at small α. Both operators, $\mathcal{C}_c$ and $\mathcal{C}_{nc}$, are hermitian with respect to the scalar product of Eq.19. Let $\varphi^n_{\mathbf{k},\lambda}$ be the orthogonal eigenfunctions of $\mathcal{C}_c$ such that

$$(\mathcal{C}_c\varphi^n)_{\mathbf{k},\lambda} = b_n\varphi^n_{\mathbf{k},\lambda}. \tag{28}$$

$\psi_{\mathbf{k},\lambda}$ is expanded in terms of these functions:

$$\psi_{\mathbf{k},\lambda} = \sum_n\gamma_n\varphi^n_{\mathbf{k},\lambda}. \tag{29}$$

Suppose, some of the orthogonal basis functions φ^n, namely those with $n < n_0$, set the collinear part of the collision operator to zero, i.e.

$$\mathcal{C}_c\varphi^{n<n_0} = 0. \tag{30}$$

Then, inserting the expansion (29) into Eq. (16) and projecting it onto the basis functions $\varphi^{n'}$, one finds

$$\gamma_{n'>n_0} = \frac{\left\langle\varphi^{n'}\,\middle|\,S\right\rangle - \left\langle\varphi^{n'}\,\middle|\,(\mathcal{L}+\mathcal{C}_{nc})\,\psi\right\rangle}{b_{n'}\log(1/\alpha)}. \tag{31}$$

Hence, zero modes of $\mathcal{C}_c$ are enhanced by factor $\log\left(1/\alpha\right)$[11]. These colinear zero modes can be found from the collision operator given in Eq.24:

$$\chi_{\mathbf{k},\lambda}^{(m,s)} = \lambda^{m} e^{im\theta}\left\{1, \lambda, \lambda\beta v\hbar k\right\}. \qquad (32)$$

Here, m labels the angular momentum, $s \in \{1, 2, 3\}$ the modes $\{1, \lambda, \lambda\beta v k\}$, and θ is the polar angle of the momentum vector $\mathbf{k}$. All modes set the integral (24) to zero for collinear processes (see Appendix B).

From Eq. (31) follows that for small values of α, only the collinear zero modes have to be retained in the expansion of the entire collision operator Eq. (29), i.e. the kinetic equation (12) can be solved using the restricted subspace of basis functions of Eq. (32). The stronger colinear scattering processes give rise to a rapid equilibration to the subset of modes given in Eq. (32) which then dominate the long-time dynamics.

In order to proceed, the matrix elements of Eq. (16) in this basis must be calculated. The matrix elements of the Liouville operator $\mathcal{L}$ are given by

$$\left\langle \chi_{\mathbf{k},\lambda}^{(m,s)} \left|\mathcal{L}\right| \chi_{\mathbf{k},\lambda}^{(m',s')} \right\rangle = \left(-i\omega\delta_{m,m'} + \frac{1}{2} ivq\left(e^{-i\vartheta_{\mathbf{q}}}\delta_{m,m'+1} + e^{i\vartheta_{\mathbf{q}}}\delta_{m,m'-1}\right)\right)\left(v\beta\hbar\right)^{-2} L_{s,s'}, \qquad (33)$$

where $\vartheta_{\mathbf{q}}$ is the polar angle of the wave-vector $\mathbf{q}$ and

$$L = \begin{bmatrix} \frac{\log(2)}{\pi} & 0 & 0 \\ 0 & \frac{\log(2)}{\pi} & \frac{\pi}{6} \\ 0 & \frac{\pi}{6} & \frac{9\zeta(3)}{2\pi} \end{bmatrix}. \qquad (34)$$

The rows and columns of the matrix notation refer to the mode index s of Eq. (32).

We calculate the matrix elements of the collision operator $\mathcal{C}$ numerically (some values are given in Appendix C). Due to the rotational invariance of the low-energy Dirac Hamiltonian (8), they are diagonal in the angular harmonic representation. Most importantly, the matrix elements rapidly approach a linear behavior for large $|m|$:

$$\left\langle \chi_{\mathbf{k},\lambda}^{(m,s)} \left|\mathcal{C}\right| \chi_{\mathbf{k},\lambda}^{(m',s')} \right\rangle = \frac{\delta_{m,m'}}{v^{2}\beta^{3}\hbar^{3}}\left(\left|m\right|\gamma_{s,s'} - \eta_{s,s'}\right). \qquad (35)$$

$\gamma_{s,s'}$ and $\eta_{s,s'}$ are numerical coefficients that are listed below Eqs. (45) and (47). This surprising result is due to the linear Dirac spectrum of the system. It allows to solve the Boltzmann equation exactly, as will be seen later. The linear behavior of the scattering rates is also shown in Fig. 2. To find closed expressions for the nonlocal transport coefficients, the scattering rates are approximated by Eq. (35) for $m > 2$. In principle, the numerically exact scattering rates up to an arbitrary m can be included. Here, the rates for $m > 2$ will be assumed to follow Eq.35 in order to keep the algebraic efforts at a minimum. The projections of the force terms (21)-(23) onto collinear zero modes read

$$\left\langle S_{E}\,|\,\chi_{\mathbf{k},\lambda}\right\rangle = -\frac{eE_{0}}{2\hbar^{2}\beta v}\delta_{|m|,1}\begin{bmatrix} \frac{\log(2)}{\pi} \\ 0 \\ 0 \end{bmatrix}, \qquad (36)$$

$$\left\langle S_{T}\,|\,\chi_{\mathbf{k},\lambda}\right\rangle = \frac{\left|\nabla T\right| k_{B}\pi^{4}}{v\beta\hbar^{2}}\delta_{|m|,1}\begin{bmatrix} 0 \\ \frac{\pi}{6} \\ \frac{9\zeta(3)}{2\pi} \end{bmatrix}, \qquad (37)$$

$$\left\langle S_{S}\,|\,\chi_{\mathbf{k},\lambda}\right\rangle = -\frac{iX_{0}}{4\left(v\beta\hbar\right)^{2}}\operatorname{sign}\left(m\right)\delta_{|m|,2}\begin{bmatrix} 0 \\ \frac{\pi}{6} \\ \frac{9\zeta(3)}{2\pi} \end{bmatrix}. \qquad (38)$$

For the Vlasov term (26) one finds

$$\begin{aligned} \left\langle \psi_{\mathbf{k},\lambda}^{(m,s)} \left|\mathcal{V}\right| \psi_{\mathbf{k},\lambda}^{(m',s')} \right\rangle &= i\alpha N\left(e^{-i\vartheta_{\mathbf{q}}}\delta_{m,1} + e^{i\vartheta_{\mathbf{q}}}\delta_{m,-1}\right) \\ &\times \frac{\delta_{1,s}\delta_{1,s'}\delta_{m',0}}{2v^{2}\beta^{3}\hbar^{3}}\begin{bmatrix} \frac{\log(2)^{2}}{\pi^{2}} \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \qquad (39)$$

The non-equilibrium part of the distribution function expanded in the subset of colinear zero modes becomes

$$\psi_{\mathbf{k},\lambda} = \sum_{m=-\infty}^{\infty}\sum_{s=1}^{3} a_{m,s}\left(\omega, \mathbf{q}\right)\chi_{\mathbf{k},\lambda}^{(m,s)}. \qquad (40)$$

Together, the expressions (16), (33), (35), (36)-(38), (39) and (40) provide a linearized kinetic equation restricted to the basis of collinear zero modes that becomes exact for small values of the fine structure constant α. Since no assumptions on the spatial dependencies were made, except that they are be within the limits of the applicability of the kinetic equation, this expansion can be used to derive the non-local transport coefficients in the linear-response regime, as well as the dispersion relations of collective excitations.

IV. NON-LOCAL TRANSPORT

A. Effects of electron-hole symmetry, momentum conservation and thermal transport

Within the kinetic approach, the charge current $\mathbf{j}_c$ and the heat current $\mathbf{j}_\varepsilon$ are given by

$$\mathbf{j}_c = e\sum_\lambda \int_{\mathbf{k}} \lambda v \frac{\mathbf{k}}{k} f_{\mathbf{k},\lambda}, \tag{41}$$

$$\mathbf{j}_\varepsilon = \sum_\lambda \int_{\mathbf{k}} v^2 \hbar \mathbf{k} f_{\mathbf{k},\lambda}. \tag{42}$$

In these expressions intra-band processes that create particle-hole pairs are neglected (see Appendix A). It follows from Eqs. (41) (42), that the even in λ part of the distribution function $f_{\mathbf{k},\lambda}$ contains information about thermal transport, whereas the odd part governs the transport of charge. Since the electric field contribution to the kinetic equation (21) is odd in λ, and the thermal gradient leads to a term that is even in λ (Eq. (22)), the phenomena of thermal and charge transport are decoupled to linear order in the external fields at the neutrality point. This can be traced back to particle-hole symmetry and is the ultimate reason why the Wiedemann-Franz law is dramatically violated in a Dirac fluid[18]. The distribution function shows a similar decoupling ocf charge and heat modes for higher m: The collinear modes of Eq. (32) are proportional to λ^m for $s = 1$ and to λ^{m+1} for $s = 2, 3$. Consequently the kinetic equation in the subspace of collinear zero modes is block diagonal in the $s = 1$ and $s = 2, 3$ modes, as can be seen from Eqs. (24), (33), (36)-(38). In the following this will further simplify the calculation of transport coefficients.

Another important consequence of the linear graphene spectrum is that the heat current $\mathbf{j}_\varepsilon$ is proportional to the momentum density $\mathbf{g} = \sum_\lambda \int_k \hbar \mathbf{k} f_{\mathbf{k},\lambda}$ and is therefore conserved. The charge current, unlike in Galilean invariant systems, is not conserved, and decays due to interactions, giving rise to a finite restistvity in the clean system.

B. Scattering times

The matrix elements of the collision operator determine the scattering rates of the three collinear zero modes in different angular harmonic channels. In the absence of spatial inhomogeneities and external forces, the kinetic equation in the basis of collinear zero modes (32) reads

$$\sum_{s'} \left(\partial_t \delta_{s,s'} + \Gamma_m^{s,s'}\right) a_{m,s'} = 0, \tag{43}$$

where the $a_{m,s}$ are the coefficients of the expansion (40). Posed as an initial value problem, this equation describes the exponential decay of collinear zero modes. This decay governs the behavior of the system at long time scales, because modes that do not set the collinear part of the collision integral to zero decay faster by a factor $\log(1/\alpha)$ (see Eq. (27)).

The scattering rates $\Gamma_m^{s,s'}$ are given by

$$\Gamma_m^{s,s'} = (v\beta\hbar)^2 L^{-1}_{s,s'} \left\langle \chi^{(m,s)}_{\mathbf{k},\lambda} \left|\mathcal{C}\right| \chi^{(m',s')}_{\mathbf{k},\lambda} \right\rangle. \tag{44}$$

Because of the definition of the scalar product in Eq. (19), the matrix elements have dimension length²/time. Vanishing scattering rates indicate conservation laws, and the corresponding modes are zero modes of the full collision operator as well as its collinear part. These modes reflect the conservation of particle density, imbalance density, energy density and momentum density:

$$\chi^{(s=1,m=0)}_{\mathbf{k},\lambda} = 1, \qquad \chi^{(s=2,m=0)}_{\mathbf{k},\lambda} = \lambda,$$

$$\chi^{(s=3,m=0)}_{\mathbf{k},\lambda} = \lambda\beta v\hbar k, \qquad \chi^{(s=3,m=1)}_{\mathbf{k},\lambda} = \lambda e^{i\theta}\beta v\hbar k.$$

The imbalance density is conserved only to order α^2, as it decays due to higher order interaction processes. An important simplification stems from the fact that all scattering rates, for large $|m|$, share the asymptotic behavior $\Gamma_m \sim |m|$. This becomes a reasonable approximation for the scattering rates with $m \geq 2$. In the next section it is shown, how this behavior allows us to obtain closed form expressions for the non-local transport coefficients. As discussed in the previous section, the matrix of scattering rates $\Gamma_m^{s,s'}$ is block diagonal in the modes describing charge ($s = 1$) and thermal excitations ($s = 2, 3$), i.e. $\Gamma_m^{1,2} = \Gamma_m^{2,1} = \Gamma_m^{1,3} = \Gamma_m^{3,1} = 0$. Therefore, the scattering times determining the non-local electric conductivity are given by $\tau_{c,m} = 1/\Gamma_m^{1,1}$: $\tau_{c,0} \to \infty$, $\tau_{c,1} = \frac{1}{\alpha^2}\frac{\hbar}{k_B T}\frac{\log 2}{0.804\pi}$, $\tau_{c,2} = \frac{1}{\alpha^2}\frac{\hbar}{k_B T}\frac{\log 2}{2.617\pi}$ as well as

$$\tau_{c,m} \approx \frac{1}{\alpha^2}\frac{\hbar}{k_B T}\frac{\log 2}{\pi}\left(\gamma_c \cdot |m| - \eta_c\right)^{-1} \text{ if } m > 2, \tag{45}$$

where $\gamma_c = 2.57$ and $\eta_c = 3.45$ (see Appendix C for more numerical values). It is also convenient to define an effective scattering time for the Vlasov term:

$$\tau_V = \frac{2\pi^2\beta\hbar}{\alpha N \log(2)}. \tag{46}$$

Notice, that $\tau_V/\tau_{c,m} \sim 1/\alpha$ is large for small α.

In the thermal sector, there are two relevant modes. However, the $s = 3$ mode is physically more important, because the vanishing of the corresponding scattering rates for the $m = 0$ and $m = 1$ channels indicate the conservation of energy and momentum. In the following, it is shown that the neglecting of the $s = 2$ imbalance mode in the calculation of the thermal conductivity and viscosity, while significantly simplifying the analysis, does only result in a small numerical error. Therefore, for the purpose of calculating the transport coefficients, only the $s = 3$ energy mode will be considered. The scattering times are then given by $\tau_{\varepsilon,m} = 1/\Gamma^{3,3}_{m}$. Because of energy and momentum conservation, we have $\tau_{\varepsilon,m=0,1} \rightarrow \infty$, and for $m = 2$, it is $\tau_{\varepsilon,2} = \frac{1}{\alpha^2}\frac{\hbar}{k_B T}\frac{9\zeta(3)}{3.341\cdot 2\pi}$. For $m > 2$ the linear approximation can be used:

$$\tau_{\varepsilon,m} \approx \frac{1}{\alpha^2}\frac{\hbar}{k_B T}\frac{9\zeta(3)}{2\pi}\left(\gamma_\varepsilon \cdot |m| - \eta_\varepsilon\right)^{-1} \;\; m > 2, \qquad (47)$$

with $\gamma_\varepsilon = 5.18$ and $\eta_\varepsilon = 11.3$.

C. Non-local transport coefficients

The linear, non-local response of a system to external forces $\mathcal{F}(\mathbf{r})$ is characterized by constitutive relations of the form

$$\mathcal{J}(\mathbf{r},t) = \int d^d r' dt' \, \nu\left(\mathbf{r}-\mathbf{r}', t-t'\right) \mathcal{F}\left(\mathbf{r}', t'\right), \qquad (48)$$

where $\mathcal{J}(\mathbf{r},t)$ is a current sourced by the field $\mathcal{F}(\mathbf{r}',t')$ and $\nu(\mathbf{r}-\mathbf{r}', t-t')$ is the corresponding transport coefficient. $\mathcal{F}$ can be a scalar potential, a vector field (an electric field or a thermal gradient), or a tensor. Eq. (48) takes a much simpler form in Fourier space:

$$\mathcal{J}(\mathbf{q},\omega) = \nu(\mathbf{q},\omega)\,\mathcal{F}(\mathbf{q},\omega). \qquad (49)$$

If the system is confined to a geometry of a characteristic size l_{geo}, the relevant wave vectors $\mathbf{q}$ in Eq. (49) will be of the order of $q_{\text{geo}} \approx 2\pi/l_{\text{geo}}$. On the other hand $\nu(\mathbf{q},\omega)$ varies on scales of the inverse mean free path $q_{mf} \approx 2\pi/l_{mf}$, where $l_{mf} = v\tau$ and τ is the relevant relaxation time. Thus if $l_{\text{geo}} \gg l_{mf}$, we can approximate $\nu(\mathbf{q}_{\text{geo}},\omega) \approx \nu(\mathbf{q}=0,\omega)$. We then have

$$\nu\left(\mathbf{r}-\mathbf{r}',\omega\right) \approx \nu_0\left(\mathbf{q}=0,\omega\right)\delta\left(\mathbf{r}-\mathbf{r}'\right) \qquad (50)$$

and the constitutive relation (48) reduces to its local form $\mathcal{J}(\mathbf{r},\omega) = \nu_0(\omega)\,\mathcal{F}(\mathbf{r},\omega)$. The non-locality of Eq. (48) matters if $l_{\text{geo}} \lesssim l_{mf}$. On scales comparable to the mean free path, transport is intrinsically non-local, because particles loose their memory of previous events through collisions with other particles or impurities - a mechanism that ceases to be efficient. A good example is the Poiseuille flow through narrow channels described in Sec. VII. We proceed with the calculation of the non-local, i.e. wavenumber dependent electric conductivity, thermal conductivity and viscosity using the kinetic equation (12) and the collinear zero mode expansion summarized in Sec. III B.

1. *Electric conductivity*

As mentioned in Sec. IV B, only the first collinear mode $s = 1$ is involved in the calculation of the electric conductivity. Inserting the expansion of the distribution function in terms of collinear zero modes (40) into the kinetic equation (16) using its matrix representation of Eqs. (33), (35), (36)-(38) and (39), the left hand side of (16) can be transformed into a recurrence relation for the coefficients $a_{1,m}$, where, for the rest of this section, the $s = 1$ index is dropped. A similar analysis for electrons in a random magnetic field was performed in Ref.[66]. For $m > 2$, Eq. (45) can be used, and the recurrence relation reads

$$a_{m+1} = \frac{2ie^{-i\vartheta_{\mathbf{q}}}}{vq}\left(i\omega - \tau^{-1}_{c,m}\right)a_m - e^{-2i\vartheta_{\mathbf{q}}}a_{m-1}. \qquad (51)$$

This recurrence relation has the form

$$a_{m+1} = \left(\alpha' m + \beta'\right)a_m - e^{i\delta}a_{m-1} \qquad (52)$$

with $\alpha' = -\frac{2ie^{-i\vartheta_{\mathbf{q}}}}{vq}\frac{k_B T}{\hbar}\frac{\pi}{\log 2}\gamma_c$, $\beta' = \frac{2ie^{-i\vartheta_{\mathbf{q}}}}{vq}\left(i\omega - \eta_c\frac{k_B T}{\hbar}\frac{\pi}{\log 2}\right)$ and $\delta = -2\vartheta_{\mathbf{q}}$. It has two solutions that can be given in terms of modified Bessel functions. The physically interesting solution is

$$a_m = c \cdot e^{i\frac{\delta}{2}\left(m+\frac{\beta'}{\alpha'}\right)}\mathrm{I}_{m+\frac{\beta'}{\alpha'}}\left(-\frac{2e^{i\delta/2}}{\alpha'}\right), \qquad (53)$$

where $\mathrm{I}_\nu(z)$ is the modified Bessel function of the first kind. Another solution that diverges for $m \rightarrow \infty$ is given by

$$c_m = c \cdot e^{i\frac{\delta}{2}\left(m+\frac{\beta'}{\alpha'}\right)}\mathrm{K}_{m+\frac{\beta'}{\alpha'}}\left(\frac{2e^{i\delta/2}}{\alpha'}\right).$$

K_ν is the modified Bessel function of the second kind. Making use of the coefficients a_m for $m > 2$ as given by Eq. (53), the kinetic equation can be reduced to a 5×5 component matrix equation:

$$\begin{bmatrix} -i\omega+M_c\left(q,\omega\right) & \frac{1}{2}ivqe^{i\vartheta_{\mathbf{q}}} & 0 & 0 & 0 \\ \frac{1}{2}ivqe^{-i\vartheta_{\mathbf{q}}} & -i\omega+\tau_{c,1}^{-1} & \frac{1}{2}ivqe^{i\vartheta_{\mathbf{q}}} & 0 & 0 \\ 0 & \frac{1}{2}ivqe^{-i\vartheta_{\mathbf{q}}} & -i\omega & \frac{1}{2}ivqe^{i\vartheta_{\mathbf{q}}} & 0 \\ 0 & 0 & \frac{1}{2}ivqe^{-i\vartheta_{\mathbf{q}}} & -i\omega+\tau_{c,1}^{-1} & \frac{1}{2}ivqe^{i\vartheta_{\mathbf{q}}} \\ 0 & 0 & 0 & \frac{1}{2}ivqe^{-i\vartheta_{\mathbf{q}}} & -i\omega+M_c\left(q,\omega\right) \end{bmatrix} \begin{bmatrix} a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{eE_0\beta v}{2} \\ 0 \\ \frac{eE_0\beta v}{2} \\ 0 \end{bmatrix}, \tag{54}$$

where $M_c\left(q,\omega\right)=\tau_{c,2}^{-1}+a_3\left(q,\omega\right)/a_2\left(q,\omega\right)$ is a memory function containing information on scattering channels with higher angular momentum numbers. Using the Eqs. (52) and (53), the memory function is written

$$M_c\left(q,\omega\right)=\tau_{c,2}^{-1}+\frac{1}{2}vq\frac{\mathrm{I}_{3+\frac{\eta_c}{\gamma_c}-i\omega\tau_c}\left(\tau_c vq\right)}{\mathrm{I}_{2+\frac{\eta_c}{\gamma_c}-i\omega\tau_c}\left(\tau_c vq\right)}, \tag{55}$$

with the abbreviation $\tau_c=\frac{\hbar}{k_BT}\frac{\log 2}{\pi}\gamma_c^{-1}$. It is now straightforward to calculate the electric conductivity from the relation

$$j_{c,x}\left(\mathbf{q},\omega\right)=\sigma_{xx}\left(\mathbf{q},\omega\right)E_x\left(\mathbf{q},\omega\right). \tag{56}$$

The non-local conductivity can be decomposed into a longitudinal part $\sigma_{\parallel}\left(\omega,q\right)$ and a transverse part $\sigma_{\perp}\left(\omega,q\right)$, both depending on the modulus of $\mathbf{q}$. The longitudinal and transverse parts describe currents that flow in the direction of $\mathbf{q}$, or orthogonal to $\mathbf{q}$, respectively:

$$\sigma_{\alpha\beta}=\frac{q_\alpha q_\beta}{q^2}\sigma_{\parallel}\left(q,\omega\right)+\left(\delta_{\alpha\beta}-\frac{q_\alpha q_\beta}{q^2}\right)\sigma_{\perp}\left(q,\omega\right). \tag{57}$$

We assumed that the electric field is parallel to the x-axis. According to Eq. (57), $\sigma_{\parallel}\left(q,\omega\right)$ can be read off from the x-component of the current density $j_{c,x}$ by letting $\mathbf{q}$ be parallel to $\mathbf{e}_x$, and $\sigma_{\perp}\left(q,\omega\right)$ by considering the case $\mathbf{q}\parallel\mathbf{e}_y$. The conductivities are then given by

$$\begin{aligned} \sigma_{\parallel} &= \frac{\sigma_0}{1-i\tau_{1,c}\omega+\frac{1}{4}v^2\tau_{c,1}q^2\left(\frac{2i}{\omega}+\frac{1}{M_c(q,\omega)-i\omega}\right)}, \\ \sigma_{\perp} &= \frac{\sigma_0}{1-i\tau_{c,1}\omega+\frac{\frac{1}{4}v^2\tau_{c,1}q^2}{M_c(q,\omega)-i\omega}}, \end{aligned} \tag{58}$$

where $\sigma_0=N\frac{e^2\log(2)\tau_{c,1}}{2\pi\beta\hbar^2}$ is the quantum critical conductivity calculated in Ref. [11]. Note that $\sigma_{\parallel}\left(q\neq 0,\omega=0\right)=0$ holds, which also follows from formula (60). If this was not the case, static currents with a finite wave-vector $\mathbf{q}$ would lead to an infinite accumulation of charge at certain points, which is forbidden by the conservation of charge. In Fig. 3 the charge conductivities are plotted as functions of ω for different values of q.

The electric conductivity tensor $\sigma_{\alpha\beta}\left(\mathbf{q},\omega\right)$ of Eq. (58) gives access to different electric response functions. The current-current correlation function is given by

$$\chi_{J_\alpha J_\beta}\left(\mathbf{q},\omega\right)=-i\omega\sigma_{\alpha\beta}\left(\mathbf{q},\omega\right), \tag{59}$$

where α, β denote the components of the current vector (see e.g. Ref. [84]). With the help of the continuity equation, the charge density-density correlation function is obtained from Eq. (59):

$$\begin{aligned} \chi_{\rho\rho}\left(q,\omega\right) &= \frac{q_\alpha q_\beta}{\omega^2}\chi_{J_\alpha J_\beta}\left(\mathbf{q},\omega\right). \\ &= \frac{q^2}{i\omega}\sigma_{\parallel}\left(q,\omega\right). \end{aligned} \tag{60}$$

The non-local conductivity is related to the dielectric constant $\varepsilon\left(\mathbf{q},\omega\right)$ which is defined as (see Eq. (13))

$$\varepsilon=\frac{\varphi_{\text{ext}}}{\varphi_{\text{tot}}}. \tag{61}$$

Observing that $\varphi_{\text{ind}}\left(\mathbf{q},\omega\right)=V\left(q\right)\delta\rho\left(\mathbf{q},\omega\right)$, where $\delta\rho$ is the induced charge density, we find

$$\varepsilon=1-V\left(q\right)\frac{\delta\rho}{\varphi_{\text{tot}}}. \tag{62}$$

In linear response it is $\delta\rho=\chi_{\rho\rho}\left(q,\omega\right)\varphi_{\text{tot}}$, so that we can write

$$\varepsilon=1-V\left(q\right)\chi_{\rho\rho}. \tag{63}$$

Taking the divergence of Ohm's law $j_\alpha\left(\mathbf{q},\omega\right)=\sigma_{\alpha\beta}\left(\mathbf{q},\omega\right)E_\beta\left(\mathbf{q},\omega\right)$, and using the continuity equation $i\omega\delta\rho=iq_\alpha j_\alpha$ to express the electric current in terms of the induced charge density, we obtain

$$\varphi_{\text{tot}}=\frac{i\omega\delta\rho}{q^2\sigma_{\parallel}}.$$

Inserting in Eq (62) we have

$$\varepsilon\left(\mathbf{q},\omega\right)=1-V\left(q\right)\frac{iq^2}{i\omega}\sigma_{\parallel}\left(\mathbf{q},\omega\right),$$

which is in accordance with Eq. (60). Notice, that both the longitudinal conductivity $\sigma_{\parallel}$ and the charge susceptibility $\chi_{\rho\rho}$ describe the response to the total potential φ_{tot}. Hence the Vlasov term does not enter these quantities explicitely (for an in-depth discussion see Ref. [101], Chapter 3, in particular Eq. (3.56)) Finally, the charge compressibility $K=\partial\rho/\partial\mu$ is given by

$$K\left(q\right)=\chi_{\rho\rho}\left(\omega=0\right). \tag{64}$$

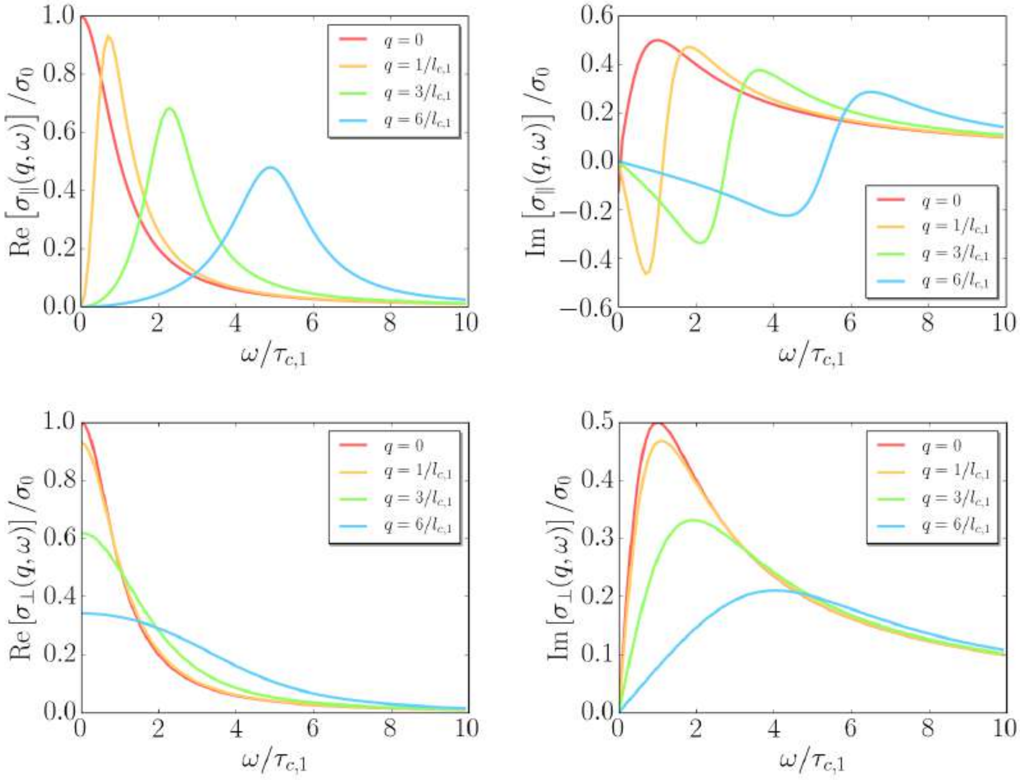


Figure 3: Longitudinal (upper row) and transverse (lower row) electric conductivities of charge neutral graphene as functions of the electric field frequency ω as given by Eqs (58). Different colors indicate different values of the wavenumber q. Frequencies and wave-numbers are normalized to the characteristic scattering times and lengths $\tau_{c,1}$, $l_{c,1} = v\tau_{c,1}$. σ_0 is the interaction induced conductivity at the neutrality point [11, 12]. The graphs show distinct resonant features at frequencies $\omega \sim q/v$, where v is the electron group velocity. Whereas the real part of the longitudinal and the imaginary part of the transverse conductivities are peaked around $\omega \sim q/v$, the imaginary part of the longitudinal conductivity exhibits a sign change indicating an abrubt phase change of the current response. The real parts approach σ_0 for $q \to 0$, $\omega \to 0$. For $q \neq 0$, $\omega = 0$ the longitudinal conductivity vanishes. This general property of the charge conductivity follows from the conservation of charge (see Eq. (60)).

The role of interaction effects for the compressibility were discussed in Ref.[17].

2. *Thermal conductivity*

Next we presemt our analysis for the non-local thermal conductivity. Since momentum conservation implies for a Dirac fluid the conservation of the heat current, thermal transport is expected to display classical hydrodynamic behavior, i.e. one expects non-local effects to be even more important than for charge transport.[14, 84].

As pointed out in Sec. IV B, the $s = 3$ energy mode must be kept in the calculation of the thermal conductivity, whereas the $s = 2$ imbalance mode can be neglected, contributing only a small correction to the overall result. With only a single mode involved, the calculation is formally analogous to the calculation of the electrical conductivity in Sec. IV C 1, even though there are crucial differences in the actual result, given the distinct role of momentum conservation. The relaxation time $\tau_{c,m}$ must be replaced by $\tau_{\varepsilon,m}$ as given by Eq. (47). The conservation of momentum is incorporated via $\tau_{\varepsilon,1} \to \infty$, whivch follows from the Boltzmann approach. The resulting longitudinal and transverse thermal conductivities read

$$\begin{aligned}\kappa_\parallel(q,\omega) &= \frac{\kappa_0}{i\omega\tau_{\varepsilon,2} - \frac{1}{4}v^2q^2\tau_{\varepsilon,2}\left(\frac{2i}{\omega} - \frac{1}{M_\varepsilon(q,\omega)+i\omega}\right)}\\ \kappa_\perp(q,\omega) &= \frac{\kappa_0}{i\omega\tau_{\varepsilon,2} + \frac{\frac{1}{4}v^2q^2\tau_{\varepsilon,2}}{M_{\varepsilon,2}(q,\omega)+i\omega}},\end{aligned} \tag{65}$$

with the memory function

$$M_\varepsilon(q,\omega) = \tau_{\varepsilon,2}^{-1} + \frac{1}{2}vq\frac{\mathrm{I}_{3+\frac{\eta_\varepsilon}{\gamma_\varepsilon}+i\omega\tau_\varepsilon}(\tau_\varepsilon vq)}{\mathrm{I}_{2+\frac{\eta_\varepsilon}{\gamma_\varepsilon}+i\omega\tau_\varepsilon}(\tau_\varepsilon vq)}.$$

The abbreviation $\tau_{\varepsilon,2} = \frac{1}{\alpha^2}\frac{\hbar}{k_BT}\frac{9\zeta(3)}{3.341\cdot 2\pi}$ is used. For convenience $\kappa_{\parallel/\perp}$ is given in units of a thermal conductivity $\kappa_0 = 9N\pi^3k_B\zeta(3)\tau_{\varepsilon,2}/2\beta^2\hbar^2$, however, $\tau_{\varepsilon,2}$ is the relaxation time in the $|m| = 2$ channel, and should not be confused with an alleged relaxation time of the energy

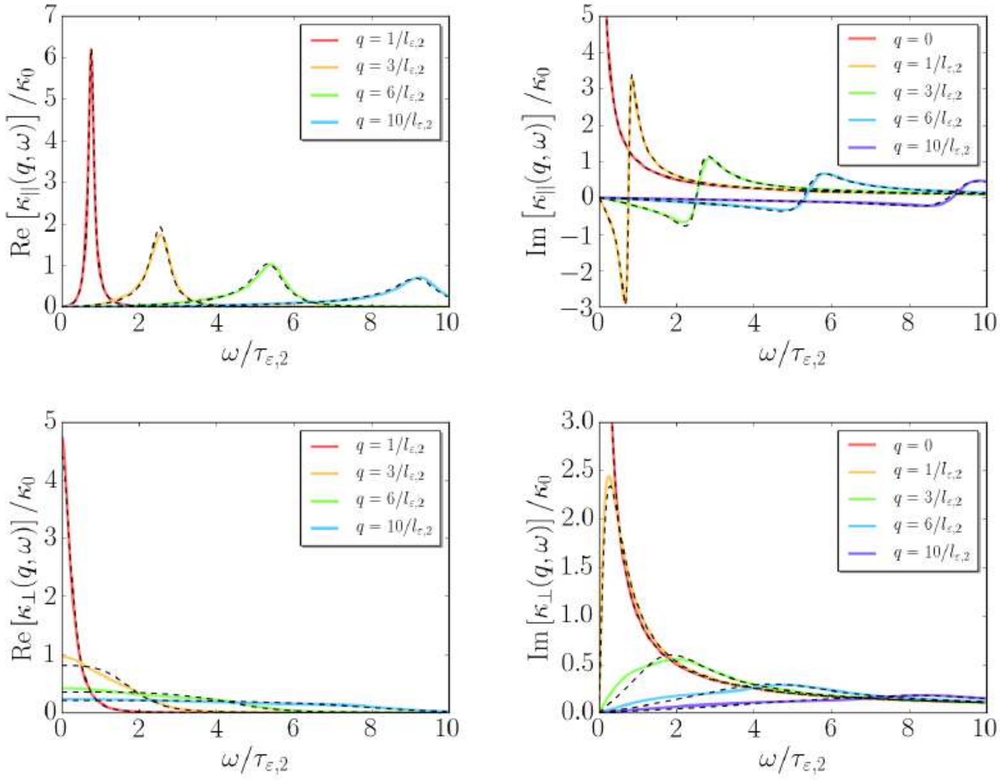


Figure 4: The figure shows the longitudinal (upper row) and transverse (lower row) thermal conductivities (58) as functions of the electric field frequency ω. Different colors indicate different values of the wavenumber q. The conductivities are normalized to $\kappa_0 = 9N\pi^3 k_B \zeta(3)\tau_{\varepsilon,2}/2\beta^2\hbar^2$. For small ω and vanishing q, the imaginary part of $\kappa_{\parallel/\perp}$ diverges as $1/\omega$, whereas the real part vanishes - a behavior indicating that thermal transport in the system is ballistic. The solid lines show the analytical result of Eq. (65), the dashed lines show the full numerical result including all modes and the exact scattering times.

current, which is infinite due to the conservation of momentum.

In Fig. 3 the thermal conductivities are plotted as functions of ω for different values of q. The fact that thermal currents are protected by momentum conservation leads to a divergence of the thermal conductivity at small frequencies: for $q = 0$, κ is purely imaginary and shows the characteristic $1/\omega$ Drude behavior.

3. Non-local shear viscosity

The non-local viscosity is defined through a constitutive relation of the form of Eq. (49), linking the shear force $X_{0,\alpha\beta}(\mathbf{r}')$ to the momentum-current tensor $\tau_{\alpha\beta}$:

$$\tau_{\alpha\beta}(\mathbf{r},t) = \int d^2r' \int dt'\, \eta_{\alpha\beta\gamma\delta}(\mathbf{r}-\mathbf{r}', t-t')\, X_{0,\gamma\delta}(\mathbf{r}', t')\,. \tag{66}$$

Since the system is isotropic, the shear force can be chosen such that the flow velocity is aligned with the x-axis, and its gradient shows in the y direction. It is assumed that the shear force is wavelike: $X_{0,xy}(\mathbf{r}) = X_{0,xy}e^{i\mathbf{q}\cdot\mathbf{r}-i\omega t}$. The wave-vector $\mathbf{q}$ can have an arbitrary direction in the xy-plane, introducing a preference direction to the system's response. In addition to τ_{xy}, this gives rise to nonzero components τ_{xx}, τ_{yy}, if $\mathbf{q}$ does not align with the x or the y-axes. The viscosity tensor $\eta_{\alpha\beta xy}$ can be decomposed into transverse and longitudinal parts (see Eq. (57)) analogously to the electric and charge conductivities. Because $\eta_{\alpha\beta xy}$ is a fourth rank tensor the decomposition is slightly more involved and the reader is referred to Appendix D for details. The general $\mathbf{q}$-dependent viscosity tensor can be constructed with the help of three rank two tensors:

$$\begin{aligned} e^{(1)}_{\alpha\beta} &= \frac{q_\alpha q_\beta}{q^2} \\ e^{(2)}_{\alpha\beta} &= \delta_{\alpha\beta} - \frac{q_\alpha q_\beta}{q^2} \\ e^{(3)}_{\alpha\beta} &= \frac{1}{\sqrt{2}}\left(q_\alpha p_\beta + p_\alpha q_\beta\right)/(pq)\,, \end{aligned} \tag{67}$$

where

$$p_\alpha = q_\gamma \varepsilon_{\gamma\alpha}. \tag{68}$$

The viscosity tensor is parameterized by two frequency and momentum dependent functions, $\eta_{\parallel}(\mathbf{q},\omega)$ and $\eta_{\perp}(\mathbf{q},\omega)$, which we will call longitudinal and transverse viscosities:

$$\eta_{\alpha\beta\gamma\delta}(\mathbf{q},\omega)=\eta_1(\mathbf{q},\omega)\left(e^{(1)}_{\alpha\beta}e^{(1)}_{\gamma\delta}+e^{(2)}_{\alpha\beta}e^{(2)}_{\gamma\delta}\right)$$
$$+\,\eta_2(\mathbf{q},\omega)\,e^{(3)}_{\alpha\beta}e^{(3)}_{\gamma\delta}.$$

Let the flow be in x-direction: $\mathbf{u}(y)=u(y)\,\hat{\mathbf{e}}_x$, and let the wave-vector be parameterized by $\mathbf{q}=q\left(\cos(\vartheta_{\mathbf{q}}),\sin(\vartheta_{\mathbf{q}})\right)^T$, where θ is measured with respect to the x-axis. For $\vartheta_{\mathbf{q}}=0$ or $\vartheta_{\mathbf{q}}=\pi/2$ follows $e^{(1,2)}_{\alpha\beta}=0$, $\eta_{xxxy}=\eta_{xxyx}=0$ and $\eta_{xyxy}=\eta_2/2$. This corresponds to the familiar shear flow in e.g. a Poiseuille geometry where $\tau_{xx}=\tau_{yy}=0$. The momentum current flows orthogonal to the direction of the momentum density. For $\vartheta_{\mathbf{q}}=\pi/4$, the viscosity is determined by η_1: $\eta_{xyxy}=\eta_1/2$.

As in the case of thermal conductivity, dropping the $s=2$ imbalance mode produces only a small numerical correction in the final result for the viscosity. With an external shear force of the form of Eqs. (23), (38) applied to the system, the kinetic equation can be written as 5×5 component matrix equation, similar to the case of an applied electric field (see Eq. (54)). The force acts in the $|m|=2$ channels, and the equation reads

$$\begin{bmatrix} -i\omega+M_\varepsilon(q,\omega) & \frac{1}{2}ivqe^{i\theta} & 0 & 0 & 0\\ \frac{1}{2}ivqe^{-i\theta} & -i\omega & \frac{1}{2}ivqe^{i\theta} & 0 & 0\\ 0 & \frac{1}{2}ivqe^{-i\theta} & -i\omega & \frac{1}{2}ivqe^{i\theta} & 0\\ 0 & 0 & \frac{1}{2}ivqe^{-i\theta} & -i\omega & \frac{1}{2}ivqe^{i\theta}\\ 0 & 0 & 0 & \frac{1}{2}ivqe^{-i\theta} & -i\omega+M_\varepsilon(q,\omega)\end{bmatrix}\begin{bmatrix}a_{-2}\\ a_{-1}\\ a_0\\ a_1\\ a_2\end{bmatrix}=\begin{bmatrix}-\frac{iX_0}{4}\\ 0\\ 0\\ 0\\ \frac{iX_0}{4}\end{bmatrix}. \tag{69}$$

Solving the matrix equation (69) for $a_{\pm2}$, the viscosity is calculated with the help of Eq. (66) which takes the form $\tau_{xy}=N\sum_\lambda\int_k v_xk_yf_{\mathbf{k},\lambda}=\eta_{xyxy}X_{0,xy}$. As explained above, the viscosity components η_1 and η_2 can be read off from the general result $\eta_{xyxy}\left(\mathbf{q}=q\left(\cos(\vartheta_{\mathbf{q}}),\sin(\vartheta_{\mathbf{q}})\right)^T,\omega\right)$ by setting $\vartheta_{\mathbf{q}}=0$ and $\vartheta_{\mathbf{q}}=\pi/2$:

$$\begin{aligned}\eta_1(q,\omega)&=\frac{2\eta_0}{-i\tau_{\varepsilon,2}\omega-q^2v^2\frac{i\tau_{\varepsilon,2}\omega}{2q^2v^2-4\omega^2}+\tau_{\varepsilon,2}M_\varepsilon(q,\omega)},\\ \eta_2(q,\omega)&=\frac{2\eta_0}{-i\tau_{\varepsilon,2}\omega-\frac{q^2v^2\tau_{\varepsilon,2}}{4i\omega}+\tau_{\varepsilon,2}M_\varepsilon(q,\omega)}.\end{aligned} \tag{70}$$

Here, η_0 is the viscosity at $q=0$, $\omega=0$, $\eta_0=N\left(k_BT\right)^3\tau_{\varepsilon,2}/\left(8\hbar^2v^2\right)$, as it was first calculated in Ref. [13] including both modes, $s=2$ and $s=3$.

V. COLLECTIVE MODES

Collective modes are solutions to the homogeneous part of the kinetic equation (12), (16) (see e.g. [39]). Consider Eq. (16). With the force terms set to zero it holds

$$(\mathcal{L}+\mathcal{V}+\mathcal{C})\,\psi=0.$$

Here, $\mathcal{L}$ and $\mathcal{C}$ have are the matrix operators of Eqs. (33) and (35). Solutions to this equation exist only if

$$\det(\mathcal{L}+\mathcal{V}+\mathcal{C})=0 \tag{71}$$

holds. This is only the case for certain values of the variable pairs ω, q. Eq. (71) is an eigenvalue problem where the eigenvalues $\omega(q)$ determine the dispersion relations of the collective modes. On the other hand, collective modes can be found from poles of response functions for an external force S. The two methods are equivalent. Within the kinetic equation formalism, response functions are calculated as averages over the distribution function $\psi=(\mathcal{L}+\mathcal{V}+\mathcal{C})^{-1}S$. If the condition (71) is fulfilled, the operator $(\mathcal{L}+\mathcal{V}+\mathcal{C})^{-1}$ is singular and thus singularities in the response to S appear. We will use Eq. (71) to study the collective modes of a Dirac fluid on an infinite domain.

As in the previous sections, the kinetic equation will be expanded in terms of collinear zero modes (32): $\chi^{(m,s)}_{\mathbf{k},\lambda}=\lambda^me^{im\theta}\{1,\lambda,\lambda\beta v\hbar k\}$. For $m=0$ these modes correspond to excitations of the charge, imbalance and energy densities; for $|m|=1$ they correspond to the associated currents. At the end of this section it will be shown that including non-collinear zero modes in the calculation does not change the result as long as the fine structure constant α is kept small.

To get a feeling for the structure of collective modes in the system, it is useful to begin with the case $q=0$. In the subspace of collinear zero modes, the kinetic equation reduces to Eq. (43) and the condition (71) reads

$$\det\left(-i\omega\delta_{s,s'}+\Gamma^{s,s'}_m\right)=0. \tag{72}$$

This is an eigenvalue equation for the frequencies of collective modes that can be solved independently for any m. Since, as pointed out in Sec. IV A, $\Gamma^{s,s'}_m$ is block-diagonal in the subspaces of electric ($s = 1$) and imbalance/energy ($s = 2, 3$) excitations, the above equation, as well as its extension to $q \neq 0$, can be solved independently in these two sectors. For $s = 1$, the eigenfrequencies are $\omega_m(q=0) = -i/\tau_{c,m}$. Since in this scenario the time evolution of the modes is given by the factor $e^{-i\omega_m t}$, all but the $m = 0$ mode, which is protected by charge conservation, exponentially decay at a rate inversely proportional to their scattering time. The $m = 0$ zero mode corresponds to the charge density, which is conserved, and therefore does not decay. In the following two sections, the collective charge, as well as energy and imbalance excitations will be described at finite $\mathbf{q}$. Figs. 5, 6, 7, 8, 9 show the dispersion relations of these modes.

A. Collective charge excitations

In general, conserved modes do not decay at $q = 0$, and therefore their dispersion relations must vanish in a spatially homogeneous system. The only conserved mode in the charge sector is the charge density mode $\chi_{\mathbf{k},\lambda}^{(m=0,s=1)} = 1$. In the limit $q \ll v\tau_{c,1}$, the memory matrix (55) reduces to $M_c(q,\omega) \approx \tau_{c,2}^{-1}$ and Eq. (71) can be solved analytically. The dispersions of the two lowest modes are

$$\omega_{\text{charge diff.}} \approx \omega_\pm = -\frac{i}{2\tau_{c,1}} \pm \sqrt{\frac{vq}{\tau_V} - \frac{1}{4\tau_{c,1}^2}}. \tag{73}$$

The conserved charge density mode is described by ω_-. The dispersion relations of Eq. (73) have a non-vanishing real part for

$$q > q_{pl}^* = \frac{\tau_V}{4v\tau_{c,1}^2}. \tag{74}$$

For wave-vectors below q_{pl}^*, the plasmon is over-damped (see Fig. 5). However, we have $vq_{pl}^* \sim \alpha^3 k_B T/\hbar$ such that the plasmon mode becomes more and more pronounced at low temperatures.

The plasmon mode is gapped out due to the intrinsic interaction induced resistivity. At $q = 0$ it has a vanishing real part and its decay rate is given by the scattering rate in the $m = 1$ channel:

$$\omega_{pl}(q \to 0) = -i/\tau_{c,1} \tag{75}$$

(see also [26]). It is the most weakly damped of an infinite set of modes corresponding to higher angular harmonics (see Fig. 5). It is clearly seen, that the modes relate to different angular harmonic channels m. For $q = 0$ their dispersions approach $\omega_m(q=0) = -i/\tau_{c,m}$. Such modes play a crucial role in the relaxation mechanism of focused current beams in graphene [1]. Similar collective modes have been argued to influence the relaxation behavior of unitary fermi gases [7] and QCD plasmas [8–10].

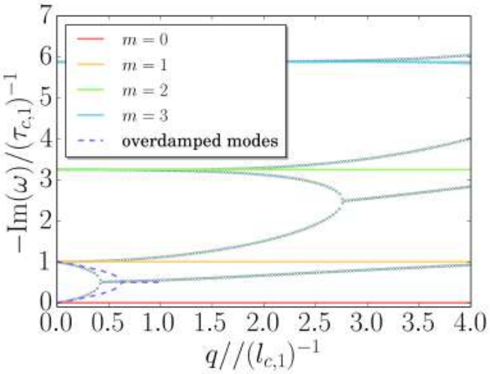


Figure 5: The imaginary parts of the dispersion relations of collective charge excitations in different angular harmonic channels m are shown. The wave-vector q is given in units of the inverse scattering length $v\tau_{c,1}^{-1}$. The grey symbols correspond to the numerical solution of Eq. (71). The purely imaginary $m = 0$ diffusive mode is the only mode approaching zero for small q - a behavior necessitated by charge conservation. Modes with a higher m are damped and approach the values $-i/\tau_{c,m}$ for $q \to 0$. The corresponding excitations decay even in the absence of spatial inhomogeneities. At a value $q = q_{pl}^*$ (Eq. 74), the dispersions of the diffusive mode and the $m = 1$ excitation merge, giving rise to a plasmon mode, which has a finite real part (see Fig 6). This value is slightly overestimated by the simplified expression of Eq. (74).

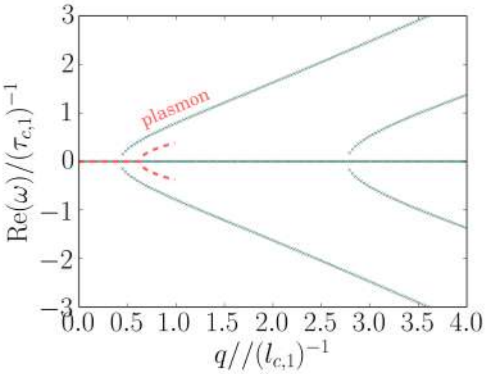


Figure 6: The figure shows the real parts of the dispersion relations of collective charge excitations in different angular harmonic channels m. The wave-vector q is given in units of the inverse scattering length $v\tau_{c,1}^{-1}$. The grey symbols correspond to the numerical solution of Eq. (71). The plasmon mode is gapped out by the interaction induced conductivity and only obtains a finite real part around $q = q_{pl}^*$ (the simplified value of q_{pl}^* given in Eq. (74) (red dashed line) overestimates the branching point). At higher q, other, strongly damped modes corresponding to higher angular harmonics appear. The dampings of these modes are given by the $m > 1$ modes of Fig. 5.

B. Collective energy and imbalance excitations

In the energy sector spanned by the modes $s = 2, 3$, the Eqs. (71) and (72) give rise three zero eigenvalues. These correspond to the conserved energy ($\chi_{\mathbf{k},\lambda}^{(m=0,s=3)} = \lambda\beta v\hbar k$) and quasiparticle (imbalance) densities ($\chi_{\mathbf{k},\lambda}^{(m=0,s=2)} = \lambda$), as well as momentum ($\chi_{\mathbf{k},\lambda}^{(m=1,s=3)} + (-)\,\chi_{\mathbf{k},\lambda}^{(m=-1,s=3)} = 2\,(i)\,\beta v\hbar k_{x(y)}$). The first two conservation laws lead to two diffusive modes. The conservation of momentum gives rise to second sound - ballistic thermal waves propagating through the two dimensional graphene plane [74]. This mode is the analogue of the density modes of a clean neutral Galilean invariant system.

Truncating the mode expansion of Eq. (71) at $m = 2$, which is a good approximation for low wave-numbers, yields the dispersions

$$\begin{aligned} \omega_{\text{heat diff.}} &\approx \frac{1}{4} v^2 q^2 \tau_{\varepsilon,2}, \\ \omega_{\text{qp diff.}} &\approx \frac{1}{8} v^2 q^2 \tau_{\varepsilon,2}, \end{aligned} \tag{76}$$

for the heat and quasiparticle (imbalance) diffusion modes, respectively. The second sound dispersion is given by

$$\omega_{\text{sec. sound}} \approx \frac{vq}{\sqrt{2}} + i\tau_{\varepsilon,2}\frac{v^2 q^2}{8}. \tag{77}$$

Second sound mediated by phonons has been previously observed in solids [102] and had a velocity comparable to the velocity of sound. Here, the second sound is carried by electrons and propagates with a velocity $v_0/\sqrt{2}$. The above dispersion relations are shown in Figs. 7 and 9.

The dispersion of the quasiparticle diffusion mode and the imaginary part of the second sound dispersion merge at low wave-numbers. As in the case of charge excitations, there exists an infinite number of damped modes associated with scattering in higher angular harmonic channels. These modes are depicted in Figs. 8 and 9. Note, that modes associated with imbalance excitations ($s = 2$) are damped stronger by an order of magnitude as compared to energy excitations ($s = 3$).

C. Validity of the collinear zero mode approximation for collective modes

The discussion so far was carried out in the restricted subspace of collinear zero modes. In this section it is shown that the results for collective excitations obtained within the restricted subspace remain valid, if this restriction is lifted, and non-collinear zero modes are added. These modes introduce large corrections to the matrix of scattering rates $\Gamma_m^{s,s'}$, and it is not obvious that they can be neglected. It is sufficient to consider the $\mathbf{q} = 0$ case. The extension to finite wave-numbers is straightforward.

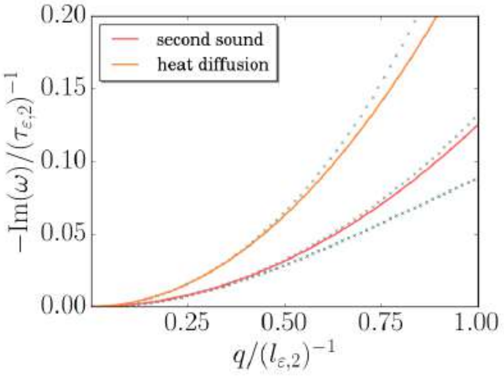


Figure 7: The figure shows the imaginary part of the dispersion relations of second sound, heat diffusion, and quasiparticle (imbalance) diffusion excitations. The wavevector q is given in units of the inverse scattering length $v\tau_{c,2}^{-1}$. The grey symbols correspond to the numerical solution of Eq. (71). The damping of second sound is due to scattering in the $m = 2$ channel and follows the dispersion $\mathrm{Im}\,(\omega_{\text{sec. sound}}) \approx \frac{1}{8} v^2 q^2 \tau_{\varepsilon,2}$ (red curve). For small q the imaginary part of the second sound dispersion and the dispersion of the quasiparticle diffusion mode merge. A third diffusive mode corresponds to the diffusion of heat (orange curve).

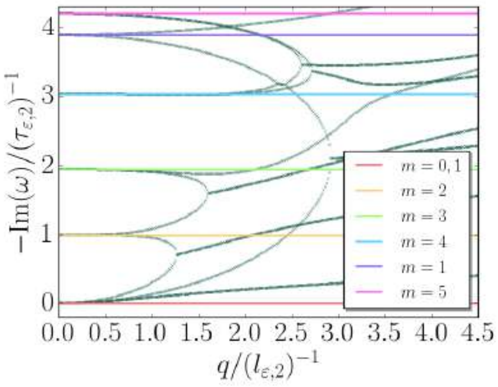


Figure 8: The imaginary part of the dispersion relations of collective charge excitations in different angular harmonic channels m are shown. The wave-vector q is given in units of the inverse scattering length $v\tau_{c,2}^{-1}$. The grey symbols correspond to the numerical solution of Eq. (71). For small q, the modes approach values given by the scattering rates $-i/\tau_{c,m}$ and are thus strongly damped. At larger values of q, the dispersions tend to merge in a complex fashion. Fig. 7 shows the weakly damped modes (second sound and diffusive modes) for small values of q.

The scattering rate matrix $\Gamma_m^{s,s'}$ of Eq. (44) is extended to include modes that are not collinear zero modes, which are labeled with indices $s > 3$. It is useful to define the

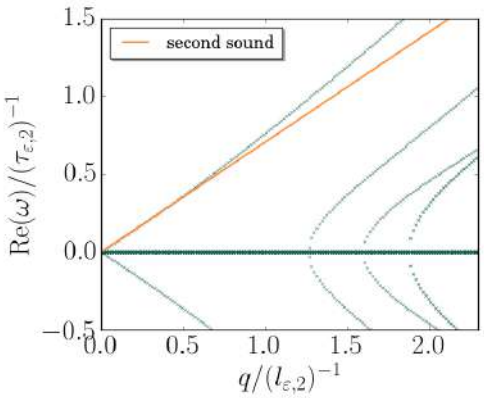


Figure 9: The real parts of the dispersion relations of collective energy and imbalance excitations are depicted. The grey symbols correspond to the numerical solution of Eq. (71). The linear dispersion of the second sound mode given by $vq/\sqrt{2}$ for small q is shown in orange color. The wave-vector q is given in units of the inverse scattering length $v\tau_{c,2}^{-1}$.

following matrices

$$\begin{aligned}
S &= (v\beta\hbar)^2 \left\langle \chi_{\mathbf{k},\lambda}^{(s<3)} \left|\mathcal{C}\right| \chi_{\mathbf{k},\lambda}^{(s'<3)} \right\rangle \\
P &= (v\beta\hbar)^2 \left\langle \chi_{\mathbf{k},\lambda}^{(s>3)} \left|\mathcal{C}\right| \chi_{\mathbf{k},\lambda}^{(s'<3)} \right\rangle \\
Q &= (v\beta\hbar)^2 \left\langle \chi_{\mathbf{k},\lambda}^{(s<3)} \left|\mathcal{C}\right| \chi_{\mathbf{k},\lambda}^{(s'>3)} \right\rangle \\
R &= (v\beta\hbar)^2 \left\langle \chi_{\mathbf{k},\lambda}^{(s>3)} \left|\mathcal{C}\right| \chi_{\mathbf{k},\lambda}^{(s'>3)} \right\rangle .
\end{aligned}$$

Here, $\chi_{\mathbf{k},\lambda}^{(s<3)}$ are the familiar collinear zero modes (32). $\chi_{\mathbf{k},\lambda}^{(s>3)}$ are modes with a different $|\mathbf{k}|$-dependence, such that the full set of modes forms a complete basis. Since $\mathcal{C}$ is Hermitian, we have $Q = P^T$. The mode expansion of the Liouville operator $L_{s,s'}$ of Eq. (33) also has to be enlarged by the $s > 3$ modes. However, we do not need to know the precise values of the corresponding elements of L. The eigenvalue equation (72) reads

$$\det\left(-i\omega L - F\right) = 0, \tag{78}$$

where F is the composite matrix

$$F = \begin{bmatrix} S & P \\ P^T & R \end{bmatrix}.$$

In the following, the Liouville matrix L will also be separated into blocks corresponding the same subspaces: $L = \left((L_S, L_P), (L_P^T, L_S)\right)$. It follows from Eq. (27) and the Hermiticity of the collinear part of the collision operator $\mathcal{C}_c$ that

$$\begin{aligned}
S &\sim P \sim 1 \\
R &\sim \log\left(1/\alpha\right),
\end{aligned}$$

meaning that non collinear zero modes are scattered faster by a factor of $\log(1/\alpha)$. The determinant can be found using the block matrix identity

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(D)\det\left(A - BD^{-1}C\right). \tag{79}$$

Applying this identity to Eq. (78) and noticing that for $\alpha \to 0$ the inverse matrix in the last determinant vanishes, one has

$$\det\left(i\omega L + F\right) \approx \det\left(i\omega L_R + R\right)\det\left(i\omega L_S + S\right).$$

Eq. (78) therefore separates into two independent parts: $\det(i\omega L_R + R) = 0$ and $\det(i\omega L_S + S) = 0$. The second equation is equivalent to the eigenvalue equation (72). In the limit of a small fine structure constant, the weakly damped collective modes can therefore be found by solving the kinetic equation in the restricted subspace of collinear zero modes, even if there is significant coupling between all modes.

VI. SURFACE ACOUSTIC WAVES

The longitudinal electrical conductivity $\sigma_{\parallel}$ is accessible through experiments with surface acoustic waves (SAWs) [70]. The simplest setup to measure $\sigma_{\parallel}$ is a sheet of graphene placed on top of a piezoelectric material. Using interdigital transducers, SAWs are induced in the piezoelectric. The real part of $\sigma_{\parallel}$ then determines the damping of the SAWs, while the imaginary part changes the SAW velocity v_s. Overall, for a small piezoelectric coupling the change of the SAW velocity Δv_s, where the imaginary part describes the damping, can be written as [67, 68]

$$\frac{\Delta v_s}{v_{s,0}} = p_e \frac{1}{1 + i\frac{\sigma_{\parallel}}{\sigma_M}}. \tag{80}$$

Here $p_e < 1$ is an effective coupling constant, and a σ_M a reference conductivity. Both p_e and σ_M depend on material parameters of the piecoelectric. A rough estimate for σ_M is given by $\sigma_M \approx v_s \epsilon_{\mathrm{eff}}$ [67, 68], where ϵ_{eff} is the effective permittivity at the surface of the piezoelectric. There has been experimental work on the coupling between SAWs and graphene [103, 104]. LiNbO$_3$ seems to be a suitable piezoelectric for such experiments[103], because it provides a relatively large coupling parameter $p_e \approx 0.03$ [71]. While there might be better choices for the piezoelectric material, here we consider LiNbO$_3$, since the feasibility of a graphene-LiNbO$_3$ device has been demonstrated in Ref. [103]. The SAW velocity is $v_s \approx 4 \cdot 10^3$ m/s and the effective dielectric constant is given by $\epsilon_{\mathrm{eff}} \approx 0.5\epsilon_0 \left(\sqrt{\epsilon_{xx}^T \epsilon_{zz}^T} + 2\right) \approx 24\epsilon_0$ (assuming that the dielectric constant above the graphene sheet is ϵ_0). One then has

$$\sigma_M \approx 10^{-6}\,\mathrm{S}.$$

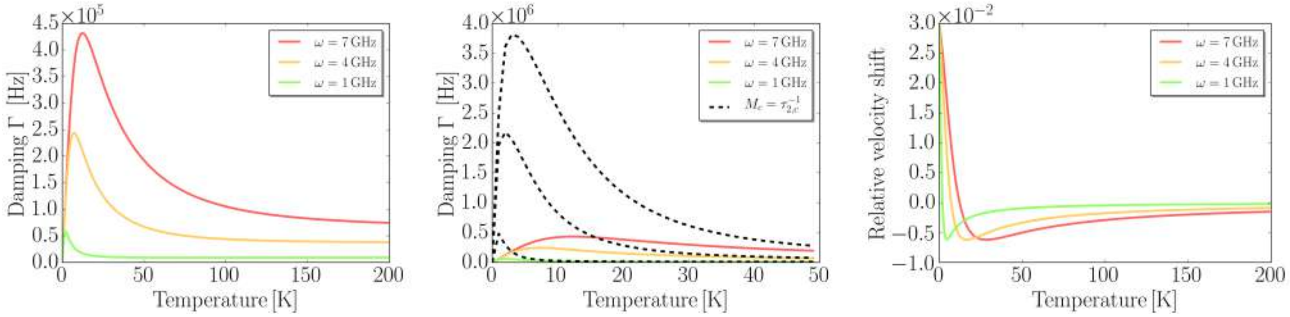


Figure 10: The Figure shows the damping coefficients and velocity shifts of $\mathrm{LiNbO_3}$ surface acoustic waves induced by a graphene sheet laying on top of the crystal. Due to the $\sim 1/T$ dependence of the scattering times $\tau_{c,m}$, changing the temperature alters the quantity $v\tau_{c,m}q$, where q is the SAW wave-vector, such that the functional dependence of $\sigma_{\parallel}\left(q,\omega=v_s q\right)$ can be investigated without switching the SAW frequencies. Here, v_s is the SAW velocity. Left figure: Damping coefficients of SAWs for three distinct frequencies. The induced damping is small (of the order of 1/cm). Middle figure: The damping coefficients at low temperatures depend sensitively on the scattering in higher angular harmonic channels. Setting $M_c = \tau_{c,2}^{-1}$, thus neglecting the scattering times $\tau_{c,m>2}$, raises the damping by an order of magnitude. Right figure: Upper right pannel: Relative velocity shifts $\Delta v_s/v_s$. The velocity shifts are large (on the order of 1%). This is a consequence of the mainly capacitive behavior of the graphene sheet at small frequencies (see main text).

The fine structure constant is small due to the large dielectric constant and renormalization effects. We estimate $\alpha \approx 0.1$. Here and in the following estimations, we assume a temperature of 50 K.

Interdigital transducers induce SAWs with sharply defined wave-vectors q_0. The frequency of the SAW ω_0 is given by

$$\omega_0 = v_s q_0.$$

ω_0 is much smaller than the characteristic hydrodynamic frequency for a wave-vector of the same magnitude $\omega_{\text{hydro}} \approx v q_0$, where $v \approx 10^6\,\text{m/s}$. It is

$$\frac{\omega_0}{\omega_{\text{hydro}}} \approx 0.005. \tag{81}$$

As shown in Fig. 3 the longitudinal conductivity $\sigma_{\parallel}$ is peaked around ω_{hydro} and vanishes in the limit $\omega = 0$, $q \to 0$. Therefore, SAW experiments are confined to a highly "off-resonant" regime due to the small ratio (81) and therefore cannot be large.

The damping coefficient is given by

$$\Gamma = -\omega \mathrm{Im}\left(\frac{\Delta v_s}{v_s}\right) = \omega p_e \frac{\mathrm{Re}\left(\sigma_{\parallel}\right)/\sigma_M}{1+\left|\frac{\sigma_{\parallel}}{\sigma_M}\right|^2}. \tag{82}$$

The relative velocity shift is

$$\mathrm{Re}\left(\frac{\Delta v_s}{v_s}\right) = p_e \frac{1+\mathrm{Im}\left(\sigma_{\parallel}\right)/\sigma_M}{1+\left|\frac{\sigma_{\parallel}}{\sigma_M}\right|^2}. \tag{83}$$

Since interdigital transducers excite SAWs of a fixed wavelength, altering q_0 is difficult. Instead, the q dependence of $\sigma_{\parallel}$ can be tested by varying the temperature, and thus the product of the wave vector and the scatterng length and $q_0 l_{c,m}$. Fig 10 shows the damping and the velocity shift induced by the graphene sheet as a function of temperature, according to Eqs. (82) and (83). As expected, the damping coefficients are very small, on the order of 10^5 Hz, corresponding to damping lengths of 1/cm. Such small damping are measurable in GaAs 2DEG structures [72], however they might be hard to observe with the more unconventional $\mathrm{LiNbO_3}$ device. On the other hand the low temperature (large q) behavior of the conductivity sensitively depends on the scattering rates in the higher angular harmonic channels (see lower left panel of Fig. 10), although the specific dependence $\tau^{-1}_{m>c,2} \sim |m|$ will be very hard to distinguish from e.g. constant scattering rates.. Finally we note, that here we considered the SAW response in the hydrodynamic regime $l_{c,1} \ll w$, where w is the sample size. For small sample sizes, the results will differ due to boundary scattering.

VII. POISEUILLE PROFILES

The wave-vector-dependence of transport coefficients is of importance when the currents in a system are spatially inhomogeneous, either because the applied fields are inhomogeneous, or because the inhomogeneity is imposed by the geometry of the system. The simplest example for the latter case is the Poiseuille flow. In undoped graphene, the energy current is conserved due to the conservation of momentum, however it is dissipated by the uneven boundaries of the sample [1]. In a Poiseuille geometry, which consists of an infinitely long, straight sample of width w, the boundaries slow down the current flow. The current profile becomes parabolic across the sample. On the other hand, charge currents decay in the bulk of undoped graphene due to the interaction induced resistivity. In this case, there exists a crossover from an almost flat current profile if $w \gg v\tau_{c,1}$ to a more parabola-like shape at $w < v\tau_{c,1}$. However, as shown in Ref. [1], the slowing down of the flow by the boundaries becomes inefficient when $w \lesssim v\tau_{c,2}$, again changing the profile. In this section we investigate the Poiseuille profiles of charge currents in undoped graphene using the full non-local conductivity (58).

A. Flow equations and boundary conditions

The thermal and charge flow is governed by the constitutive relations

$$\kappa^{-1}\left(\mathbf{q},\omega\right)_{\alpha\beta} j_{\varepsilon,\beta} = -\partial_\alpha T \tag{84}$$

and

$$\sigma^{-1}\left(\mathbf{q},\omega\right)_{\alpha\beta} j_{c,\beta} = E_\alpha, \tag{85}$$

where $j_{\varepsilon,\beta}$ is the thermal current and $j_{c,\beta}$ the electric current. With the thermal and electric conductivities κ and σ depending on the wave vector $\mathbf{q}$, these equations can be seen as Fourier transforms of differential equations. Similar equations have been studied to describe non-localities induced by vortices in type II superconductors [105]. The temperature gradient $-\partial_\alpha T$ and the electric field E_α act as source terms. In a Poiseuille geometry, the force fields act perpendicular to the gradient of the flow velocity, i.e. it is $\mathbf{E} \perp \mathbf{q}$, $\nabla T \perp \mathbf{q}$. Therefore, the currents are determined by the transverse conductivities. Let the sample be oriented in y-direction and centered around $x = 0$. The equations then read

$$\kappa_T^{-1}\left(q_x,\omega\right) j_{\varepsilon,y}\left(q_x,\omega\right) = -\partial_y T, \tag{86}$$

$$\sigma_T^{-1}\left(q_x,\omega\right) j_{c,y}\left(q_x,\omega\right) = E_y. \tag{87}$$

To solve the above equations, boundary conditions at the sample boundaries at $\pm w/2$ are needed. As discussed in Ref. [97], partial slip boundary conditions are appropriate:

$$j_{\varepsilon/c,y}\left(x = \pm w/2,\omega\right) = \mp\zeta \left.\frac{\partial j_{\varepsilon/c,y}}{\partial x}\right|_{x=\pm w/2}. \tag{88}$$

ζ is the so called slip length parameterizing the momentum charge (current) dissipation at the sample boundaries. If the boundaries are sufficiently rough, ζ is of the order of the mean free path associated with the $m = 2$ scattering time: $\zeta \sim v\tau_{\varepsilon/c,2}$. In principle, the Eqs. (86), (87) represent infinite order differential equations and require infinitely many boundary conditions. However, this problem does not appear explicitly in the calculation. The finite width of the sample w sets a natural cut-off for the wave-numbers q, and therefore only the low powers of q are relevant on the right hand side of Eqs. (86), (87). For simplicity, the boundary condition (88) is used, which is reasonable for not too small widths.

The Eqs. (86), (87) now can be solved by performing a Fourier transform. To fix the boundary conditions two point-like delta-function inhomogeneities are positioned at $\pm w$. In real space the equations take the form

$$\kappa_T^{-1}\left(\partial_x,\omega\right) j_{\varepsilon,y}\left(x,\omega\right) = -\partial_y T - \alpha\delta\left(x-w\right) - \beta\delta\left(x+w\right) \tag{89}$$

$$\sigma_T^{-1}\left(\partial_x,\omega\right) j_{c,y}\left(x,\omega\right) = E_y - \alpha\delta\left(x-w\right) - \beta\delta\left(x+w\right). \tag{90}$$

If the constants α, β are chosen such that Eq. (88) is satisfied, the solution inside the sample will be identical to the solution of the homogeneous equations with the matching boundary conditions.

Here, the profiles of electric current flows through samples of different widths will be calculated. Solving the Eq. (89) in Fourier space one obtains

$$j_{c,y}\left(q_x,\omega\right) = \left(2\pi E_y \delta\left(q_x\right) - \alpha e^{-iwq_x} - \beta e^{iwq_x}\right)\sigma_T\left(q_x,\omega\right). \tag{91}$$

Inserting this result into Eq. (88) gives two algebraic equations, from which α and β can be determined:

$$\zeta \int \frac{dq_x}{2\pi}\left(iq_x\right)\left(\alpha e^{-iq_x\frac{3w}{2}} + \beta e^{iq_x\frac{w}{2}}\right)\sigma_T\left(q_x,\omega\right) = \int \frac{dq_x}{2\pi}\left(\alpha e^{-iq_x\frac{3w}{2}} + \beta e^{iq_x\frac{w}{2}}\right)\sigma_T\left(q_x,\omega\right) - E_y\sigma_T\left(0,\omega\right)$$

$$\zeta \int \frac{dq_x}{2\pi}\left(iq_x\right)\left(\alpha e^{-iq_x\frac{w}{2}} + \beta e^{iq_x\frac{3w}{2}}\right)\sigma_T\left(q_x,\omega\right) = -\int \frac{dq_x}{2\pi}\left(\alpha e^{-iq_x\frac{w}{2}} + \beta e^{iq_x\frac{3w}{2}}\right)\sigma_T\left(q_x,\omega\right) + E_y\sigma_T\left(0,\omega\right).$$

The above integrals are calculated with the FFT algorithm. Once α, β are found, a Fourier transform the of the solution (91) gives the desired flow profiles.

Figs. 11 and 12 show the results for different widths w. For demonstration purposes no-slip boundary conditions ($\zeta = 0$) were assumed in Fig. 11. Here, for $w > v\tau_{c,1}$ the flow profile turns flat in the middle of the sample and steeply descends to zero at the boundaries (as necessitated by the no-slip boundary conditions). This behavior is due to the interaction-induced conductivity that dissipates current uniformly across the sample - at a distance $d > v\tau_{c,1}$ away from the boundary, a uniform

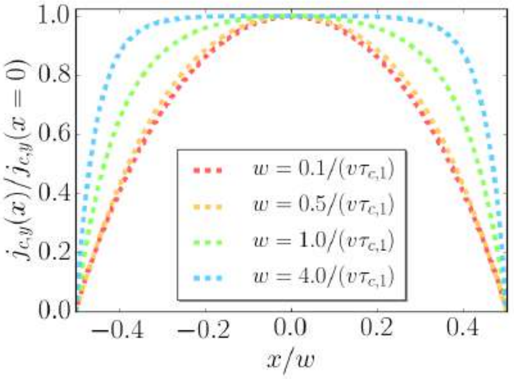


Figure 11: Poiseuille profiles of charge currents in undoped graphene samples of different widths w. Although physically incorrect, no-slip boundary conditions were assumed for clarity. The profiles are normalized to the current at $x = 0$. At large widths $w > v\tau_{c,1}$, the flow profiles turn flat. In the bulk they resemble Ohmic flow. For small widths $w < v\tau_{c,1}$, the momentum non-conserving scattering becomes inefficient. The electrons travel a distance corresponding to several width before loosing their momentum. Consequently, the profiles take a parabolic form, resembling classical Poiseuille flow. The profiles were calculated from Eq. (91).

flow is restored. On the other hand, for $w < v\tau_{c,1}$ the current-relaxing scattering processes in the $m = 1$ channel become less and less important. The scattering in the $m = 2$ channel dominates. It acts in the same way viscous forces act in ordinary flows. Current is transported from the middle of the sample, where it is maximal, to the sample edges, where it is dissipated. A finite slip length (as discussed, $\zeta = v\tau_{c,2}$ was chosen for simplicity) alters these results (see Fig. 12): Whereas for widths $w > v\tau_{c,1}$ the finite slip gives the current a non-negligible velocity at the sample boundary, for small widths $w < v\tau_{c,2}$, the flow profiles are rendered flatter, and the boundary effects become negligible. In the crossover region $w \sim v\tau_{c,1}$, the profiles are curved and resemble a parabola. This takes place around $w \sim 0.5 v\tau_{c,2}$ and is in accordance with the general expectations [97]: for $w < v\tau_{c,2}$ the quasi-viscous transport of currents from the middle of the sample towards the boundaries becomes inefficient, and the boundary does efficiently dissipate the current.

An interesting question is how the collective modes investigated in Sec. V are changed when the Dirac fluid is confined to a Poiseuille type sample with the boundary conditions of Eq. (88). For large sample sizes one can expect that e.g. the charge modes will exhibit a small correction of the order of $l_{c,1}/w$. The effects for small w should be more interesting. They are, however, beyond the scope of the present study.

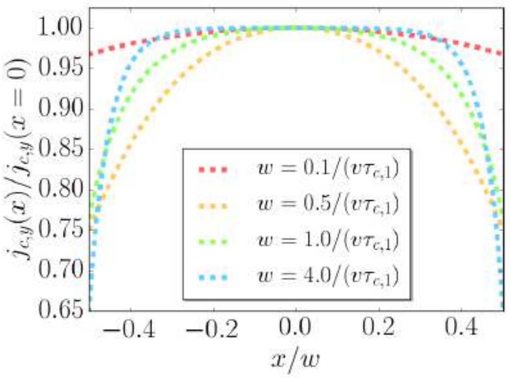


Figure 12: Poiseuille profiles of charge currents in undoped graphene samples of different widths w, normalized to the current at $x = 0$. Partial slip boundary conditions with a slip length $\zeta = v\tau_{c,2}$ were applied. At very small widths $w \ll \zeta$, boundary scattering ceases to be an efficient mechanism for the dissipation of electric current. The profiles turn flat, as they do in the nearly Ohmic regime $w > v\tau_{c,1}$. In the crossover regime at widths $w \sim 0.5 v\tau_{c,1}$, profile curvature is most pronounced. The profiles were calculated from Eq. (91).

VIII. CONCLUSION

In conclusion we have developed a kinetic theory of non-local charge and thermal transport in a clean Dirac fluid in the hydrodynamic regime. We obtained closed analytic expressions for the frequency and wave-vector-dependent, charge and thermal conductivities as well as the non-local viscosity due to electron-electron Coulomb interactions. Our solution is possible due to the dominance of so-called colinear zero modes. In the limit os a small fine-structure constant of graphene, all other mode relax more rapidly, limiting the phase space of the collective excitations that dominate the long-time dynamics. One aspect of the same physics, that was discussed previously by us in Ref.[1], is the onset of superdiffusion in phase space, where Lévy-flight behavior on the Dirac cone emerges. Frequent small angle scattering events are interrupted by rate large-angle scattering processes. We made specific predictions for measurements such as the velocity shift of surface acoustic waves and for inhomogeneous flow pattern. Those become identical to the one that follow from the solution of the Navier-Stokes equations in the long wavelength limit, but include higher order gradients that come into play as the sample geometry becomes smaller. In particular, we have demonstrated how the non-local transport coefficients determine the profiles of a hydrodynamic flow through narrow channels. In addition we determined the collective mode spectrum of the system including plasma waves and second sound like thermal waves. We find a complex structure of damped collective excitations. These excitations are similar to the so-called "non-hydrodynamic" modes that were shown to be relevant for the equilibration of other collission-domuinated quantum fluids [7] [8–10].

Acknowledgments

This work was supported by the European Commission's Horizon 2020 RISE program Hydrotronics (Grant No. 873028). We thank L. Levitov, A. Lucas, and J. F. Karcher for interesting discussions and P. Witkowski for drawing our attention to Refs. [7–10]. We are grateful to B. Jeevanesan for pointing out that the identity (79) provides a simple proof of the dominance of collinear zero modes in Sec. V C and to I. V. Gornyi for clarifying to us the role of the Vlasov term in the conductivity.

Appendix A: The collision operator

Transformed to the band basis, the interaction part of the Hamilton operator (8) reads

$$H_{\text{int}} = \frac{1}{2}\int_{k,k',q}\sum_{\alpha\beta} T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\psi^{\dagger}_{\lambda'}\left(\mathbf{k}+\mathbf{q},t\right)\psi^{\dagger}_{\mu}\left(\mathbf{k}'-\mathbf{q},t\right)\psi_{\mu'}\left(\mathbf{k}',t\right)\psi_{\lambda}\left(\mathbf{k},t\right) \tag{A1}$$

where the matrix elements $T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)$

$$T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = V\left(q\right)\left(U_{\mathbf{k}+\mathbf{q}}U^{-1}_{\mathbf{k}}\right)_{\lambda'\lambda}\left(U_{\mathbf{k}'-\mathbf{q}}U^{-1}_{\mathbf{k}'}\right)_{\mu\mu'}. \tag{A2}$$

U is the usual transformation from sub-lattice space to the band space (see Eq. 11). For the derivation of the quantum Boltzmann equation, the self energies $\Sigma^{\gtrless}_{\lambda}$ and the Green's functions $g^{\gtrless}_{\lambda'}$ are of interest (the small g is used for the Green's function transformed to the band basis $g^{\gtrless}\left(\mathbf{X},T;\mathbf{k},\omega\right) = U_{\mathbf{k}}G^{\gtrless}\left(\mathbf{X},T;\mathbf{k},\omega\right)U^{\dagger}_{\mathbf{k}}$, where $\left(\mathbf{X},T\right)$ are the center of mass coordinates, and $\left(\mathbf{k},\omega\right)$ are the relative coordinates after the Wigner transform). For details on the Wigner transform and the definitions of $G^{\gtrless}$, $\Sigma^{\gtrless}$ see e.g. [82, 83, 106]). The off diagonal elements of greens functions in band space can be neglected if the frequencies of interest are smaller than the energies of thermally excited particles: $\omega \ll k_B T$. In the following, only the weak space and time dependencies induced by external forces and represented by the center of mass coordinates will be of interest. For simplicity, the dependence on $\left(\mathbf{X},T\right)$ will be suppressed. The Green's functions $g^{\gtrless}_{\lambda'}\left(\mathbf{k},\omega\right)$ can be related to the distribution function:

$$\begin{aligned} g^{>}_{\lambda}\left(\mathbf{k},\omega\right) &= -i2\pi\delta\left(\omega-\varepsilon_{\lambda}\left(\mathbf{k}\right)-U_{\text{pot}}\right)\left(1-f_{\lambda,\mathbf{k}}\left(\omega\right)\right) \\ g^{<}_{\lambda}\left(\mathbf{k},\omega\right) &= i2\pi\delta\left(\omega-\varepsilon_{\lambda}\left(\mathbf{k}\right)-U_{\text{pot}}\right)f_{\lambda,\mathbf{k}}\left(\omega\right). \end{aligned} \tag{A3}$$

To second order in perturbation theory, for the self-energies

$$\begin{aligned} \Sigma^{\gtrless}_{\lambda}\left(\mathbf{k},\omega\right) &= N\sum_{\mu\mu'\lambda'}\int\frac{d^2q d^2k' d\omega_1 d\omega_2}{\left(2\pi\right)^6}\left|T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\ &\times\ g^{\gtrless}_{\lambda'}\left(\mathbf{k}+\mathbf{q},\omega_1\right)g^{\gtrless}_{\mu}\left(\mathbf{k}'-\mathbf{q},\omega_2\right)g^{\lessgtr}_{\mu'}\left(\mathbf{k}',\omega_1+\omega_2-\omega\right) \\ &-\ \sum_{\mu\mu'\lambda'}\int\frac{d^2q d^2k'}{\left(2\pi\right)^4}\int\frac{d\omega_1 d\omega_2}{\left(2\pi\right)^2}T_{\lambda\lambda'\mu'\mu}\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right)T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)^{*} \\ &\times\ g^{\gtrless}_{\lambda'}\left(\mathbf{k}+\mathbf{q},\omega_1\right)g^{\gtrless}_{\mu}\left(\mathbf{k}'-\mathbf{q},\omega_2\right)g^{\lessgtr}_{\mu'}\left(\mathbf{k}',\omega_1+\omega_2-\omega\right) \end{aligned} \tag{A4}$$

holds. $N=4$ accounts for the spin-valley degeneracy.

The collision operator, as it appears in Eq. (12), can now be determined from the self energies $\Sigma^{<}$ and $\Sigma^{>}$. It can then be written in terms of the distribution function $f_{\lambda}\left(\mathbf{k}\right)$:

$$\mathcal{C}_{\lambda}\left(\mathbf{k}\right) = -i\Sigma^{<}_{\lambda}\left(\mathbf{k},\varepsilon_{\lambda}\left(\mathbf{k}\right)\right)\left(1-f_{\lambda}\left(\mathbf{k}\right)\right) - i\Sigma^{>}_{\lambda}\left(\mathbf{k},\varepsilon_{\lambda}\left(\mathbf{k}\right)\right)f_{\lambda}\left(\mathbf{k}\right). \tag{A5}$$

The delta function $\delta\left(\omega-\varepsilon_{\lambda}\left(\mathbf{k}\right)-U_{\text{pot}}\left(\mathbf{x}\right)\right)$ sets the left hand side of the quantum Boltzmann equation to zero and therefore cancels out. Inserting Eqs. (A3) into the self energies, parameterizing the deviations of $f_{\lambda}\left(\mathbf{k}\right)$ from the equilibrium distribution function as shown in Eq. (14), and linearizing in $\psi_{\mathbf{k}\lambda}\left(\mathbf{x},t\right)$ leads to the collision operator of Eq. (24). The matrix elements $\gamma^{(1,2)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}$ of Eq. (24) are given by:

$$\begin{aligned} \gamma_1\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= \left(N-1\right)\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 + \frac{1}{2}\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right) - T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\ &\quad -\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right)\right|^2 \\ \gamma_2\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= \left(N-1\right)\left|T_B\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{k}-\mathbf{q}\right)\right|^2 + \left(N-1\right)\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\ &\quad +\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) - T_B\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right)\right|^2, \end{aligned} \tag{A6}$$

with

$$\begin{aligned} T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= T_{++++}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{----}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\ &= T_{+--+}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{-++-}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\ &= \frac{V\left(q\right)}{4}\left(1+\frac{\left(K+Q\right)K^{*}}{\left|\mathbf{k}+\mathbf{q}\right|k}\right)\left(1+\frac{\left(K'-Q\right)K'^{*}}{\left|\mathbf{k}'-\mathbf{q}\right|k'}\right) \end{aligned}$$

and

$$\begin{aligned} T_B\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= T_{++--}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{--++}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\ &= \frac{V\left(q\right)}{4}\left(1-\frac{\left(K+Q\right)K^{*}}{\left|\mathbf{k}+\mathbf{q}\right|k}\right)\left(1-\frac{\left(K'-Q\right)K'^{*}}{\left|\mathbf{k}'-\mathbf{q}\right|k'}\right) \end{aligned} \tag{A7}$$

Upper-case letters like $K = k_x + ik_y$ etc. combine the two components of the momentum vector onto a complex variable.

Since the quantum Boltzmann equation only accounts for the diagonal in λ components of the distribution function, the currents also have to be decomposed into contributions that involve particle-hole pair creation ($\mathbf{j}_{\text{inter}}$) and those who do not ($\mathbf{j}_{\text{intra}}$). Here, the identity

$$U_{\mathbf{k}}\sigma U_{\mathbf{k}}^{-1} = \frac{\mathbf{k}}{k}\sigma_z - \frac{\mathbf{k}\times\mathbf{e}_z}{k}\sigma_y \tag{A8}$$

is useful. The charge current

$$\mathbf{j}_c = ev\int_{\mathbf{k}}\psi^{\dagger}\left(\mathbf{k}\right)\sigma\psi\left(\mathbf{k}\right) \tag{A9}$$

can be written as

$$\mathbf{j}_c = \mathbf{j}_{c,\text{intra}} + \mathbf{j}_{c,\text{inter}}, \tag{A10}$$

where the two contributions are given by

$$\begin{aligned} \mathbf{j}_{c,\text{intra}} &= ev\int_{\mathbf{k}}\sum_{\lambda=\pm}\frac{\lambda\mathbf{k}}{k}\gamma^{\dagger}_{\mathbf{k},\lambda}\gamma_{\mathbf{k},\lambda} \\ \mathbf{j}_{c,\text{inter}} &= iev\int_{\mathbf{k}}\frac{\mathbf{k}\times\mathbf{e}_z}{k}\left(\gamma^{\dagger}_{\mathbf{k},+}\gamma_{\mathbf{k},-} - \gamma^{\dagger}_{\mathbf{k},-}\gamma_{\mathbf{k},+}\right). \end{aligned} \tag{A11}$$

The energy current $\mathbf{j}_{\varepsilon}$ and the momentum current tensor τ_{xy} can be decomposed in a similar manner. This leads to the expressions (41) and (42) of the main text and the expression that is used for τ_{xy} in Sec. IV C 3. As discussed above, in the hydrodynamic regime, it is legitimate to focus on the intra-band contributions, which dominate the transport behavior of the system.

Appendix B: Collinear scattering and collinear zero modes

Here, the logarithmic divergence of the collision operator for collinear processes is demonstrated following Ref. [11]. We then show, that the m-dependent collinear zero modes are those given in Eq. (32).

The essential mathematics behind the divergence is contained in phase space density available for two particle collisions. The phase space is restricted by the delta function ensuring energy conservation: $\delta\left(k + k_1 - \left|\mathbf{k}+\mathbf{q}\right| - \left|\mathbf{k}_1-\mathbf{q}\right|\right)$. This can be seen from power counting in Eq. (24) using Eqs. (A6), (A7).

Choosing $\mathbf{k} = (k, 0)$ with $k > 0$, and writing $\mathbf{k}_1 = (k_1, k_\perp)$, $\mathbf{q} = (q, q_\perp)$, collinear scattering occurs when $k_1 > 0$, $k + q > 0$, $k_1 - q > 0$ and $q_\perp \approx 0$, $k_\perp \approx 0$. For small $q_\perp$, $k_\perp$ the argument of the delta function can be approximated as

$$k + k_1 - \left|\mathbf{k}+\mathbf{q}\right| - \left|\mathbf{k}_1-\mathbf{q}\right| \approx \frac{k_\perp^2}{2k_1} - \frac{q_\perp^2}{2\left(k+q\right)} - \frac{\left(k_\perp - q_\perp\right)^2}{2\left(k_1-q\right)}. \tag{B1}$$

The right hand side of this equation is a polynomial in $q_\perp$, and can be written in terms of linear factors as

$$\frac{k_\perp^2}{2k_1} - \frac{q_\perp^2}{2\left(k+q\right)} - \frac{\left(k_\perp - q_\perp\right)^2}{2\left(k_1-q\right)} = -\frac{k_1+k}{2\left(k+q\right)\left(k_1-q\right)}\left(q_\perp - \zeta_1 k_\perp\right)\left(q_\perp - \zeta_2 k_\perp\right).$$

It is then easy to see by performing the $q_\perp$ integration that

$$\int dk_\perp dq_\perp \delta\left(-\frac{k_1+k}{2\left(k+q\right)\left(k_1-q\right)}\left(q_\perp-\zeta_1 k_\perp\right)\left(q_\perp-\zeta_2 k_\perp\right)\right) \propto \int \frac{dk_\perp}{k_\perp}.$$

This behavior leads to a logarithmic divergence. The divergence is however cut off by the screening of the Coulomb potential [65]

$$V\left(|\mathbf{q}|\right) \to V\left(|\mathbf{q}|+q_{TF}\right),$$

where q_{TF} is the Thomas Fermi screening length. In the case of charge neutral graphene $q_{FT}=\alpha k_B T/v$. If the screening is included, the integral of (24) vanishes in the infrared. Thus, the contribution of collinear processes to the scattering rates is enhanced by the large factor

$$\log\left(1/\alpha\right).$$

It was demonstrated in sec. III B of the main text, that relaxation processes in the hydrodynamic regime are dominated by collinear zero modes. As demonstrated above, these modes describe scattering events in which all particle velocities show in the same direction. Examining the delta function responsible for energy conservation $\delta\left(k+k_1-|\mathbf{k}+\mathbf{q}|-|\mathbf{k}_1-\mathbf{q}|\right)$, we see that, if all momenta are parallel to each other, energy is only conserved, if the above conditions $k>0$, $k_1>0$, $k+q>0$, $k_1-q>0$ apply (except for unimportant isolated points in phase space). The exchange momentum q, however, can be positive or negative. To find those $\psi_{\mathbf{k}\lambda}$ that correspond to collinear zero modes, two terms in the collision operator Eq. (24) have to be considered:

$$\begin{aligned} A^{(1)}_{\mathbf{k},\mathbf{k}_1,\mathbf{q},\lambda} &= \psi_{\mathbf{k}+\mathbf{q}\lambda}+\psi_{\mathbf{k}_1-\mathbf{q}\lambda}-\psi_{\mathbf{k}_1\lambda}-\psi_{\mathbf{k}\lambda} \\ A^{(2)}_{\mathbf{k},\mathbf{k}_1,\mathbf{q},\lambda} &= \psi_{\mathbf{k}+\mathbf{q}\lambda}-\psi_{-\mathbf{k}_1+\mathbf{q}\bar{\lambda}}+\psi_{-\mathbf{k}_1\bar{\lambda}}-\psi_{\mathbf{k}\lambda}. \end{aligned} \tag{B2}$$

Using the parameterization

$$\psi_{\mathbf{k},\lambda}=a_{\lambda,m}\left(k\right)e^{im\theta_{\mathbf{k}}} \tag{B3}$$

yields

$$\begin{aligned} A^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= \left(a_{\lambda,m}\left(k+q\right)+a_{\lambda,m}\left(k_1-q\right)-a_{\lambda,m}\left(k_1\right)-a_{\lambda,m}\left(k\right)\right)e^{im\theta_{\mathbf{k}}} \\ A^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= \left(a_{\lambda,m}\left(k+q\right)-\left(-1\right)^m a_{\bar{\lambda},m}\left(k_1-q\right)+\left(-1\right)^m a_{\bar{\lambda},m}\left(k_1\right)-a_{\lambda,m}\left(k\right)\right)e^{im\theta_{\mathbf{k}}}. \end{aligned} \tag{B4}$$

For collinear zero modes

$$\begin{aligned} A^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= 0 \\ A^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= 0 \end{aligned}$$

has to hold. $A^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda}$ is set to zero by $a_{\lambda,m}\left(k\right)=\{1,\lambda,\beta v\hbar k,\lambda\beta v\hbar k\}$. $A^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda}$ is more restrictive. For even m its zero modes are given by $a_{\lambda,m}\left(k\right)=\{1,\lambda,\lambda\beta v\hbar k\}$, for odd m the zero modes are $a_{\lambda,m}\left(k\right)=\{1,\lambda,\beta v\hbar k\}$. Summing up, the collinear zero modes are given by

$$a_{\lambda,m}=\lambda^m\{1,\lambda,\lambda\beta v\hbar k\}e^{im\theta_{\mathbf{k}}}.$$

Appendix C: Matrix elements of the collision operator

The values of some matrix elements are shown in Table I. For $m\geq 2$ the values can be approximated by

$$\begin{aligned} \left\langle \chi^{(m,s=1)}_{\mathbf{k},\lambda}\left|\mathcal{C}\right|\chi^{(m,s=1)}_{\mathbf{k},\lambda}\right\rangle &= 2.574\cdot|m|-3.456 \\ \left\langle \chi^{(m,s=2)}_{\mathbf{k},\lambda}\left|\mathcal{C}\right|\chi^{(m,s=2)}_{\mathbf{k},\lambda}\right\rangle &= 1.825\cdot|m|-2.741 \\ \left\langle \chi^{(m,s=3)}_{\mathbf{k},\lambda}\left|\mathcal{C}\right|\chi^{(m,s=3)}_{\mathbf{k},\lambda}\right\rangle &= 5.184\cdot|m|-11.37 \\ \left\langle \chi^{(m,s=2)}_{\mathbf{k},\lambda}\left|\mathcal{C}\right|\chi^{(m,s=3)}_{\mathbf{k},\lambda}\right\rangle &= 2.042\cdot|m|-4.398. \end{aligned} \tag{C1}$$

All values are given in units of $\frac{1}{v^2\beta^3\hbar^3}$.

m	s	s'	$\left\langle \chi_{\mathbf{k},\lambda}^{(m,s)} \left\lvert \mathcal{C} \right\rvert \chi_{\mathbf{k},\lambda}^{(m,s')} \right\rangle$	m	s	s’	$\left\langle \chi_{\mathbf{k},\lambda}^{(m,s)} \left\lvert \mathcal{C} \right\rvert \chi_{\mathbf{k},\lambda}^{(m,s')} \right\rangle$	m	s	s’	$\left\langle \chi_{\mathbf{k},\lambda}^{(m,s)} \left\lvert \mathcal{C} \right\rvert \chi_{\mathbf{k},\lambda}^{(m,s')} \right\rangle$
0	1	1	0	2	1	1	2.617	4	1	1	6.988
0	1	2	0	2	1	2	0	4	1	2	0
0	1	3	0	2	1	3	0	4	1	3	0
0	2	2	0	2	2	2	1.745	4	2	2	4.722
0	2	3	0	2	2	3	1.243	4	2	3	4.122
0	3	3	0	2	3	3	3.341	4	3	3	10.456
1	1	1	0.804	3	1	1	4.728	5	1	1	9.345
1	1	2	0	3	1	2	0	5	1	2	0
1	1	3	0	3	1	3	0	5	1	3	0
1	2	2	0.463	3	2	2	3.167	5	2	2	6.351
1	2	3	0	3	2	3	2.573	5	2	3	5.800
1	3	3	0	3	3	3	6.647	5	3	3	14.610

Table I: Matrix elements of the collision operator (24) with respect to the collinear zero modes $\chi_{\mathbf{k},\lambda}^{(m,s)} = \lambda^m e^{im\theta}\left\{1, \lambda, \lambda\beta v\hbar k\right\}$. The index m labels the angular harmonic and s one of the modes in curved brackets.

Appendix D: Decomposition of the viscosity tensor into longitudinal and transverse parts

Consider a system with a preference direction introduced by the wave-vector $\mathbf{q}$. It is useful to define the orthogonal tensor basis

$$\begin{aligned} e_{\alpha\beta}^{(1)} &= \frac{q_\alpha q_\beta}{q^2} \\ e_{\alpha\beta}^{(2)} &= \delta_{\alpha\beta} - \frac{q_\alpha q_\beta}{q^2} \\ e_{\alpha\beta}^{(3)} &= \frac{1}{\sqrt{2}}\left(q_\alpha p_\beta + p_\alpha q_\beta\right)/\left(pq\right), \end{aligned} \tag{D1}$$

which is normalized according according to

$$\sum_{\alpha\beta} e_{\alpha\beta}^{(i)} e_{\alpha\beta}^{(j)} = \delta_{ij}.$$

Here it is

$$p_\alpha = q_\gamma \varepsilon_{\gamma\alpha}.$$

In this basis, the symmetric shear force tensor $X_{0,\alpha\beta}$ can be written

$$X_{0,\alpha\beta} = X^{(1)} e_{\alpha\beta}^{(1)} + X^{(2)} e_{\alpha\beta}^{(2)} + X^{(3)} e_{\alpha\beta}^{(3)}. \tag{D2}$$

The same holds for the momentum current (stress) tensor

$$\tau_{\alpha\beta} = \tau^{(1)} e_{\alpha\beta}^{(1)} + \tau^{(2)} e_{\alpha\beta}^{(2)} + \tau^{(3)} e_{\alpha\beta}^{(3)}. \tag{D3}$$

Since the system is fully isotropic, except for the preference direction set by $\mathbf{q}$, the response of the system to different components of $X_{0,\alpha\beta}$ can only be distinct as far as these components relate differently to the direction of $\mathbf{q}$. Eqs (D2) and (D3) are decompositions of the shear force and momentum current tensors into such components. The fourth rank viscosity tensor $\eta_{\alpha\beta\gamma\delta}$ is defined through the constitutive relation

$$\tau_{\alpha\beta} = \eta_{\alpha\beta\gamma\delta} X_{0,\gamma\delta}.$$

In general, such a tensor connecting the quantities $\tau_{\alpha\beta}$ and $X_{0,\alpha\beta}$ as given by Eqs. (D2), (D3) can be written as $\eta_{\alpha\beta\gamma\delta} = \sum_{ij} e_{\alpha\beta}^{(i)} e_{\gamma\delta}^{(j)} \eta^{(ij)}$. However it follows from an Onsager reciprocity relation that $\eta_{\alpha\beta\gamma\delta}$ has to be symmetric with respect to an interchange of the first and last pairs of indices:

$$\eta_{(\alpha\beta)(\gamma\delta)} = \eta_{(\gamma\delta)(\alpha\beta)}.$$

This condition further restricts the form of $\eta_{\alpha\beta\gamma\delta}$ to

$$\eta_{\alpha\beta\gamma\delta} = \sum_i e^{(i)}_{\alpha\beta} e^{(i)}_{\gamma\delta} \eta^{(i)}. \tag{D4}$$

Calculating the scalars $\eta^{(i)}$ using the quantum Boltzmann equation, one finds $\eta^{(1)} = \eta^{(2)} \neq \eta^{(3)}$. For reasons explained in the main text, we call $\eta^{(1)} = \eta^{(2)} = \eta_\perp$ the transverse, and $\eta^{(3)} = \eta_\|$ the longitudinal viscosity. In the sense that $\eta_{\alpha\beta\gamma\delta}$ is spanned by projection operators onto the tensorial subspaces which span the force and current tensors and are given in Eqs. (D1), the decomposition (D4) is completely analogous to the decomposition of a conductivity tensor into transverse and longitudinal parts (see Eq. (57)).

[1] E. I. Kiselev and J. Schmalian, *Lévy flights and hydrodynamic superdiffusion on the dirac cone of graphene*, Phys. Rev. Lett. **123**, 195302 (2019).

[2] A. B. Pippard, *The surface impedance of superconductors and normal metals at high frequencies. ii. the anomalous skin effect in normal metals*, Proc. R. Soc. Lond. A **191**, 385 (1947).

[3] A. B. Pippard, *Trapped flux in superconductors*, Philosophical Transactions of the Royal Society A: Math., Phys. and Eng. Sci. **248**, 97 (1955).

[4] J. R. Waldram, A. B. Pippard, and J. Clarke, *Theory of the current-voltage characteristics of sns junctions and other superconducting weak links*, Philosophical Transactions of the Royal Society A: Math., Phys. and Eng. Sci. **268**, 265 (1970).

[5] D. Svintsov, *Hydrodynamic-to-ballistic crossover in dirac materials*, Phys. Rev. B **97**, 121405 (2018).

[6] I. Torre, L. V. de Castro, B. V. Duppen, D. B. Ruiz, F. M. Peeters, F. H. L. Koppens, and M. Polini, *Acoustic plasmons at the crossover between the collisionless and hydrodynamic regimes in two-dimensional electron liquids*, Phys. Rev. B **99**, 144307 (2019).

[7] J. Brewer and P. Romatschke, *Nonhydrodynamic transport in trapped unitary fermi gases*, Phys. Rev. Lett. **115**, 190404 (2015).

[8] P. Romatschke, *Retarded correlators in kinetic theory: branch cuts, poles and hydrodynamic onset transitions*, Eur. Phys. J. C **76**, 352 (2016).

[9] P. Romatschke, *Relativistic fluid dynamics far from local equilibrium*, Phys. Rev. Lett. **120**, 12301 (2018).

[10] M. P. Heller, A. Kurkela, M. Spalinski, and V. Svensson, *Hydrodynamization in kinetic theory: Transient modes and the gradient expansion*, Phys. Rev. D **97**, 91503 (2018).

[11] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, *Quantum critical transport in clean graphene*, Phys. Rev. B **78**, 85416 (2008).

[12] A. B. Kashuba, *Conductivity of defectless graphene*, Phys. Rev. B **78**, 85415 (2008).

[13] M. Müller, J. Schmalian, and L. Fritz, *Graphene: A nearly perfect fluid*, Phys. Rev. Lett. **103**, 25301 (2009).

[14] M. S. Foster and I. L. Aleiner, *Slow imbalance relaxation and thermoelectric transport in graphene*, Phys. Rev. B **79**, 85415 (2009).

[15] K. Damle and S. Sachdev, *Nonzero-temperature transport near quantum critical points*, Phys. Rev. B **56**, 8714 (1997).

[16] S. A. Hartnoll, P. K. Kovtun, M. Müller, and S. Sachdev, *Theory of the nernst effect near quantum phase transitions in condensed matter and in dyonic black holes*, Phys. Rev. B **76**, 144502 (2007).

[17] D. E. Sheehy and J. Schmalian, *Quantum critical scaling in graphene*, Phys. Rev. Lett. .

[18] J. Crossno, J. K. Shi1, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, *Observation of the dirac fluid and the breakdown of the wiedemann-franz law in graphene*, Science **351**, 1058 (2016).

[19] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, *Quantum-critical conductivity of the dirac fluid in graphene*, Science **364**, 158 (2019).

[20] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A. K. Geim, and S. Ilani, *Visualizing poiseuille flow of hydrodynamic electrons*, Nature **576**, 75 (2019).

[21] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. B. Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, *Fluidity onset in graphene*, Nature Communications **9**, 4533 (2018).

[22] D. A. Bandurin, I. Torre, R. K. Kumar, M. B. Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, L. A. Ponomarenko, A. K. Geim, and M. Polini, *Negative local resistance caused by viscous electron backflow in graphene*, Science **351**, 1055 (2016).

[23] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. K. Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin, *Measuring hall viscosity of graphene's electron fluid*, Science **364**, 162 (2019).

[24] R. K. Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. B. Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, and A. K. Geim, *Superballistic flow of viscous electron fluid through graphene constrictions*, Nature Physics **13**, 1182 (2017).

[25] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, *Hydrodynamics in graphene: Linear-response transport*, Phys. Rev. B **91**, 035414 (2015).

[26] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, *Collision-dominated nonlinear hydrodynamics in graphene*, Phys. Rev. B **92**, 115426 (2015).

[27] A. K. G. I. Torre, A. Tomadin and M. Polini, *Nonlo-*

cal transport and the hydrodynamic shear viscosity in graphene, Phys. Rev. B **92**, 165433 (2015).

[28] B. N. Narozhny, I. V. Gornyi, and A. D. Mirlin, *Hydrodynamic approach to electronic transport in graphene*, Ann. d. Phys. **529**, 1700043 (2017).

[29] H.-Y. Xie and A. Levchenko, *Negative viscosity and eddy flow of the imbalanced electron-hole liquid in graphene*, Phys. Rev. B **99**, 045434 (2019).

[30] M. J. Klug, M. S. Scheurer, and J. Schmalian, *Hierarchy of information scrambling, thermalization, and hydrodynamic flow in graphene*, Phys. Rev. B **98**, 045102 (2018).

[31] B. N. Narozhny, *Optical conductivity in graphene: Hydrodynamic regime*, Phys. Rev. B **100**, 115434 (2019).

[32] J. M. Link, D. E. Sheehy, B. N. Narozhny, and J. Schmalian, *Elastic response of the electron fluid in intrinsic graphene: The collisionless regime*, Phys. Rev. B **98**, 195103 (2018).

[33] L. S. Levitov, A. V. Shtyk, and M. V. Feigelman, *Electron-electron interactions and plasmon dispersion in graphene*, Phys. Rev. B **88**, 235403 (2013).

[34] D. S. Borgnia, T. V. Phan, and L. S. Levitov, *Quasi-relativistic doppler effect and non-reciprocal plasmons in graphene*, arXiv:1512.09044 (2015).

[35] B. N. Narozhny, *Electronic hydrodynamics in graphene.*, Annals of Physics **411**, 167979 (2019).

[36] B. N. Narozhny and M. Schütt, *Magnetohydrodynamics in graphene: Shear and hall viscosities*, Phys. Rev. B **100**, 035125 (2019).

[37] S. Danz and B. N. Narozhny, *Vorticity of viscous electronic flow in graphene*, 2D Materials **7**, 035001 (2020).

[38] O. Kashuba, B. Trauzettel, and L. W. Molenkamp, *Relativistic gurzhi effect in channels of dirac materials*, Phys. Rev. B **97**, 205129 (2018).

[39] A. Lucas and K. C. Fong, *Hydrodynamics of electrons in graphene*, Journal of Physics: Condensed Matter **30**, 53001 (2018).

[40] Z. Sun, D. N. Basov, and M. M. Fogler, *Universal linear and nonlinear electrodynamics of a dirac fluid*, PNAS **115**, 3285 (2018).

[41] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, *Evidence for hydrodynamic electron flow in* $\mathrm{PdCoO_2}$, Science **351**, 1061 (2016).

[42] A. P. Mackenzie, *The properties of ultrapure delafossite metals*, Rep. Prog. Phys. **80**, 32501 (2017).

[43] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, *Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide*, Nat. Commun **9**, 4093 (2018).

[44] G. M. Gusev, A. D. Levin, E. V. Levinson, and A. K. Bakarov, *Viscous transport and hall viscosity in a two-dimensional electron system*, Phys. Rev. B **98**, 161303 (2018).

[45] A. V. Andreev, S. A. Kivelson, and B. Spivak, *Hydrodynamic description of transport in strongly correlated electron systems*, Phys. Rev. Lett. **106**, 256804 (2011).

[46] A. Principi and G. Vignale, *Violation of the wiedemann-franz law in hydrodynamic electron liquids*, Phys. Rev. Lett. **115**, 056603 (2015).

[47] P. Alekseev, A. Dmitriev, I. Gornyi, V. Kachorovskii, B. Narozhny, M. Schütt, and M. Titov, *Magnetoresistance in two-component systems*, Phys. Rev. Lett. **114**, 156601 (2015).

[48] P. Alekseev, *Negative magnetoresistance in viscous flow of two-dimensional electrons*, Phys. Rev. Lett. **117**, 166601 (2016).

[49] H. Guo, E. Ilseven, G. Falkovich, and L. S. Levitov, *Higher-than-ballistic conduction of viscous electron flows*, PNAS **114**, 3068 (2017).

[50] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, *Hydrodynamic electron flow and hall viscosity*, Phys. Rev. Lett. **118**, 226601 (2017).

[51] P. S. Alekseev, *Magnetic resonance in a high-frequency flow of a two-dimensional viscous electron fluid*, Phys. Rev. B **98**, 165440 (2018).

[52] R. Moessner, P. Surówka, and P.Witkowski, *Pulsating flow and boundary layers in viscous electronic hydrodynamics*, Phys. Rev. B **97**, 161112 (2018).

[53] R. Cohen and M. Goldstein, *Hall and dissipative viscosity effects on edge magnetoplasmons*, Phys. Rev. B **98**, 235103 (2018).

[54] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, *Dissipative and hall viscosity of a disordered 2d electron gas*, Phys. Rev. Lett. **123**, 026804 (2019).

[55] R. Moessner, N. Morales-Durán, P. Surówka, and P. Witkowski, *Boundary-condition and geometry engineering in electronic hydrodynamics*, Phys. Rev. B **100**, 155115 (2019).

[56] T. Zdyrski and J. McGreevy, *Effects of dissipation on solitons in the hydrodynamic regime of graphene*, Phys. Rev. B **99**, 235435 (2019).

[57] P. Alekseev and A. Alekseeva, *Transverse magnetosonic waves and viscoelastic resonance in a two-dimensional highly viscous electron fluid*, Phys. Rev. Lett. **123**, 236801 (2019).

[58] D. Svintsov, *Emission of plasmons by drifting dirac electrons: A hallmark of hydrodynamic transport*, Phys. Rev. B **100**, 195428 (2019).

[59] T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, *Ballistic and hydrodynamic magnetotransport in narrow channels*, Phys. Rev. B **100**, 245305 (2019).

[60] C. Q. Cook and A. Lucas, *Electron hydrodynamics with a polygonal fermi surface*, Phys. Rev. B **99**, 235148 (2019).

[61] F. M. D. Pellegrino, I. Torre, and M. Polini, *Nonlocal transport and the hall viscosity of two-dimensional hydrodynamic electron liquids*, Phys. Rev. B **96**, 195401 (2017).

[62] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, *Counterflows in viscous electron-hole fluid*, Phys. Rev. B **98**, 125111 (2018).

[63] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, *Out-of-bounds hydrodynamics in anisotropic dirac fluids*, Phys. Rev. Lett. **120**, 196801 (2018).

[64] I. Matthaiakakis, D. R. Fernández, C. Tutschku, E. M. Hankiewicz, J. Erdmenger, and R. Meyer, *Functional dependence of the hall viscosity-induced transverse voltage in two-dimensional fermi liquids*, Phys. Rev. B **101**, 045423 (2020).

[65] M. Müller, L. Fritz, and S. Sachdev, *Quantum-critical relativistic magnetotransport in graphene*, Phys. Rev. B **78**, 115406 (2008).

[66] A. D. Mirlin and P. Wölfle, *Composite fermions in the*

fractional quantum hall effect: Transport at finite wave vector, Phys. Rev. Lett. .

[67] K. A. Ingebrigtsen, *Surface waves in piezoelectrics*, J. Appl. Phys. **40**, 2681 (1969).

[68] S. H. Simon, *Coupling of surface acoustic waves to a two-dimensional electron gas*, Phys. Rev. B **54**, 13878 (1996).

[69] A. L. Efros and Y. M. Galperin, *Quantization of the acoustoelectric current in a two-dimensional electron system in a strong magnetic field*, Phys. Rev. Lett. **41**, 1959 (1990).

[70] A. Wixforth, J. Scriba, M. Wassermeier, J. P. Kotthaus, G. Weimann, and W. Schlapp, *Surface acoustic waves on gaas/al_xga_{1-x}*, Phys. Rev. B **40**, 7874 (1989).

[71] M. Rotter, A. Wixforth, W. Ruile, D. Bernklau, and H. Riechert, *Giant acoustoelectric effect in gaas/$linbo_3$ hybrids*, Appl. Phys. Lett. **73**, 2128 (1998).

[72] A. O. Govorov, A. V. Kalameitsev, M. Rotter, A. Wixforth, J. P. Kotthaus, K.-H. Hoffmann, and N. Botkin, *Nonlinear acoustoelectric transport in a two-dimensional electron system*, Phys. Rev. B **62**, 2659 (2000).

[73] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, and J. A. Sulpizio, *Simultaneous voltage and current density imaging of flowing electrons in two dimensions*, Nature Nanotechnology **14**, 480 (2019).

[74] T. V. Phan, J. C. W. Song, and L. S. Levitov, *Ballistic heat transfer and energy waves in an electron system*, arXiv:1306.4972 (2016).

[75] R. N. Gurzhi, *Minimum of resistance in impurity-free conductors*, Zh. Eksp. Teor. Fiz. **44**, 771 (1963).

[76] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *The electronic properties of graphene*, Rev. Mod. Phys. **81**, 109 (2009).

[77] L. D. Landau, *The theory of a fermi liquid*, Sov. Phys. JETP **3**, 920 (1957).

[78] L. D. Landau, *Oscillations in a fermi liquid*, Sov. Phys. JETP **5**, 101 (1957).

[79] A. A. Abrikosov and I. M. Khalatnikov, *The theory of a fermi liquid (the properties of liquid 3he at low temperatures)*, Reports on Progress in Physics **22**, 329 (1959).

[80] C. Hodges, H. Smith, and J. W. Wilkins, *Effect of fermi surface geometry on electron-electron scattering.*, Physical Review B **4**, 302 (1971).

[81] D. C. Elias, R. V. Gorbachev, A. S. Mayorov, S. V. Morozov, A. A. Zhukov, P. Blake, L. A. Ponomarenko, I. V. Grigorieva, K. S. Novoselov, F. Guinea, and A. K. Geim, *Dirac cones reshaped by interaction effects in suspended graphene*, Nature Physics **7**, 701 (2011).

[82] L. P. Kadanoff and G. Baym, *Quntum Statistical Mechanics*, Addison-Wesley (1994). ISBN: 978-0201410464.

[83] G. D. Mahan, *Many-Particle Physics*, Springer (2000). ISBN: 978-0306463389.

[84] J. M. Link, P. P. Orth, D. E. Sheehy, and J. Schmalian, *Universal collisionless transport of graphene*, Phys. Rev. B **93**, 235447 (2016).

[85] L. Wang, I. Meric, P. Y. Huang, Q. Gao, Y. Gao, H. Tran, T. Taniguchi, K. Watanabe, L. M. Campos, D. A. Muller, J. Guo, P. Kim, J. Hone, K. L. Shepard, and C. R. Dean, *One-dimensional electrical contact to a two-dimensional material*, Science **342**, 614 (2013).

[86] E. H. Hwang, S. Adam, and S. D. Sarma, *Carrier Transport in Two-Dimensional Graphene Layers*, Phys. Rev. Lett. **98**, 186806 (2007).

[87] J. Martin, N. Akerman, G. Ulbricht, T. Lohmann, J. H. Smet, K. von Klitzing, and A. Yacoby, *Observation of electron-hole puddles in graphene using a scanning single-electron transistor*, Nature **4**, 144 (2008).

[88] V. V. Cheianov, V. I. Fal'ko, B. L. Altshuler, and I. L. Aleiner, *Random Resistor Network Model of Minimal Conductivity in Graphene*, Phys. Rev. Lett. **99**, 176801 (2007).

[89] A. Lucas, J. Crossno, K. C. Fong, P. Kim, and S. Sachdev, *Transport in inhomogeneous quantum critical fluids and in the Dirac fluid in graphene*, Phys. Rev. B **93**, 075426 (2016).

[90] M. Gibertini, A. Tomadin, F. Guinea, M. I. Katsnelson, and M. Polini, *Electron-hole puddles in the absence of charged impurities*, Phys. Rev. B **85**, 201405 (2012).

[91] N. Sule, S. Hagness, and I. Knezevic, *Clustered impurities and carrier transport in supported graphene*, Phys. Rev. B **89**, 165402 (2014).

[92] A. Levchenko and J. Schmalian, *Transport properties of strongly coupled electron–phonon liquids*, Annals of Physics (168218) (2020).

[93] D. K. Efetov and P. Kim, *Controlling electron-phonon interactions in graphene at ultrahigh carrier densities*, Phys. Rev. Lett. **105**, 256805 (2010).

[94] J.-H. Chen, C. Jang, S. Xiao, M. Ishigami, and M. S. Fuhrer, *Intrinsic and extrinsic performance limits of graphene devices on sio 2*, Nature nanotechnology **3**, 206 (2008).

[95] E. Hwang, R. Sensarma, and S. D. Sarma, *Plasmon-phonon coupling in graphene*, Physical Review B **82**, 195406 (2010).

[96] F. Karimi and I. Knezevic, *Plasmons in graphene nanoribbons*, Physical Review B **96**, 125417 (2017).

[97] E. I. Kiselev and J. Schmalian, *Boundary conditions of viscous electron flow*, Phys. Rev. B **99**, 35430 (2019).

[98] Y.-W. Son, M. L. Cohen, and S. G. Louie, *Energy gaps in graphene nanoribbons*, Phys. Rev. Lett. **97**, 216803 (2006).

[99] B. Wild, L. Cao, Y. Sun, B. P. Khanal, E. R. Zubarev, S. K. Gray, N. F. Scherer, and M. Pelton, *Propagation lengths and group velocities of plasmons in chemically synthesized gold and silver nanowires*, ACS nano **6**, 472 (2012).

[100] A. A. Vlasov, *On vibration properties of electron gas*, J. Exp. Theor. Phys. **8**, 291 (1938).

[101] D. Pines and P. Nozières, *Theory Of Quantum Liquids: Normal Fermi Liquids*, CRC Press (1989). ISBN 978-0201407747.

[102] V. Narayanamurti and R. Dynes, *Observation of second sound in bismuth*, Phys. Rev. Lett. **28**, 1461 (1972).

[103] L. Bandhu, L. M. Lawton, and G. R. Nash, *Macroscopic acoustoelectric charge transport in graphene*, Appl. Phys. Lett. **103**, 133101 (2013).

[104] V. Miseikis, J. E. Cunningham, K. Saeed, R. O'Rorke, and A. G. Davies, *Acoustically induced current flow in graphene*, Appl. Phys. Lett. **100**, 133105 (2012).

[105] D. A. Huse and S. N. Majumdar, *Nonlocal resistivity in the vortex liquid regime of type-ii superconductors*, Phys. Rev. Lett. **71**, 2473 (1993).

[106] T. Kita, *Introduction to non-equilibrium statistical mechanics with quantum field theory*, Prog. Theor. Phys. **123**, 581 (2010).

Lévy flights and hydrodynamic superdiffusion on the Dirac cone of Graphene

Egor I. Kiselev[1] and Jörg Schmalian[1, 2]
[1] *Institut für Theorie der Kondensierten Materie,*
Karlsruher Institut für Technologie, 76131 Karlsruhe, Germany
[2] *Institut für Festkörperphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe, Germany*

We show that hydrodynamic collision processes of graphene at the neutrality point can be described in terms of a Fokker-Planck equation with fractional derivative, corresponding to a Lévy flight in momentum space. Thus, electron-electron collisions give rise to frequent small-angle scattering processes that are interrupted by rare large-angle events. The latter give rise to superdiffusive dynamics of collective excitations. We argue that such superdiffusive dynamics is of more general importance to the out-of-equilibrium dynamics of quantum-critical systems.

The kinetics of large gravitational systems such as globular clusters in galaxies or of a classical charged plasma are governed by continuous collisions with small-angle scatterings. The origin for this behavior is the long-range character of the Newton or Coulomb force, respectively. Such small-angle collisions behave in velocity space like drag and diffusion events, where a Fokker-Planck equation offers an efficient description[1–3]. Collisions can thus be seen as a Gaussian random walk in phase space. The velocity of a plasma or gravitational dust particle undergoes ordinary Brownian motion.

Quantum many-body systems that are near a quantum-critical point are governed by soft modes that will also induce effective long-range interactions[4]. This begs the question whether such quantum-critical systems also allow for an effective Fokker-Planck description of the non-equilibrium kinetics; in the collision-dominated hydrodynamic regime and in the crossover regime from hydrodynamic to ballistic dynamics. Candidate systems are itinerant electrons near magnetic or nematic quantum phase transitions[5–14], the superconductor-insulator phase transition[15], or graphene near the Dirac point[16]. Anomalous diffusion was even shown to be present in two-dimensional Fermi liquids[17–22].

In this paper we analyze the quantum kinetics of graphene near the Dirac point with electron-electron Coulomb interaction. We show that the kinetic theory at charge neutrality[23–27] can be expressed in terms of a Fokker-Planck equation, yet with fractional derivative with respect to the momentum direction. The underlying random processes are Lévy flights[28–30], non-Gaussian random walks whose step widths are distributed according to a powerlaw. The slowly decaying tail of the step-width distribution makes it impossible to define a diffusion constant or to use a conventional Fokker-Planck equation. However, a diffusion equation of the form

$$\frac{\partial \rho}{\partial t} + D_\mu \left|\triangle\right|^{\frac{\mu}{2}} \rho = 0, \qquad (1)$$

with appropriately generalized fractional derivative[31, 32] can be used to describe such random walks. Lévy flights have been discussed to model the migration pattern of animals as they search for resources[33, 34], the high-frequency index dynamics of the stock market[35], or to describe durations between consecutive earthquakes[36]. In our system they correspond to random walks in momentum space with powerlaw weight for large momentum-transfer processes. We demonstrate that the collision operator due to electron-electron interactions in graphene takes the form of a fractional derivative. Then the Boltzmann equation becomes a fractional Fokker-Planck equation, similar to Eq.1 with exponent $\mu = 1$:

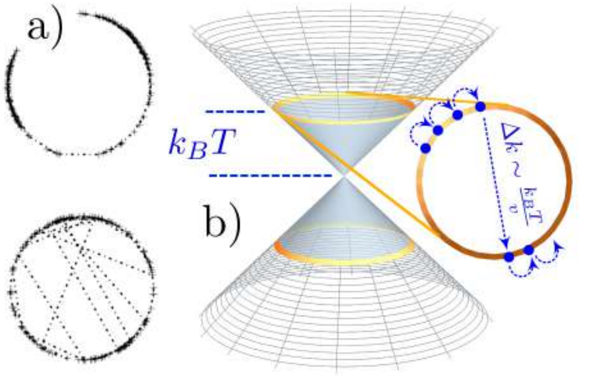


Figure 1: a) A wrapped Gaussian flight (upper circle) and a wrapped Cauchy flight (lower circle) with rare large momentum-transfer processes. b) Illustration of the Lévy flight in momentum space for graphene at the Dirac point. Electrons and holes that are thermally excited collide into each other. Most of the time the momentum transfer due to the electron-electron Coulomb interaction leads to small-angle scattering. However, those processes are interrupted by rare processes with large momentum transfer. The latter change the dynamics of the system qualitatively, leading to an accelerated or superdiffusive dynamics.

$$\left(\partial_t + \mathbf{v}_{\mathbf{k}\lambda} \cdot \nabla_{\mathbf{x}} - \tau_L^{-1} \left(\frac{\partial^2}{\partial \theta^2}\right)^{1/2}\right) f_{\mathbf{k}\lambda} = S_{\mathbf{k}\lambda}, \qquad (2)$$

where θ determines the electron momentum direction: $\mathbf{k} = k(\cos\theta, \sin\theta)$. The precise definition of the fractional derivative is given below. This result implies that the out-of-equilibrium dynamics of graphene in the hydrodynamic regime is governed by a wrapped Cauchy flight[37, 38], a specific Lévy flight on the Dirac cone. In Fig. a we show a simulation of ordinary Brownian motion on a ring and of the wrapped Cauchy flight. Details of this simulation are summarized in[39]. The occurrence of rare large-angle jumps is clearly visible. The corresponding phase-space dynamics is sketched in Fig. b. While the direction of $\mathbf{k}$ undergoes anomalous diffusion, its magnitude $k \equiv |\mathbf{k}|$ is of the order of $k_B T/v_0$ with the graphene group velocity $v_0 \approx 10^8$cm/s. The characteristic time of the process is τ_L with

$$\hbar \tau_L^{-1} \approx 11.66 \alpha^2 k_B T, \tag{3}$$

where the fine-structure constant of graphene is $\alpha = e^2/(\hbar \epsilon v_0)$. τ_L agrees up to a numerical coefficient with the collision time for the hydrodynamic transport behavior of graphene at the Dirac point[23–25]. Below we discuss how τ_L is determined. Such a time scale was recently observed experimentally in THz spectroscopy of graphene at charge neutrality[40].

Lévy flights in graphene have been discussed in Ref.[41], where an egineered distribution of adatoms was shown to result in a superdiffusive behavior of charge carriers, and in Ref.[42] in the context of highly photoexcited carriers that relax according to a cascade of processes - a behavior with interesting implications for pump-probe experiments. This can be seen as a superdiffusion in energy space far from equilibrium. It affects the magnitude of the momentum. Here we focus on the low-energy hydrodynamic regime and find a very different behavior for the directional diffusion in momentum space. Nevertheless, these results strongly suggest that superdiffusive phase-space dynamics is a more common phenomenon in quantum-critical systems.

We start from the Boltzmann equation

$$\left(\partial_t + \mathbf{v}_{\mathbf{k}\lambda} \cdot \nabla_{\mathbf{x}} + \mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{k}} + \mathcal{C}\right) f_{\mathbf{k}\lambda}(\mathbf{x}, t) = 0 \tag{4}$$

for the electron distribution function $f_{\mathbf{k}\lambda}(\mathbf{x}, t)$ where $\mathbf{k}$ refers to the momentum and $\lambda = \pm 1$ labels the upper and lower cone of the Dirac spectrum $\varepsilon_{\mathbf{k}\lambda} = \lambda v_0 |\mathbf{k}|$. $\mathbf{v}_{\mathbf{k}\lambda} = \partial \varepsilon_{\mathbf{k}\lambda}/\partial \mathbf{k}$ is the velocity vector and $\mathbf{F}(\mathbf{x}, t)$ some external force, e.g. due to an external electric field. $\mathcal{C}$ is the Boltzmann collision operator due to electron-electron interactions and was derived to order α^2 in Ref.[23] from a Keldysh-Schwinger approach; see also in[39]. It takes the usual form of a two-body interaction:

$$\begin{aligned} \mathcal{C} f_1 &= -\sum_{2,3,4} W_{12,34} \left[f_1 f_2 (1-f_3)(1-f_4) \right| \\ &- \left| (1-f_1)(1-f_2) f_3 f_4 \right]. \end{aligned} \tag{5}$$

The transition probability $W_{12,34}$ is due to the electron-electron Coulomb interaction e^2/ϵ of Dirac fermions that are confined to a two-dimensional system. ϵ is the dielectric constant determined by the substrate. For free standing graphene $\epsilon = 1$ and the fine-structure constant $\alpha \approx 2.2$ is of order unity. A renormalization group analysis shows that α flows towards weak coupling, justifying our perturbative approach[16].

As usual, the kinetic distribution function $f_{\lambda,\boldsymbol{k}}$ is expanded around the local equilibrium distribution $f_{k\lambda}^0 = \left(e^{\beta(\epsilon_{\lambda,\boldsymbol{k}} - \mu)} + 1\right)^{-1}$ and parametrized as ($f_k^{(0)} = f_{k+}^{(0)}$):

$$f_{\mathbf{k}\lambda}(\mathbf{x}, t) = f_{k\lambda}^{(0)} + f_k^{(0)} \left(1 - f_k^{(0)}\right) \psi_{\mathbf{k}\lambda}(\mathbf{x}, t). \tag{6}$$

We linearize the Boltzmann equation with respect to $\psi_{\mathbf{k}\lambda}(\mathbf{x}, t)$. With the Liouville operator

$$\mathcal{L} = \left(\partial_t + \mathbf{v}_{\mathbf{k}\lambda} \cdot \nabla_{\mathbf{x}}\right) f_k^{(0)} \left(1 - f_k^{(0)}\right) \tag{7}$$

we obtain a compact formulation of the Boltzmann equation: $(\mathcal{L} + \mathcal{C}) \psi = S$. $S_{\mathbf{k}\lambda}(\mathbf{x}, t)$ contains external perturbations, such as those due to a space and time dependent electric field or flow-velocity gradient. The operators $\mathcal{L}$ and $\mathcal{C}$ act on the momentum and band indices $\mathbf{k}$ and λ, respectively. Taking into account the kinematic constraints of the linear Dirac spectrum, the collision operator becomes:

$$\begin{aligned} (\mathcal{C}\psi)_{\mathbf{k}\lambda} &= \frac{2\pi}{\hbar} \int_{k'q} \delta\left(k + k' - |\mathbf{k} + \mathbf{q}| - |\mathbf{k}' - \mathbf{q}|\right) \\ &\times \left(1 - f_k^{(0)}\right) \left(1 - f_{k'}^{(0)}\right) f_{|\mathbf{k}+\mathbf{q}|}^{(0)} f_{|\mathbf{k}'-\mathbf{q}|}^{(0)} \\ &\times \Big\{ \gamma_{\mathbf{k},\mathbf{k}',\mathbf{q}}^{(1)} \left(\psi_{\mathbf{k}+\mathbf{q}\lambda} + \psi_{\mathbf{k}'-\mathbf{q}\lambda} - \psi_{\mathbf{k}'\lambda} - \psi_{\mathbf{k}\lambda}\right) \\ &+ \gamma_{\mathbf{k},\mathbf{k}',\mathbf{q}}^{(2)} \left(\psi_{\mathbf{k}+\mathbf{q}\lambda} - \psi_{-\mathbf{k}'+\mathbf{q}\bar{\lambda}} + \psi_{-\mathbf{k}'\bar{\lambda}} - \psi_{\mathbf{k}\lambda}\right) \Big\}, \end{aligned} \tag{8}$$

where the matrix elements $\gamma_{\mathbf{k},\mathbf{k}',\mathbf{q}}^{(1,2)}$ are given in Ref.[39] and $\int_k \cdots = \int \frac{d^2k}{(2\pi)^2} \cdots$. One easily finds the zero modes that correspond to the conservation laws[23]. Eq.4 was obtained by projecting the distribution function onto the helical eigenstates of the problem. The same projection was performed in the derivation of the collision operator[23, 39].

The usual analysis of the Boltzmann equation proceeds as follows: One performs a Fourier transformation from $(\mathbf{x}, t)$ to $(\mathbf{q}, \omega)$ and introduces a complete set of states $\chi_{\mathbf{k}\lambda}^{(s)}$ to evaluate the matrix elements $\langle s | \mathcal{L} + \mathcal{C} | s' \rangle$ with scalar product $\langle s | s' \rangle = \sum_{\mathbf{k}\lambda} \chi_{\mathbf{k}\lambda}^{(s)*} \chi_{\mathbf{k}\lambda}^{(s')}$. The Liouville operator becomes $\mathcal{L} = \left(-i\omega + i\mathbf{v}_{\mathbf{k}\lambda} \cdot \mathbf{q}\right) f_k^{(0)} \left(1 - f_k^{(0)}\right)$. The distribution function then follows as $\psi = (\mathcal{L} + \mathcal{C})^{-1} S$. For finite ω or $\mathbf{q}$ the operator $\mathcal{L} + \mathcal{C}$ is nonsingular. This program is somewhat simplified for graphene at charge neutrality. As shown in Refs.[23–25, 27], scattering processes where all momenta are collinear are enhanced by a factor $\log(1/\alpha)$. This can be used to identify the dominant modes, derived in the supplementary material:

$$\chi_{\mathbf{k}\lambda}^{(m,s)} = \lambda^m e^{im\theta} \left\{1, \lambda, \lambda v_0 k/(k_B T)\right\}, \tag{9}$$

where $m \in \mathbb{Z}$ is the angular momentum quantum number while $s = 1 \cdots 3$ labels the collinear modes for given m. We solve the kinetic equation by projecting it onto the dominant collinear modes $\chi_{\mathbf{k}\lambda}^{(m,s)}$, but checked that our key conclusions are unchanged if we chose a larger set of basis functions. Also, if we restrict our considerations to the transport of charge due to external electric fields, it suffices to consider the modes $\chi_{\mathbf{k}\lambda}^{(m,1)} = \lambda^m e^{im\theta}$ of Eq. (9). For simplicity we confine ourselves to electric-field source terms and only discuss this mode. The generalization to other modes is straightforward.

The low-energy Dirac Hamiltonian is rotationally invariant such that the collision operator becomes diagonal in the angular momentum representation

$$\langle m \left|\mathcal{C}\right| m' \rangle = \frac{\ln 2}{\pi} \delta_{m,m'} \tau_m^{-1}. \tag{10}$$

The diagonal elements are, besides a convenient prefactor, the scattering rates of the corresponding angular momentum channel. $\tau_0^{-1} = 0$ due to charge conservation, while the collision rate

$$\hbar \tau_1^{-1} = 3.646 \alpha^2 k_B T \tag{11}$$

for $m = 1$ was determined in Ref.[23] to yield the optical conductivity $\sigma(\omega) = \frac{e^2}{h} 4 \ln 2 k_B T \left(-i\hbar\omega + \hbar\tau_1^{-1}\right)^{-1}$. τ_1^{-1} was recently observed in Ref.[40] using a waveguide setup; a demonstration of quantum-critical hydrodynamic transport. The dramatic violation of the Wiedemann-Franz law at charge neutrality is another important indication for electronic hydrodynamics at charge neutrality[44].

We evaluated the matrix elements $\langle m \left|\mathcal{C}\right| m \rangle$ and obtain

$$\tau_m^{-1} = \tau_1^{-1} \left(\kappa \left|m\right| - \kappa'\right), \tag{12}$$

where the two numerical constants are given as $\kappa \approx 3.199$ and $\kappa' \approx 4.296$, see also Fig. . This behavior is asymptotically exact at large m but valid with good accuracy already for $m > 2$. The most important aspect of this result is that the dependence of the scattering rate on the angular momentum m is non-analytic. To simplify the analysis we assume in the following that $\tau_m^{-1} = \tau_L^{-1} \left|m\right|$, where $\tau_L^{-1} = \kappa \tau_1^{-1}$ is the characteristic time of the of the Lévy flight process, given in Eq. (3).

The implication of the $\left|m\right|$-dependence of τ_m^{-1} becomes evident if we consider the scattering between two distinct momentum directions. Fourier transformation of τ_m^{-1} yields:

$$\langle \theta \left|\mathcal{C}\right| \theta' \rangle = -\frac{\ln 2 \tau_L^{-1}}{(2\pi)^2 \sin^2\left(\frac{\theta-\theta'}{2}\right)}. \tag{13}$$

Thus, we obtain a slowly-decaying powerlaw $\sim (\theta - \theta')^{-2}$ for scattering processes away from forward scattering. Using this result for $\langle \theta \left|\mathcal{C}\right| \theta' \rangle$ we can rewrite the Boltzmann equation in the form Eq. (2) with characteristic time τ_L of Eq. (3) for the Lévy flight. To arrive at Eq. (2) we used that the convolution of the distribution function with $\langle \theta \left|\mathcal{C}\right| \theta' \rangle$ can be expressed as a fractional derivative

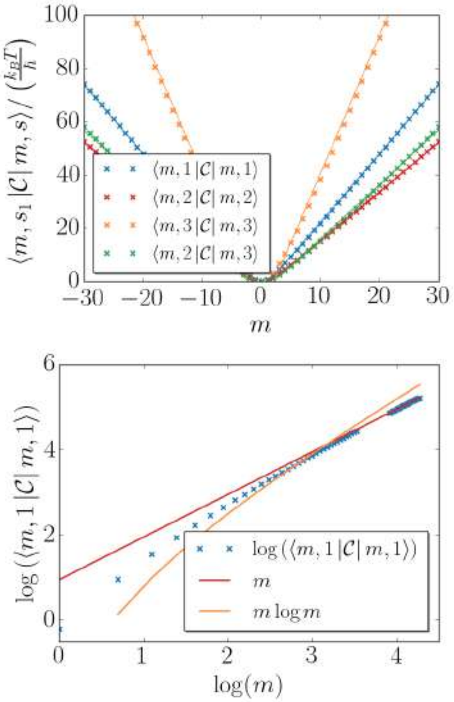


Figure 2: Upper panel: Angular momentum dependence of the matrix elements of the collision operator $\langle m, s \left|\mathcal{C}\right| m, s' \rangle$ where $s = 1 \cdots 3$ refers to the collinear eigenmodes of Eq.9. In the text we discuss, for simplicity, only $\langle m \left|\mathcal{C}\right| m \rangle \equiv \langle m, 1 \left|\mathcal{C}\right| m, 1 \rangle$. Lower panel: log-log plot of the matrix element to demonstrate that we can distinguish the $\left|m\right|$-dependence from, e.g. $\left|m\right| \log \left|m\right|$.

$$\left(\frac{\partial^2 g(\theta)}{\partial \theta^2}\right)^{1/2} = \frac{\partial}{\partial \theta} \int_0^{2\pi} \frac{g(\theta')}{\tan\left(\frac{\theta-\theta'}{2}\right)} d\theta', \tag{14}$$

a special case of the Riesz-Feller derivative $\triangle^{\mu/2}$[31, 32].

There are some profound implications that this fractional Fokker-Planck formulation immediately reveals. For example, we consider a scenario where we inject a highly directed excitation[22]. To this end we consider a source term in the Boltzmann equation that causes this excitation:

$$S_{\mathbf{k}\lambda}(t) = \delta(t) f_k^{(0)} \left(1 - f_k^{(0)}\right) \sum_m \delta h_{\lambda m} e^{im\theta}. \tag{15}$$

We assumed that we will only inject excitations in a window $\pm k_B T$ near the Dirac point, hence the factor $f_k^{(0)} \left(1 - f_k^{(0)}\right)$. In addition we decomposed the source

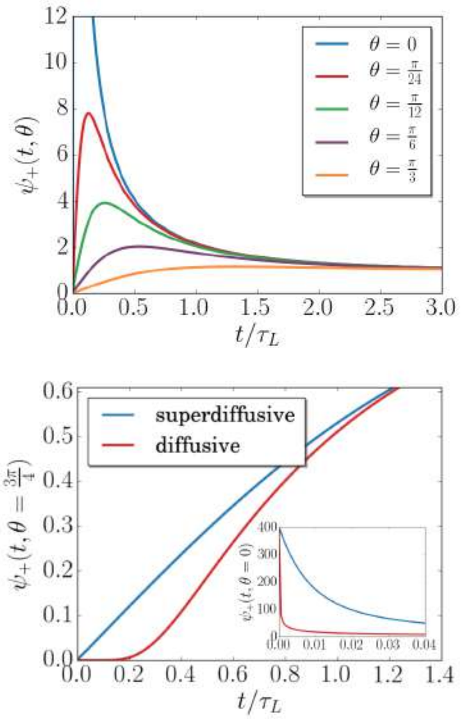


Figure 3: Upper panel: Post-injection distribution function that follows from the fractional Fokker-Planck equation, Eq.2, with external perturbation of Eq.15. Notice the superdiffusive dynamics at short times. Lower panel: Comparison of superdiffusive and diffusive dynamics at short times. At angles away from the peak at $\theta = 0$ superdiffusion leads to a faster growth of the distribution function. Inlet: the initial peak at $\theta = 0$ decays as $1/t$ for superdiffusion and $1/\sqrt{t}$ for ordinary diffusion. This behavior dominates the heating of the system (see main text).

term into its angular momentum modes. The linearized Boltzmann equation is applicable if $|\delta h_{\lambda m}| \ll 1$. To describe an excitation that is peaked along an axis given by a certain momentum direction, we use $\delta h_{\lambda,m} = \delta h \lambda^m$, which has a λ-dependence of the $s = 1$ mode of Eq. (9). The solution of the fractional Fokker-Planck equation for a homogeneous case $\mathbf{q} = 0$ is then given as

$$\psi_\lambda(\theta, t) = \delta h \Theta(t) \frac{\sinh(t/\tau_L)}{\cosh(t/\tau_L) - \lambda \cos(\theta)}. \tag{16}$$

This function is known as wrapped Cauchy distribution with circular variance $1 - e^{-t/\tau_L}$[37, 38]. $\Theta(t)$ is the step function. $\psi_+(\theta, t)$ is shown in the upper panel of Fig. .

For $t = 0$, $\psi_\lambda(\theta, t)$ corresponds to two delta functions due to particle and hole flows in opposite directions. Let us concentrate on the particle channel $\lambda = +1$. For short times $t \ll \tau_L$, the peak in the initial current direction decays as

$$\psi_+(t, \theta = 0) \approx \delta h \frac{\tau_L}{\pi t}, \tag{17}$$

while the distribution function grows linearly for all nonzero angles:

$$\psi_+(t, \theta \neq 0) \approx \frac{\delta h}{4\pi \sin^2(\theta/2)} \frac{t}{\tau_L}. \tag{18}$$

The same behavior occurs for $\lambda = -1$ if we shift $\theta \to \theta + \pi$. This behavior in contrast to the one that follows from usual Fokker-Planck diffusion. The latter we obtain for example from collision rates $\tau_m^{-1} \sim m^2$. Then the usual spreading of a Gaussian wave package occurs with $\psi_+(t, \theta = 0) \propto t^{-1/2}$ and $\psi_+(t, \theta \neq 0) \propto t^2$ (lower panel of Fig.). While the forward direction of a Levy flight decays more slowly than in usual diffusion, the growth at larger angles is much faster, hence the name superdiffusion.

A tangible implication of this superdiffusive charge motion is the heating of the system after the injection. To this end we determine the time dependence of the entropy density

$$\frac{\partial s(t)}{\partial t} = 4k_B \sum_\lambda \int_k \log\left(\frac{1 - f_{\mathbf{k}\lambda}}{f_{\mathbf{k}\lambda}}\right) \frac{\partial f_{\mathbf{k}\lambda}}{\partial t}. \tag{19}$$

The heat density caused by the injection is given by $\delta q(t) = T(s_{\text{eq}} - s(t))$. Inserting the distribution function of Eq. (16) we obtain

$$\frac{\partial s(t)}{\partial t} = \frac{4 \log 2}{9\zeta(3)} \frac{s_{\text{eq}}}{\tau_L} \frac{(\delta h)^2}{\sinh^2(t/\tau_L)}, \tag{20}$$

where s_{eq} is the equilibrium entropy density. In order to stay within the regime of linear response, we are confined to $t > \delta h \tau_L$. For $t \to \infty$ one finds $s \to s_{\text{eq}}$, and we obtain $s(t) = s_{\text{eq}}\left(1 - (\delta h)^2 \frac{4\log(2)}{9\zeta(3)}\left(\coth\left(\frac{t}{\tau_L}\right) - 1\right)\right)$. Thus, initial heating occurs according to

$$\delta q(t) \propto T s_{\text{eq}} \delta h^2 \frac{\tau_L}{t}. \tag{21}$$

This result is a direct consequence of the superdiffusive behavior, in particular of the slow decay along the forward direction. In case of ordinary diffusion follows $\delta q(t) \propto t^{-1/2}$ which is much faster (see Fig.). The m-dependence of τ_m^{-1} that is responsible for the Lévy flight behavior can also be seen in non-local transport coefficients since the conductivity at finite momentum $\mathbf{q}$ couples the different harmonics of the distribution function. As an example we show in the supplementary material the transverse optical conductivity at finite $\mathbf{q}$. Nevertheless, experiments with directed electron beams [46], which in the past have been used to investigate electron-electron scattering effects [45], seem to offer a more direct way of testing the short time behavior of Eq. (21).

The occurrence of Lévy flights to describe scattering processes in momentum space is a more general phenomenon and not restricted to graphene at the neutrality point. In two-dimensional Fermi liquids with characteristic rate $\hbar\tau_{FL}^{-1} \sim k_B T^2/T_F$, it holds for $|m| < M \sim \sqrt{T_F/T}$ that $\tau_m{}^{-1} \sim \tau_{FL}^{-1}\frac{m^p}{M^p}\log|m|$ with $p = 2\left(1+(-1)^m\right)$, while $\tau_m^{-1} \sim \tau_{FL}^{-1}$ for $|m| > M$[18, 22]. T_F is the Fermi temperature. This yields superdiffusive behavior in a wide time window. Another system that also shows $\tau_m^{-1} \propto |m|$ for arbitrarily large m consists of electrons in a random magnetic field, important for the description of composite fermions in the fractional quantum Hall regime[47]. Our analysis implies that this system should also undergo a wrapped Cauchy flight in momentum space. Large classes of quantum-critical systems, discussed e.g. in Refs.[5–15] are governed by long-ranged soft-mode interactions. An analysis of collision processes along the lines discussed here may reveal a non-analytic dependence of the scattering rates on angular momentum quantum number according to $\tau_m^{-1} \propto |m|^{\mu/2}$. This would give rise to a more general class of wrapped Lévy flights, a consequence of the power-law behavior $\langle\theta\,|\mathcal{C}|\,\theta'\rangle \propto |\theta-\theta'|^{-1-\frac{\mu}{2}}$ near forward scattering. This could occur on the Fermi surface for itinerant quantum critical systems or near a soft momentum in critical bosonic systems. If a fractional Fokker-Planck formulation, along the lines of our Eq. (2), can be derived, it will be significantly easier to draw conclusions about the out-of-equilibrium dynamics of the system such as a focussed injection of collective excitations. Finally we mention that the formulation of the Boltzmann equation presented here can also be used to study the nonlocal electric and thermal conductivities and viscosities, allowing insight into the diffusive and sound excitations in the hydrodynamic regime[48].

Acknowledgements: We are grateful to Andrey V. Chubukov, Leonid S. Levitov, Dimitrii L. Maslov, and Alexander D. Mirlin for stimulating discussions and to the European Commission's Horizon 2020 RISE program *Hydrotronics* for support.

---

[1] S. Chandrasekhar, *Stochastic Problems in Physics and Astronomy*, Rev. Mod. Phys. 15, 1 (1943).

[2] M. N. Rosenbluth, W. M. MacDonald, and D. L. Judd, *Fokker-Planck Equation for an Inverse-Square Force*, Phys. Rev. **107** 1 (1957).

[3] D. F. Chernoff and M. D. Weinberg, Evolution of globular clusters in the Galaxy, Astrophysical Journal **351**,121 (1990).

[4] S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press 1999).

[5] J. A. Hertz, *Quantum critical phenomena*, Phys. Rev. B **14**, 1165 (1976).

[6] T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism*, Springer Series in Solid-State Science 52, Springer, Berlin-Heidelberg (1984).

[7] A. J. Millis, *Effect of a nonzero temperature on quantum critical points in itinerant fermion systems,* Phys. Rev. B **48**, 7183 (1993).

[8] R. B. Laughlin, G. G. Lonzarich, P. Monthoux, and D. Pines, *The Quantum Criticality Conundrum*, Adv. Phys. 50, 361 (2001).

[9] A. J. Millis, A. J. Schofield, G. G. Lonzarich, and S. A. Grigera, *Metamagnetic Quantum Criticality*, Phys. Rev. Lett. 88, 217204 (2002).

[10] Ar. Abanov, A. V. Chubukov, and J. Schmalian, *Quantum-critical theory of the spin-fermion model and its application to cuprates: Normal state analysis*, Adv. Phys. **52**, 119 (2003).

[11] H. v. Löhneysen, A. Rosch, M. Vojta, and P. Wölfle, *Fermi-liquid instabilities at magnetic quantum phase transitions*, Rev. Mod. Phys. **79**, 1015 (2007).

[12] L. Dell'Anna and W. Metzner, Electrical Resistivity near Pomeranchuk Instability in Two Dimensions, Phys. Rev. Lett. 98, 136402 (2007).

[13] M. A. Metlitski and S. Sachdev, Quantum phase transitions of metals in two spatial dimensions. I. Ising-nematic order, Phys. Rev. B **82**, 075127 (2010).

[14] Y. Schattner, S. Lederer, S. A. Kivelson, and E. Berg, *Ising Nematic Quantum Critical Point in a Metal: A Monte Carlo Study*, Phys. Rev. X **6**, 031028 (2016).

[15] K. Damle and S. Sachdev, *Nonzero-temperature transport near quantum critical points,* Phys. Rev. B **56**, 8714 (1997).

[16] Daniel E. Sheehy and Jörg Schmalian, Quantum Critical Scaling in Graphene, Phys. Rev. Lett. 99, 226803 (2007).

[17] R. N. Gurzhi, A. N. Kalinenko, and A. I. Kopeliovich Phys. Rev. Lett. **74**, 3872 (1995).

[18] R.N. Gurzhi, A.N. Kalinenko, A.I. Kopeliovich, Surface Science **361/362,** 497 (1996).

[19] D. L. Maslov, V. I. Yudson, and A. V. Chubukov, Phys. Rev. Lett. **106**, 106403 (2011).

[20] H. K. Pal, V. I. Yudson, and D. L. Maslov, *Resistivity of non-Galileian invariant Fermi- and Non-Fermi liquids*, Lithuanian Journal of Physics, **52**, 142 (2012).

[21] D. L. Maslov and A. V. Chubukov, Optical response of correlated electron systems, Rep. Prog. Phys. **80**, 026503 (2017).

[22] P. J. Ledwith, H. Guo, L. Levitov, *The Hierarchy of Excitation Lifetimes in Two-Dimensional Fermi Gases*, arXiv:1905.03751.

[23] L. Fritz, J. Schmalian, M. Müller and S. Sachdev, *Quantum critical transport in clean graphene*, Phys. Rev. B **78**, 085416 (2008).

[24] A. B. Kashuba, *Conductivity of defectless graphene*, Phys. Rev. B 78, 085415 (2008).

[25] M. Müller, J. Schmalian and L. Fritz, *Graphene: A Nearly Perfect Fluid*, Phys. Rev. Lett. **103**, 025301 (2009).

[26] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, *Coulomb interaction in graphene: Relaxation rates and transport*, Phys. Rev. B 83, 155441 (2011).

[27] E. I. Kiselev, J. Schmalian, *Boundary conditions of viscous electron flow*, Phys. Rev. B **99**, 035430 (2019).

[28] P. Lévy, P. *Théorie de l'Addition des Variables Aléatoires* Gauthier-Villars, Paris, (1954).

[29] B. Mandelbrot, T*he Fractal Geometry of Nature*, Freeman, New York, (1977).

[30] W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed., Vol. 2, Wiley, New York, (1971).

[31] S. G. Samko, A. A. Kilbas and O. I. Marichev, *Fractional Integrals and Derivatives, Theory and Applications,* Gordon and Breach, Amsterdam, (1993).

[32] R. Hermann, *Fractional Calculus: An Introduction for Physicists* (2nd ed.). New Jersey: World Scientific Publishing (2014).

[33] F. Bartumeus, M. G. E. Da Luz, G. M. Viswanathan, J. Catalan, *Animal search strategies: A quantitative random-walk analysis.* Ecology **86**, 3078 (2005).

[34] A. M. Reynolds and C. J. Rhodes, *The Lévy flight paradigm: random search patterns and mechanisms,* Ecology, 90,, 877 (2009).

[35] R. N. Mantegna and H. E. Stanley, *Econophysics: Scaling and its breakdown in finance. Journal of statistical Physics*, **89**, 469 (1997).

[36] A. Corral, *Universal earthquake-occurrence jumps, correlations with time, and anomalous diffusion* Phys. Rev. Lett. **97**, 178501 (2006).

[37] P. Levy, *L'addition des variables aléatoires définies sur une circonférence,* Bulletin de la Societe mathematique de France, 67, 1 (1939),

[38] Kanti V. Mardia and Peter E. Jupp, *Directional Statistics*, Wiley (1999) ISBN 978-0-471-95333-3.

[39] see supplementary material.

[40] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, *Quantum-critical conductivity of the Dirac fluid in graphene,* Science **364**, 158 (2019).

[41] S. Gattenlöhner, I. V. Gornyi, P. M. Ostrovsky, B. Trauzettel, A. D. Mirlin, M. Titov, *Lévy flights due to anisotropic disorder in graphene.* Phys. Rev. Lett. **117**, 046603 (2016).

[42] U. Briskot, I. A. Dmitriev, and A. D. Mirlin, *Relaxation of optically excited carriers in graphene: Anomalous diffusion and Lévy flights*, Phys. Rev. B **89**, 075414 (2014).

[43] P. Barthelemy, J. Bertolotti, and D. S. Wiersma, *A Lévy flight for light,* Nature **453**, 495 (2008).

[44] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, K. C. Fong, *Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene*, Science **351**, 1058 (2016).

[45] H. Predel, H. Buhmann, L. W. Molenkamp, R. N. Gurzhi, A. N. Kalinenko, A. I. Kopeliovich, A. V. Yanovsky, Phys. Rev. B **62**, 2057 (2000).

[46] K. Wang, M. M. Elahi, L. Wang, K. M. M. Habib, T. Taniguchi, K. Watanabe, J. Hone, A. W. Ghosh, G.-H. Lee, P. Kim, PNAS **116**, 6575 (2019).

[47] A. D. Mirlin and P. Wölfle, *Composite fermions in the fractional quantum Hall effect: Transport at finite wave vector*, Phys. Rev. Lett. **78**, 371 (1997).

[48] E. I. Kiselev and J. Schmalian, unpublished.Supplementary material: Lévy flights and hydrodynamic superdiffusion on the Dirac cone of graphene

Supplementary material

I. SUMMARY OF THE SIMULATIONS SHOWN IN FIG.1

Superdiffusion on the Dirac cone can be seen as a random walk of particles, where the step sizes are distributed according to a wrapped Cauchy distribution. This distribution solves the fractional Fokker-Planck equation (2) of the main text (see [1, 2]). The anglular distance on the Dirac cone travelled by an electron during a time interval Δt is therefore distributed according to

$$\tilde{\psi}\left(\theta,\Delta t\right)=\frac{\sinh\left(\Delta t/\tau_{L}\right)}{\cosh\left(\Delta t/\tau_{L}\right)-\cos\left(\theta\right)}. \tag{22}$$

To generate Figs. 1 a) and b) of the main text we created a sequence of random angles $\Delta\theta_i$ using the distribution (22). The position of the particle after N steps, i.e. after a time intervall $N\Delta t$, then is

$$\theta_N=\sum_{i=0}^{N}\Delta\theta_i. \tag{23}$$

Fig. 4 depicts a wrapped Cauchy random walk with $N=500$ steps. In the case of ordinary diffusion, the step size distribution of Eq. (22) must be replaced by a wrapped normal distribution, which is written in terms of the Jacobi theta function, but can be closely approximated by the van Mises distribution (see e.g. Ref. [3]):

$$\tilde{\psi}_{\text{normal}}\left(\theta,\Delta t\right)=\frac{e^{\cos(\theta)/\Delta t}}{2\pi I_0\left(1/\Delta t\right)}.$$

Using the described procedure we obtain the wrapped Gaussian random flight shown in Fig. 5.

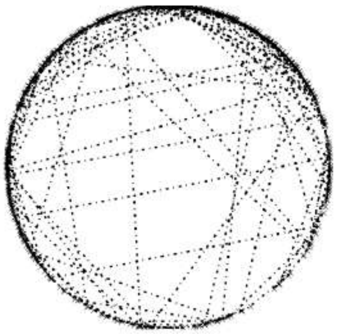


Figure 4: 500 steps of a superdiffusive wrapped Cauchy random walk of an electron on the Dirac cone.

Figure 5: 500 steps of a wrapped Gaussian random walk. The wrapped normal distribution was approximated by the von Mises distribution.

II. COLLISION OPERATOR DUE TO ELECTRON-ELECTRON COULOMB INTERACTION

We briefly summarize the main steps in deriving the collision operator of the Boltzmann equation used in this paper. The collision operator is determined from the larger and smaller self energies on the Keldysh contour. For further details, see Ref. [4].

The non-interacting part of the Hamiltonian is

$$H_0 = v\hbar \int_k \sum_{\alpha\beta} \psi^\dagger_\alpha (\mathbf{k}) \left(\mathbf{k}\cdot\sigma\right)_{\alpha\beta} \psi_\beta (\mathbf{k}) \tag{24}$$

which is diagonalized by the unitary transformation

$$U_{\mathbf{k}} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{k_x+ik_y}{k} & 1 \\ -\frac{k_x+ik_y}{k} & 1 \end{pmatrix} \tag{25}$$

with

$$U_{\mathbf{k}} v\hbar\mathbf{k}\cdot\sigma U^{-1}_{\mathbf{k}} = \begin{pmatrix} v\hbar k & 0 \\ 0 & -v\hbar k \end{pmatrix}. \tag{26}$$

The eigenvalues of the Hamiltonian are $\pm v\hbar k$. Thus we obtain quasiparticle states for the two bands: $\gamma_{\mathbf{k}} = U_{\mathbf{k}}\psi_{\mathbf{k}}$ with

$$H_0 = v\hbar \int_{\mathbf{k}} \sum_{\lambda=\pm} \lambda k \gamma^\dagger_{\mathbf{k},\lambda} \gamma_{\mathbf{k},\lambda}. \tag{27}$$

The electron-electron Coulomb interaction is

$$H_{\text{int}} = \frac{1}{2} \int_{k,k',q} \sum_{\alpha\beta} V(\mathbf{q})\, \psi^{\dagger}_{\alpha}(\mathbf{k}+\mathbf{q},t)\, \psi^{\dagger}_{\beta}(\mathbf{k}'-\mathbf{q},t)\, \psi_{\beta}(\mathbf{k}',t)\, \psi_{\alpha}(\mathbf{k},t) \tag{28}$$

with $V(q) = \frac{e^2}{2\pi\varepsilon|\mathbf{q}|}$. Transforming the interaction into the band, or helical representation, which takes into account the locking between momentum and pseudo-spin that originates from the two sub-lattice structure of graphene. It follows

$$H_{\text{int}} = \frac{1}{2} \int_{k,k',q} \sum_{\alpha\beta} T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \gamma^{\dagger}_{\lambda'}(\mathbf{k}+\mathbf{q},t)\, \gamma^{\dagger}_{\mu}(\mathbf{k}'-\mathbf{q},t)\, \gamma_{\mu'}(\mathbf{k}',t)\, \gamma_{\lambda}(\mathbf{k},t) \tag{29}$$

where

$$T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = V(q) \left(U_{\mathbf{k}+\mathbf{q}} U^{-1}_{\mathbf{k}}\right)_{\lambda'\lambda} \left(U_{\mathbf{k}'-\mathbf{q}} U^{-1}_{\mathbf{k}'}\right)_{\mu\mu'}. \tag{30}$$

Within second order perturbation theory it holds for the self energies for occupied and unoccupied states, respectively.

$$\begin{aligned}
\Sigma^{\gtrless}_{\lambda}(\mathbf{k},\omega) &= N \sum_{\mu\mu'\lambda'} \int \frac{d^2q d^2k' d\omega_1 d\omega_2}{(2\pi)^6} \left|T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\
&\times G^{\gtrless}_{\lambda'}(\mathbf{k}+\mathbf{q},\omega_1)\, G^{\gtrless}_{\mu}(\mathbf{k}'-\mathbf{q},\omega_2)\, G^{\lessgtr}_{\mu'}(\mathbf{k}',\omega_1+\omega_2-\omega) \\
&- \sum_{\mu\mu'\lambda'} \int \frac{d^2q d^2k'}{(2\pi)^4} \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} T_{\lambda\lambda'\mu'\mu}\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right) T_{\lambda\mu\mu'\lambda'}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)^{*} \\
&\times G^{\gtrless}_{\lambda'}(\mathbf{k}+\mathbf{q},\omega_1)\, G^{\gtrless}_{\mu}(\mathbf{k}'-\mathbf{q},\omega_2)\, G^{\lessgtr}_{\mu'}(\mathbf{k}',\omega_1+\omega_2-\omega).
\end{aligned} \tag{31}$$

N combines the valley and spin degrees of freedom and takes the value $N = 4$. Next we use the fact that within a quasiparticle description the upper and lower propagators are expressed in terms of the distribution functions $f_{\lambda}(\mathbf{k},\mathbf{r},t)$ as

$$\begin{aligned}
G^{>}_{\lambda}(\mathbf{k},\mathbf{r},\omega,t) &= -i2\pi\delta\left(\omega-\varepsilon_{\lambda}(\mathbf{k})\right)\left(1-f_{\lambda}(\mathbf{k},\mathbf{r},t)\right) \\
G^{<}_{\lambda}(\mathbf{k},\mathbf{r},\omega,t) &= i2\pi\delta\left(\omega-\varepsilon_{\lambda}(\mathbf{k})\right) f_{\lambda}(\mathbf{k},\mathbf{r},t)
\end{aligned} \tag{32}$$

As usual, $\mathbf{k}$ and ω stand for the Fourier-transformed variables of the relative coordinates and times while $\mathbf{r}$ and t stand for the center of gravity or mean time.

The collision operator can now we determined from the self energies $\Sigma^{<}$ and $\Sigma^{>}$:

$$\mathcal{C}_{\lambda}(\mathbf{k}) = -i\Sigma^{<}_{\lambda}(\mathbf{k},\varepsilon_{\lambda}(\mathbf{k}))\left(1-f_{\lambda}(\mathbf{k})\right) - i\Sigma^{>}_{\lambda}(\mathbf{k},\varepsilon_{\lambda}(\mathbf{k}))\, f_{\lambda}(\mathbf{k}). \tag{33}$$

For simplicity we only keep the momentum $\mathbf{k}$ and band index λ. Inserting $G^{>}$ and $G^{<}$ into the self energies yields with the linearization Eq.(6) of the main paper the result for the collision operator given in Eq.(8) of the main paper. The matrix elements $\gamma^{(1,2)}_{\mathbf{k},\mathbf{k}',\mathbf{q}}$ of that equation are given as:

$$\begin{aligned}
\gamma_1\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= (N-1)\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 + \frac{1}{2}\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right) - T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\
&\quad - \left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right)\right|^2 \\
\gamma_2\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= (N-1)\left|T_B\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{k}-\mathbf{q}\right)\right|^2 + (N-1)\left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right)\right|^2 \\
&\quad + \left|T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) - T_B\left(\mathbf{k},\mathbf{k}',\mathbf{k}'-\mathbf{q}-\mathbf{k}\right)\right|^2,
\end{aligned} \tag{34}$$

with

$$\begin{aligned}
T_A\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= T_{++++}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{----}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\
&= T_{+--+}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{-++-}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\
&= \frac{V(q)}{4}\left(1+\frac{(K+Q)\,K^{*}}{|\mathbf{k}+\mathbf{q}|\,k}\right)\left(1+\frac{(K'-Q)\,K'^{*}}{|\mathbf{k}'-\mathbf{q}|\,k'}\right)
\end{aligned}$$

and

$$\begin{aligned} T_B\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) &= T_{++--}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) = T_{--++}\left(\mathbf{k},\mathbf{k}',\mathbf{q}\right) \\ &= \frac{V\left(q\right)}{4}\left(1-\frac{\left(K+Q\right)K^{*}}{\left|\mathbf{k}+\mathbf{q}\right|k}\right)\left(1-\frac{\left(K'-Q\right)K'^{*}}{\left|\mathbf{k}'-\mathbf{q}\right|k'}\right) \end{aligned} \tag{35}$$

Upper-case letters like $K=k_x+ik_y$ etc. combine the two components of the momentum onto a complex variable.

From the same unitary transformation also follows that

$$U_{\mathbf{k}}\sigma U_{\mathbf{k}}^{-1}=\frac{\mathbf{k}}{k}\sigma_z-\frac{\mathbf{k}\times\mathbf{e}_z}{k}\sigma_y. \tag{36}$$

This can be used to analyze the current

$$\mathbf{j}=ev\int_{\mathbf{k}}\psi^{\dagger}\left(\mathbf{k}\right)\sigma\psi\left(\mathbf{k}\right) \tag{37}$$

of Dirac particles which consists of intra- and inter-band contributions:

$$\mathbf{j}=\mathbf{j}_{\text{intra}}+\mathbf{j}_{\text{inter}}. \tag{38}$$

The two terms are given as

$$\begin{aligned} \mathbf{j}_{\text{intra}} &= ev\int_{\mathbf{k}}\sum_{\lambda=\pm}\frac{\lambda\mathbf{k}}{k}\gamma_{\mathbf{k},\lambda}^{\dagger}\gamma_{\mathbf{k},\lambda} \\ \mathbf{j}_{\text{inter}} &= iev\int_{\mathbf{k}}\frac{\mathbf{k}\times\mathbf{e}_z}{k}\left(\gamma_{\mathbf{k},+}^{\dagger}\gamma_{\mathbf{k},-}-\gamma_{\mathbf{k},-}^{\dagger}\gamma_{\mathbf{k},+}\right). \end{aligned} \tag{39}$$

Thus, the velocity used in our Eq. (4) of the main paper is precisely the expression $\mathbf{v}_{\mathbf{k}\lambda}=v\frac{\lambda\mathbf{k}}{k}$ of the intraband current $\mathbf{j}_{\text{intra}}$. Spin-momentum locking is included naturally, if one goes to the helical states of the upper and lower Dirac cone. The hydrodynamic response is governed by strong collisions of intraband excitations.

III. IDENTIFICATION OF THE COLLINEAR MODES AT FINITE ANGULAR MOMENTUM

In this section we determine the zero modes of the collision operator of Eq.(8) if we confine ourselves to collinear collision processes. To this end we need to find under what conditions the two expressions

$$\begin{aligned} A_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda}^{(1)} &= \psi_{\mathbf{k}+\mathbf{q}\lambda}+\psi_{\mathbf{k}'-\mathbf{q}\lambda}-\psi_{\mathbf{k}'\lambda}-\psi_{\mathbf{k}\lambda} \\ A_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda}^{(2)} &= \psi_{\mathbf{k}+\mathbf{q}\lambda}-\psi_{-\mathbf{k}'+\mathbf{q}\bar{\lambda}}+\psi_{-\mathbf{k}'\bar{\lambda}}-\psi_{\mathbf{k}\lambda}, \end{aligned} \tag{40}$$

that occur in Eq.(8), vanish. Here we have to include the additional constrain

$$\left|\mathbf{k}+\mathbf{q}\right|+\left|\mathbf{k}'-\mathbf{q}\right|=\left|\mathbf{k}\right|+\left|\mathbf{k}'\right| \tag{41}$$

that follows from energy conservation.

By collinear modes we mean that all involved momenta are either parallel or antiparallel. As discussed in Ref.[4] we consider such zero modes of collinear processes because all other processes are suppressed by $1/\log\left(1/\alpha\right)$ where α is the fine-structure constant. Of course, the analysis allows for scattering processes that are not collinear; the issue is merely that distribution functions that become zero modes for collinear scattering are enhanced relative to those that are no such zero modes. Finally we comment that the main conclusion of our paper, namely that the scattering rate depends on angular momentum like $\tau_m^{-1}\propto k_BT\left|m\right|$, is unchanged if we go beyond the collinear mode regime.

One immediately finds that $A^{(1)}=B^{(1)}=0$ subject to Eq.41 is obeyed by $\psi_{\mathbf{k},\lambda}=1$, $\psi_{\mathbf{k},\lambda}=\mathbf{k}$, and $\psi_{\mathbf{k},\lambda}=\lambda\left|\mathbf{k}\right|$, regardless of whether we confine ourselves to collinear modes. These modes correspond to the conservation of charge, momentum, and energy, respectively. In addition to these modes one also finds $\psi_{\mathbf{k},\lambda}=\lambda$ is a zero mode. It corresponds to the fact that second order perturbation theory does not relax a charge imbalance between the upper and lower Dirac cone.

Next we consider distribution functions

$$\psi_{\mathbf{k},\lambda}=a_{\lambda,m}\left(k\right)e^{im\theta_{\mathbf{k}}}, \tag{42}$$

where $k = |\mathbf{k}|$ is the magnitude of the momentum and $\theta_{\mathbf{k}}$ its polar angle: $\mathbf{k} = k\left(\cos\theta_{\mathbf{k}}, \sin\theta_{\mathbf{k}}\right)$. Collinear scattering corresponds to

$$\theta_{\mathbf{k}} = \theta_{\mathbf{k}'} + \mathbf{s_1}\pi = \theta_{\mathbf{k}+\mathbf{q}} + \mathbf{s_2}\pi = \theta_{\mathbf{k}'-\mathbf{q}} + \mathbf{s_3}\pi, \tag{43}$$

where even or odd s_i correspond to parallel and antiparallel momenta relative to $\mathbf{k}$. We first show that all s_i are even due to energy conservation. To this end we assume without restriction that $\mathbf{k} = (k, 0)$ with $k > 0$. Then $\mathbf{k}' = u\,(k, 0)$ and $\mathbf{q} = w\,(k, 0)$, where we do not assume that u and w are positive. Energy conservation now implies

$$1 + |u| = |1 + w| + |u - w|\,. \tag{44}$$

We need to fulfill this condition for an extended set of variables, not just for an isolated point in momentum space. This implies that $1 + w > 0$ so we can cancel the "1" on both sides. Then, to be able to cancel w it must hold that $u > w$, which in turn implies $u > 0$ to cancel u. Thus, we find that the momenta $\mathbf{k}'$, $\mathbf{k} + \mathbf{q}$, and $\mathbf{k}' - \mathbf{q}$ point in the same direction as $\mathbf{k}$ even though $\mathbf{q}$ is allowed to point in the opposite direction. It follows that we can assume without restriction that

$$\theta_{\mathbf{k}} = \theta_{\mathbf{k}'} = \theta_{\mathbf{k}+\mathbf{q}} = \theta_{\mathbf{k}'-\mathbf{q}}. \tag{45}$$

If we use that $e^{im\theta_{-\mathbf{p}}} = (-1)^m\, e^{im\theta_{\mathbf{p}}}$we obtain

$$\begin{aligned}
A^{(1)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= \left(a_{\lambda,m}\left(|\mathbf{k}+\mathbf{q}|\right) + a_{\lambda,m}\left(|\mathbf{k}'-\mathbf{q}|\right) - a_{\lambda,m}\left(k'\right) - a_{\lambda,m}\left(k\right)\right) e^{im\theta_{\mathbf{k}}} \\
A^{(2)}_{\mathbf{k},\mathbf{k}',\mathbf{q},\lambda} &= \left(a_{\lambda,m}\left(|\mathbf{k}+\mathbf{q}|\right) - (-1)^m\, a_{\bar{\lambda},m}\left(|\mathbf{k}'-\mathbf{q}|\right) + (-1)^m\, a_{\bar{\lambda},m}\left(k'\right) - a_{\lambda,m}\left(k\right)\right) e^{im\theta_{\mathbf{k}}}.
\end{aligned} \tag{46}$$

It is now easy so find that there are the following solutions that yield $A^{(1)} = B^{(1)} = 0$ subject to Eq.41:

$$\begin{aligned}
a_{\lambda,m}\left(k\right) &= 1 \\
a_{\lambda,m}\left(k\right) &= \lambda.
\end{aligned} \tag{47}$$

In addition one finds $a_{\lambda,m}\left(k\right) = \lambda\,|\mathbf{k}|$ if m is even and $a_{\lambda,m}\left(k\right) = |\mathbf{k}|$ if m is odd.

Thus, we can write that the following modes are zero modes in the collinear scattering limit

$$\psi_{\mathbf{k},\lambda} = \lambda^m e^{im\theta_{\mathbf{k}}}\left(1, \lambda, \lambda\,|\mathbf{k}|\right) \tag{48}$$

which is Eq.(9) of the main paper.

IV. SUPERDIFFUSION AND NON-LOCAL TRANSPORT

The Fokker-Planck equation (2) of the main text

$$\left(\partial_t + \mathbf{v}_{\mathbf{k}\lambda}\cdot\nabla_{\mathbf{x}} - \tau_L^{-1}\left(\frac{\partial^2}{\partial\theta^2}\right)^{1/2}\right) f_{\mathbf{k}\lambda} = S_{\mathbf{k}\lambda} \tag{49}$$

can be used to calculate the response of graphene electrons to an external perturbation, such as for example an electric field. In this case the force term is given by

$$S_{\mathbf{k}\lambda} = -e\mathbf{E}\left(\mathbf{q},\omega\right)\cdot\frac{\partial f_{\mathbf{k}\lambda}}{\partial\mathbf{k}},$$

where $\mathbf{E}\left(\mathbf{q},\omega\right) = E_0 e^{i(\mathbf{q}\cdot\mathbf{x}-\omega t)}\mathbf{e}_x$ is the electric field. To first order in the electric field it is

$$S_{\mathbf{k}\lambda} = -eE_0 e^{i\mathbf{q}\cdot\mathbf{x}}\cos\theta\left(\lambda\hbar v\beta\right) f_k^{(0)}\left(1 - f_k^{(0)}\right).$$

We perform a Fourier transform $t \to \omega$, $\mathbf{x} \to \mathbf{q}$ and project the equation (49) onto the collinear zero modes (48) using the scalar product $\langle\phi\,|\,\chi\rangle = \sum_\lambda \int \frac{d^2k}{(2\pi)^2}\phi_{\boldsymbol{k},\lambda}\chi_{\boldsymbol{k},\lambda}$. The result is a simplified version of the Boltzmann equation:

$$-i\omega\delta_{m,m'} + \frac{1}{2}ivq\left(e^{-i\vartheta_q}\delta_{m,m'+1} + e^{i\vartheta_q}\delta_{m,m'-1}\right) + \frac{1}{\tau_L\,|m|} = -\frac{1}{2}eE_0 v\beta\delta_{|m|,1}, \tag{50}$$

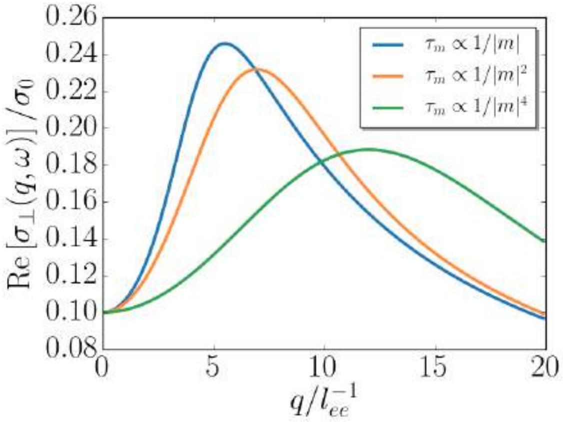


Figure 6: Real part of the transverse conductivity $\sigma_\perp(\omega, q)$ for different dependecies of the scattering times τ_m on m. This result was obtain by solving Eq. (50) numerically.

where m labels the angular harmonic of the collinear zero mode and $\vartheta_{\boldsymbol{q}}$ is the angle of the wave vector $\mathbf{q}$ with respect to the x-axis. This equation is exact in the limit of a small coupling constant α, where collinear events dominate the electron-electron scattering [4]. Notice, that the electric field only couples to angular harmonics with $|m| = 1$, however for a spatially inhomogeneous field with $q \neq 0$, the second right hand side term of Eq. (50) couples all angular harmonics. Therefore, information on the m-dependence of the scattering times can be extracted from the non-local (i.e. $\mathbf{q}$-dependent) electric conductivity $\sigma_{\alpha\beta}(\mathbf{q}, \omega)$, which is defined through

$$j_\alpha(\mathbf{q}, \omega) = \sigma_{\alpha\beta}(\mathbf{q}, \omega) E(\mathbf{q}, \omega).$$

The conductivity tensor $\sigma_{\alpha\beta}(\mathbf{q}, \omega)$ can be decomposed into longitudinal and transverse parts according to

$$\sigma_{\alpha\beta} = \frac{q_\alpha q_\beta}{q^2}\sigma_\parallel(\omega, q) + \left(\delta_{\alpha\beta} - \frac{q_\alpha q_\beta}{q^2}\right)\sigma_\perp(\omega, q),$$

where $\sigma_{\parallel/\perp}(\omega, q)$ only depend on the magnitude of $\mathbf{q}$. Fig 6 shows the influence of the m-dependence of the scattering time τ_m on the real part of the transverse conductivity $\sigma_\perp(\omega, q)$. We conclude that the non-local conductivity, playing an important role in experiments on surface acoustic waves [5, 6], provides a possibility to detect the peculiar dependence of the scattering times τ_m on m, and to confirm the Lévy flight behavior predicted in the main text. For completeness, we mention that the Boltzmann equation (50) can be solved exactly, and the expressions for the non-local conductivities can be written down in closed form [7]:

$$\begin{aligned} \sigma_\parallel(q, \omega) &= \frac{\sigma_0}{1 + i\tau_1\omega - \frac{1}{4}v^2\tau_1 q^2\left(\frac{2i}{\omega} - \frac{1}{M(q,\omega)+i\omega}\right)}, \\ \sigma_\perp(q, \omega) &= \frac{\sigma_0}{1 + i\tau_1\omega + \frac{\frac{1}{4}v^2\tau_1 q^2}{M(q,\omega)+i\omega}}. \end{aligned} \tag{51}$$

Here, $\sigma_0 = N\frac{e^2\log(2)\tau_{\sigma,1}}{2\pi\beta\hbar^2}$ is the quantum critical conductivity calculated in Ref. [4] and $M(q, \omega)$ is the memory function summerizing the effects of higher angular harmonics:

$$M(q, \omega) = \tau_2^{-1} + \frac{1}{2}vq\frac{\mathrm{I}_{3+i\omega\tau_L}(\tau_L vq)}{\mathrm{I}_{2+i\omega\tau_L}(\tau_L vq)},$$

where $I_\nu(z)$ is the modified Bessel function.

---

[1] A. V. Chechkin, V. Y. Gonchar, J. Klafter, R. Metzler, L. V. Tanatarov, J. Stat. Phys. **115**, 1505 (2004).

[2] R. Metzler, A. V. Chechkin, J. Klafter, Levy Statistics and Anomalous Transport, Computational Complexity, Springer, 2012.

[3] G. Kurz, I. Gilitschenski and U. D. Hanebeck, "Efficient evaluation of the probability density function of a wrapped normal distribution" in 2014 Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE (2014).

[4] L. Fritz, J. Schmalian, M. Müller and S. Sachdev, *Quantum critical transport in clean graphene*, Phys. Rev. B **78**, 085416 (2008).

[5] A. Wixforth, J. P. Kotthaus, and G. Weimann Phys. Rev. Lett. **56**, 2104 (1986).

[6] H. S. Simon, Phys. Rev. B **54**, 13878 (1996).

[7] E. I. Kiselev, J. Schmalian, unpublished.

arXiv:2105.06497v1 [cond-mat.mes-hall] 13 May 2021

Heating of inhomogeneous electron flow in the hydrodynamic regime

Gu Zhang,[1, *] Valentin Kachorovskii,[2] Konstantin Tikhonov,[3, 4] and Igor Gornyi[1, 2]

[1] *Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2] *Ioffe Institute, 194021 St. Petersburg, Russia*
[3] *Skolkovo Institute of Science and Technology, Moscow, 121205, Russia*
[4] *Condensed-matter Physics Laboratory, National Research University Higher School of Economics, 101000 Moscow, Russia*

(Dated: May 17, 2021)

We study the electron temperature profiles for an inhomogeneous electron flow in the hydrodynamic regime. We assume that the inhomogeneity is due to a weakly non-uniform distribution of the momentum relaxation time within a spherically constricted area. We show that the temperature profile dramatically depends on the drive strength and the viscosity of the electron liquid. In the absence of viscosity, a Landauer-dipole-like temperature distribution, asymmetrically deformed along the current by the inelastic electron-phonon scattering, emerges around the inhomogeneity. We find that both the Landauer-dipole temperature profile and its asymmetry in the direction of the driving electric field exist in all dimensionalities and are, therefore, universal features of inhomogeneous hydrodynamic electron flow. We further demonstrate that the electron viscosity suppresses the thermal Landauer dipole and leads to the appearance of a "hot spot" exactly at the center of the constriction. We also calculate the phonon temperature distribution, which can be directly measured in experiments on thermal nanoimaging.

I. INTRODUCTION

The study of low-dimensional electronic systems is a key direction in the condensed matter physics in the last few decades. This is dictated by the general trend in the reduction of the sizes of the electronic devices, and is supported by significant advances in modern technology. The semiconducting heterostructures [1] and graphene [2] are among these technological developments, through which the two-dimensional (2D) electron gas has been experimentally realized and employed for designing nanoelectronic devices.

One of the most important properties of an electronic circuit is its ability to cool efficiently and operate under a sufficiently strong drive, when transport becomes substantially nonequilibrium. Under the non-equilibrium conditions, effects related to overheating, dissipation, and thermalization become decisive for the device functioning. The variety of setups in which such phenomena define the physical properties of the system is quite wide, and surprises arise even when studying more conventional structures subject to the drive. The need of developing a comprehensive theory of heat transfer in nanosystems has become particularly evident in recent years. Indeed, miniaturization of electronic devices down to the nanoscale and the use of new materials with unique properties are expected to affect the thermal properties of nanostructures in a crucial way.

As the nanoscales are reached, new effects come into play owing to the increased role of disorder, electron-electron interactions, and their interplay. Further, it is now possible to change smoothly the dimensionality of the system. For example, through etched gates, a 2D electron gas can be divided into multiple areas that are connected by point contacts [3, 4]. Properly selected configuration of gates allows one to create a contact in the form of a quasi-one-dimensional constriction and control the number of channels responsible for current transfer through such a constriction. By using a selective doping, one can engineer inhomogeneous low-dimensional gate-controllable structures to probe charge and heat transport.

At the same time, the development of the SQUID-on-tip (tSOT) [5, 6] and the cryogenic quantum magnetometry [7] techniques enables precise measurement of the temperature profiles and electric current distributions in nanostructures. The former technique has already been successfully applied for the imaging of impurities [8, 9] and the quantum Hall edges states of graphene samples [10–12]. These modern experimental techniques can be applied to the analysis of the influence of various types of nanoscale inhomogeneities (intrinsic, geometrical, artificial) on heat balance in nanodevices.

Recently, a new direction in the physics of low-dimensional systems has emerged—electronic hydrodynamics, which was discussed for many decades [13–16], but was scarcely studied because of the lack of experimental realization at that time. Now this direction is booming thanks to the technological advances in the production of ultraclean ballistic systems, and a number of new hydrodynamic regimes were theoretically predicted (for review, see, e.g., Refs. [17–20]) and experimentally discovered [7, 21–42]. In particular, with the use of modern nanoimaging techniques, it became possible to visualize hydrodynamic flows in 2D materials [7, 27, 32–34].

Despite the experimental advances in the thermal detection, the studies of the hydrodynamic phenomena and transition regimes between hydrodynamics, in which the electron-electron scattering is the fastest, and the drift-diffusion regime, where scattering by disorder dominates, mostly focus on the charge transport. The heat transport features, however, are comparatively less visited [43–45] and overheating of the sufficiently small electron devices—the issue of crucial importance for possible applications—in fact, remains very poorly understood. Importantly, as was pointed out more than twenty years ago [43], the local Joule heating approximation

---

* gu.zhang@kit.edu

does not work for the description of small devices like point contacts. The heat generation there can be governed by nonlocal dissipation processes. However, nonlocality is not the only specific feature of the heat dissipation at the nanoscales.

A recent analysis [46] shows a great variety of different overheating regimes in a quasi-one-dimensional (quasi-1D) constriction [see Fig. 1(a)] with an inhomogeneous distribution of transport scattering rate. One of the distinguishing features of all these regimes is the presence of strongly asymmetric temperature profiles [see Fig. 1(b)]. The heat transfer in 2D graphene with local defects was addressed theoretically in Refs. [47, 48] within the concept of supercollisions—the impurity-stimulated electron-phonon scattering [49, 50] applied to resonant scatterers, as suggested by experimental observations [9]. The spatial distribution of dissipation power was linked there to the formation of Landauer dipoles [51] around the local defect [cf. Fig. 1(d)]. However, only the total dissipation power was calculated in Ref. [47] (in the linear-response regime), without exploring the spatial structure of the local temperature profiles.

In this paper, we study in detail the thermal characteristics of an inhomogeneous 2D sample in the hydrodynamic regime, focusing on the temperature profiles induced by inhomogeneities. We also discuss the thermal properties of 3D inhomogeneous systems within a simplified model. The inhomogeneity is introduced by the presence of a constricted area, where the impurity scattering rate is different from its uniform value (i.e, its value far away from the constriction). We consider a circular-shape constriction [Fig. 1(c)], where the constriction has a higher impurity scattering rate. Similarly to the quasi-1D case, we predict a strong asymmetry in the electronic temperature profiles, even for a perfectly symmetric constriction.

We assume that the inhomogeneity is weak, and study the dc-current-induced variation of temperature, density and drift velocity distributions. One of our main results is the prediction of a the Landauer-dipole character of the temperature distribution (similar to the distribution of the electric field in the original Landauer's consideration [51]). This Landauer-dipole temperature profile is further shifted asymmetrically with increasing the driving current, as in the quasi-1D setting [46]. Remarkably, this Landauer-dipole asymmetric feature universally exists in all dimensionalities [see Fig. 1]. Our second key result is that the temperature profile is dramatically modified by the viscosity of the electron liquid. We demonstrate that the electron viscosity suppresses the Landauer-dipole-like structure and creates a "hot spot" exactly at the center of the constriction. We also discuss the corresponding phonon-temperature profiles that can be experimentally observed using the tSOT technique [5, 8].

The paper is organized as follows. In Sec. II, we describe our model and present the basic equations. In Sec. III, we analyze the current-induced overheating for different dimensionalities, geometries, and interaction models within the ideal fluid model. In Sec. IV, we discuss the viscous case. Section V addresses the phonon temperature. Our results are summarized in Sec. VI. The details of calculations are discussed in Appendices A and B.

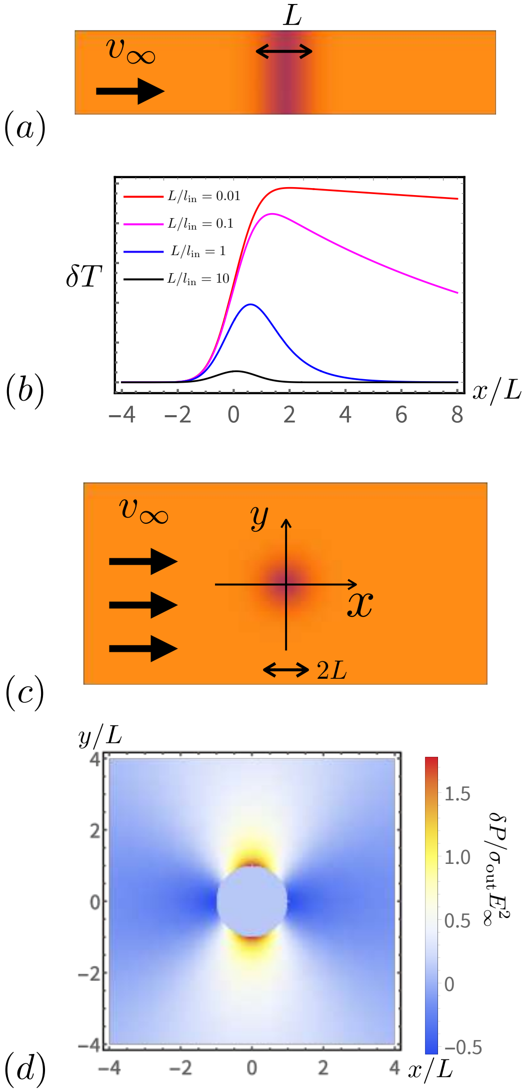


FIG. 1. (a) Quasi-1D symmetric constriction with increased scattering rate and the corresponding asymmetric overheating profile (b). The temperature asymmetry increases when the ratio between the constriction size L and the drift inelastic length $l_{\rm in}$ decreases (see Ref. [46]). The circular-shape constriction with increased impurity scattering in 2D or 3D electron gas (c) driven by the flow along the x direction, with the homogeneous drift velocity v_∞ at $r\to\infty$. The magnitude of the momentum relaxation time τ is shown by the color (the lighter the color, the larger the value of τ). (d) Constriction-induced local heating in a 2D or 3D sample, calculated with the local-heating approximation, see Appendix B.

II. MODEL AND BASIC EQUATIONS

In this section, we introduce the hydrodynamic equations of ideal electron fluid (Sec. II A) and the constriction model (Sec. II B). The formalism adopted in the present paper closely follows and generalizes the one of Ref. [46], where an ideal-fluid description of a quasi-1D constriction was developed.

A. Hydrodynamic approach

We consider a 2D spatially inhomogeneous, weakly disordered electron system, where the inhomogeneity is induced by the dependence of the electron transport scattering time τ on position $\mathbf{r}$. The distribution function $f(\mathbf{r},\mathbf{V})$ obeys the stationary kinetic equation

$$\mathbf{V}\frac{\partial f}{\partial \mathbf{r}}+\frac{\mathbf{F}}{m}\frac{\partial f}{\partial \mathbf{V}}=\widehat{\mathrm{St}}f. \tag{1}$$

Here $\mathbf{V}$ and m are the electron velocity and mass, respectively, $\mathbf{F}$ is the total local force. The right-hand side (r.h.s.) of Eq. (1) is determined by the collision operator

$$\widehat{\mathrm{St}}=\widehat{\mathrm{St}}_{\mathrm{imp}}+\widehat{\mathrm{St}}_{\mathrm{ph}}+\widehat{\mathrm{St}}_{\mathrm{ee}},$$

which contains three terms describing electron-impurity, electron-phonon, and electron-electron scattering. The impurity part of the collision integral reads

$$\widehat{\mathrm{St}}_{\mathrm{imp}}=\frac{f_0-f}{\tau(\mathbf{r})}, \tag{2}$$

where $f_0=\langle f\rangle_\theta$ is the distribution function averaged over the velocity angle θ. Below, we assume that the transport relaxation time $\tau(\mathbf{r})$ is constant everywhere except for a certain constricted area. We model the electron-phonon scattering by the following Fokker-Planck-type collision integral

$$\widehat{\mathrm{St}}_{\mathrm{ph}}=\gamma\frac{\partial}{\partial\epsilon}\left\{\epsilon\left[f_0(1-f_0)+T_0\frac{\partial}{\partial\epsilon}f_0\right]\right\}, \tag{3}$$

where γ is the electron-phonon scattering rate, ϵ is the single-particle energy, and T_0 is the phonon temperature. We assume $\gamma\tau\ll 1$. The specific form of the electron-electron collision operator, $\widehat{\mathrm{St}}_{\mathrm{ee}}$, is not important to us. We consider the hydrodynamic limit where the electron-electron collision time, τ_{ee}, is much shorter than the momentum relaxation time: $\tau_{\mathrm{ee}}\ll\tau$. We also assume that $\tau_{\mathrm{ee}}\ll 1/\gamma$.

First, we consider an ideal electron liquid neglecting the electron viscosity and the heat conductivity, which are both proportional to τ_{ee}. The viscous liquid will be discussed in Sec. IV. In the limit $\tau_{\mathrm{ee}}\to 0$, one can use the hydrodynamic Ansatz for the distribution function

$$f(\mathbf{r},\mathbf{V})=\frac{1}{\exp\left[\frac{m[\mathbf{V}-\mathbf{v}(\mathbf{r})]^2/2-\mu(\mathbf{r})}{T(\mathbf{r})}\right]+1}, \tag{4}$$

where functions $\mu(\mathbf{r})$, $T(\mathbf{r})$, and $\mathbf{v}(\mathbf{r})$ are the local values of chemical potential, electron temperature, and the drift velocity, respectively. The equations governing these collective (hydrodynamic) variables are obtained by averaging Eq. (1) multiplied by 1, $\mathbf{V}$, and mV^2, respectively, over the velocity $\mathbf{V}$ (see, for example, Appendix A in Ref. [46]). After some algebra, we get the following set of equations:

$$\mathrm{div}\,(\mathbf{v}N)=0, \tag{5}$$

$$(\mathbf{v}\nabla)\mathbf{v}+\frac{\mathbf{v}}{\tau}=\frac{1}{m}\left(\mathbf{F}_0+\delta\mathbf{F}-\frac{1}{N}\nabla W\right), \tag{6}$$

$$C\,\mathrm{div}(\mathbf{v}T)=N\left[\frac{mv^2}{\tau}-\gamma\,(T-T_0)\right]. \tag{7}$$

Here, $N=N(\mathbf{r})$ is the local electron concentration, related to the local chemical potential as follows (ν is the 2D density of states):

$$\mu=T\ln\left[\exp\left(\frac{N}{\nu T}\right)-1\right]. \tag{8}$$

The total force is written as a sum of the external force including the driving homogeneous electric force $e\mathbf{E}_0$ and the Lorenz force in the external homogeneous magnetic field $\mathbf{B}$,

$$\mathbf{F}_0=e\left(\mathbf{E}_0+\left[\frac{\mathbf{v}}{c}\times\mathbf{B}\right]\right),$$

and the inhomogeneity-induced correction $\delta\mathbf{F}$ that depends on the electrostatics of the problem. For the gated case, when interaction is screened beyond the gate-to-channel distance d, this correction reads

$$\delta\mathbf{F}=-\frac{e^2\nabla N}{\mathcal{C}}, \tag{9}$$

where $\mathcal{C}=\varepsilon/4\pi d$ is the gate-to-channel capacitance per unit area, and ε is the dielectric constant. For comparison, in the case of unscreened long-range Coulomb interaction, the correction reads:

$$\delta\mathbf{F}=-e^2\nabla\int d^2\mathbf{r}'\,\frac{N(\mathbf{r}')}{\varepsilon|\mathbf{r}-\mathbf{r}'|}.$$

The term $N^{-1}\nabla W$ on the r.h.s. of Eq. (6) represents the thermoelectric force, where

$$W=W(N,T)=\int_0^\infty d\epsilon\;\epsilon\;f_F(\epsilon)$$

is the density of energy in the frame moving with flow. Here, $f_F=1/\exp[(\epsilon-\mu)/T+1]$ is the Fermi function and μ should be expressed via N by using Eq. (8). In the limiting cases, this function is given by

$$W\approx\begin{cases}\dfrac{N^2}{2\nu}+\dfrac{\pi^2\nu T^2}{6}, & \text{for}\quad \nu T\ll N,\\[2mm] TN, & \text{for}\quad \nu T\gg N.\end{cases} \tag{10}$$

The temperature balance equation (7) contains the heat capacity defined as

$$\begin{aligned}C&=C(N,T)=(\partial W/\partial T)_{N=\mathrm{const}}\\&\approx\begin{cases}\dfrac{\pi^2\nu T}{3}, & \text{for}\quad \nu T\ll N,\\[2mm] N, & \text{for}\quad \nu T\gg N.\end{cases}\end{aligned} \tag{11}$$

The term $-\gamma(T-T_0)$ on the r.h.s. of Eq. (7) governs the heat transfer from the electron system with temperature T to the phonon bath with temperature T_0. Strictly speaking, the phonon temperature T_0 also depends on $\mathbf{r}$. Throughout most of the paper we neglect this dependence assuming that T_0 is fixed by fast heat exchange with the substrate. A brief discussion of the dependence $T_0(\mathbf{r})$ is presented in Sec. V.

B. Model of inhomogeneity

In this paper, we consider a simple model of inhomogeneity, which allows for an analytical solution. First of all, we assume that the sample inhomogeneity is completely governed by the spatial dependence of the momentum relaxation time, $\tau(\mathbf{r})$, within a certain constricted area, while other coupling parameters and the external fields are position-independent. We limit ourselves to the weak-inhomogeneity limit, for which the dimensionless parameter

$$\xi(\mathbf{r})=\frac{\tau_\infty}{\tau(\mathbf{r})}-1 \tag{12}$$

is small, $\xi \ll 1$, and search for solutions of Eqs. (5), (6), and (7) perturbatively, to the leading order in ξ. In Eq. (12), τ_∞ is the momentum relaxation time far away from the constricted area.

Recently, the quasi-1D case was analyzed in Ref. [46], where the temperature distribution of an infinitely long stripe-shape sample was found. In the quasi-1D geometry, the parameter ξ only depends on the coordinate x along the strip and is nonzero only in a limited area (see Fig. 1a). The asymmetry of temperature distribution in a variety of different overheating regimes, both in the hydrodynamic and in the impurity-dominated (the so-called drift-diffusion) cases, was predicted [see Fig. 1(b)]. This asymmetry is captured neither by the conventional local Joule heating approximation, nor a more advanced approach of Ref. [43]), and reveals itself in the regime of sufficiently strong non-equilibrium.

In this paper, we consider the hydrodynamic heat transport in higher dimensionalities, mostly focusing on 2D cases. We study the heat flow through a spherical-symmetric constriction characterized by the transport relaxation time τ_0 at the center of the constriction and the radius L. For definiteness, when illustrating our results we use the Gaussian shape of the inhomogeneity:

$$\xi(\mathbf{r})=\left(\frac{\tau_\infty}{\tau_0}-1\right)\exp\left(-\frac{r^2}{L^2}\right). \tag{13}$$

III. CURRENT-INDUCED HEATING

A. Homogeneous heating

We start our analysis with the homogeneous case, $\tau(\mathbf{r})=\tau_\infty$, $\xi\equiv 0$. We fix the electric field at $|\mathbf{r}|\to\infty$ and assume that the magnetic field is parallel to the z axis: $\mathbf{B}=-B\mathbf{e}_z$. From Eqs. (5), (6), and (7) we find the homogeneous velocity $\mathbf{v}_\infty$ and the homogeneous temperature T_∞

$$\mathbf{v}_\infty=\chi\frac{\mathbf{E}_0-\beta\,\mathbf{E}_0\times\mathbf{e}_z}{1+\beta^2},\qquad T_\infty=T_0+\frac{mv_\infty^2}{\tau_\infty\gamma}. \tag{14}$$

Here, $\chi=e\tau_\infty/m$, $\beta=\omega_c\tau_\infty=\tan\theta$, $\omega_c=eB/mc$, θ is the Hall angle,

$$v_\infty=\frac{\chi E_0}{\sqrt{1+\beta^2}}=\chi E_\parallel \tag{15}$$

is the absolute value of velocity, and $E_\parallel=E_0\cos\theta$ is the projection of the electric field on the direction of electric current. Following Ref. [46] we define the dimensionless overheating parameter

$$\alpha=1-\frac{T_0}{T_\infty}=\left(1+\frac{m\gamma T_0}{e^2E_\parallel^2\tau_\infty}\right)^{-1}, \tag{16}$$

which scales quadratically with a weak electric field and saturates at $\alpha=1$ in strong fields.

From Eqs. (14) and (15), we realize that the application of the magnetic field for a fixed direction of external electric field simply introduces a rotation of all the profiles by the Hall angle θ. In what follows, we thus focus only on the case of zero magnetic field.

B. Inhomogeneous heating

In this section, we study the spatial profiles of the hydrodynamic variables (density, hydrodynamic velocity, and temperature) after the introduction of a weak inhomogeneity. For $B=0$, $\theta=\beta=0$, $\mathbf{v}_\infty=\mu\mathbf{E}_0$, and $E_\parallel=E_0$. Below, we consider the two distinct interacting models.

1. *Gated 2D liquid*

Let us assume that the electron-electron interaction is screened by a gate, so that $\delta\mathbf{F}$ is given by Eq. (9) . We introduce small inhomogeneity-induced corrections δn, $\delta\mathbf{v}$, and δT, which are proportional to ξ:

$$N=N_\infty(1+\delta n), \tag{17}$$

$$\mathbf{v}=\mathbf{v}_\infty+\delta\mathbf{v}, \tag{18}$$

$$T=T_\infty+\delta T. \tag{19}$$

Linearizing Eqs. (5), (6), and (7) with respect to ξ and taking Fourier transform we get (for the Fourier transformed quantities):

$$\mathbf{q}\mathbf{v}_\infty\delta n_\mathbf{q}+\mathbf{q}\delta\mathbf{v}_\mathbf{q}=0, \tag{20}$$

$$\left(\frac{1}{\tau_\infty}+i\mathbf{q}\mathbf{v}_\infty\right)\delta\mathbf{v}_\mathbf{q}+i\mathbf{q}\left(s^2\delta n_\mathbf{q}+\frac{\delta T_\mathbf{q}}{M}\right)=-\frac{\mathbf{v}_\infty}{\tau_\infty}\xi_\mathbf{q}, \tag{21}$$

$$\left(\frac{\gamma}{c_\infty}+i\mathbf{q}\mathbf{v}_\infty\right)\delta T_\mathbf{q}+\left(i\mathbf{q}T_\infty-\frac{2M\mathbf{v}_\infty}{\tau_\infty}\right)\delta\mathbf{v}_\mathbf{q}=\frac{Mv_\infty^2}{\tau_\infty}\xi_\mathbf{q}. \tag{22}$$

Here

$$s^2=\frac{e^2N_\infty}{m\mathcal{C}}+\frac{1}{m}\left(\frac{\partial W}{\partial N}\right)_{T=\mathrm{const}} \tag{23}$$

is the plasma wave velocity, with the second term representing contribution from the Fermi-liquid velocity. We also define

$$M=\frac{m}{c_\infty} \tag{24}$$

as the “thermal” mass, and

$$c_\infty = \frac{C_\infty}{N_\infty} = \frac{1}{N_\infty}\left(\frac{\partial W}{\partial T}\right)_{N=\text{const}} \tag{25}$$

is the specific heat capacity.

2. Weakly compressible liquid

General solutions of Eqs. (20), (21), and (22) are rather cumbersome. Below we present solutions for the most interesting case, where the liquid is almost incompressible. This case is realized for strong electron-electron interaction, when the plasma wave velocity is large:

$$s \gg \sqrt{\frac{T_\infty}{M}}, \quad s \gg v_\infty. \tag{26}$$

We introduce two characteristic lengths: the elastic drift length determined by the scattering off disorder,

$$l = v_\infty \tau_\infty, \tag{27}$$

and the inelastic drift length [46] characterizing the electron-phonon scattering,

$$l_{\text{in}} = \frac{v_\infty c_\infty}{\gamma}. \tag{28}$$

We further assume that the elastic drift length is the shortest lengthscale,

$$L \gg l, \qquad l_{\text{in}} \gg l, \tag{29}$$

while the relation between the constriction size, L, and l_{in} can be arbitrary.

From Eq. (21) one can conclude that, in the limit $s \to \infty$, the correction to the electron density is small, $\delta n \propto 1/s^2$. Then, from Eq. (20), we find that $\mathbf{q}\delta\mathbf{v}_\mathbf{q} \propto 1/s^2$, so that the electron liquid is almost incompressible: div$\delta\mathbf{v} \to 0$. As a consequence, to the zeroth order with respect to $1/s^2$, one can replace Eq. (20) with

$$\mathbf{q}\delta\mathbf{v}_\mathbf{q} = 0. \tag{30}$$

Hence, the velocity correction in the momentum space is perpendicular to $\mathbf{q}$:

$$\delta\mathbf{v}_\mathbf{q} \propto \mathbf{t}_\mathbf{q},$$

where $\mathbf{t}_\mathbf{q} = [\mathbf{e}_z \times \mathbf{q}/q]$, and $\mathbf{e}_z$ is the unit vector in the z direction. This correction can be found by taking the scalar product of $\mathbf{t}_\mathbf{q}$ and Eq. (21). Substituting $\delta\mathbf{v}_\mathbf{q}$ into Eq. (22) and neglecting the small term $\mathbf{q}\delta\mathbf{v}_\mathbf{q}$, one can find $\delta T_\mathbf{q}$. Having in mind Eq. (29), we finally arrive at the following expressions for the velocity and temperature corrections:

$$\delta\mathbf{v}_\mathbf{q} = -\xi_\mathbf{q}\mathbf{t}_\mathbf{q}(\mathbf{t}_\mathbf{q}\mathbf{v}_\infty), \tag{31}$$

$$\delta T_\mathbf{q} = \frac{M\xi_\mathbf{q} v_\infty}{\tau_\infty}\frac{q_\parallel^2 - q_\perp^2}{q^2(iq_\parallel + 1/l_{\text{in}})}. \tag{32}$$

Here, $q_\parallel$ and $q_\perp$ are, respectively, the parallel and perpendicular components of the vector $\mathbf{q}$ with respect to the drift velocity $\mathbf{v}_\infty$.

A correction to the concentration arises only in the first order with respect to $1/s^2$. It can be obtained by taking the scalar product of Eq. (21) and $\mathbf{q}$:

$$\delta n_\mathbf{q} = \frac{\xi_\mathbf{q} v_\infty}{s^2\tau_\infty q^2}\left(iq_\parallel - \frac{q_\parallel^2 - q_\perp^2}{iq_\parallel + 1/l_{\text{in}}}\right). \tag{33}$$

Importantly, although $\delta n_\mathbf{q}$ approaches zero for large s, one cannot neglect $\delta n_\mathbf{q}$ from the very beginning, because $\delta F_\mathbf{q} \propto s^2\delta n_\mathbf{q}$ remains finite for $s \to \infty$.

We see that the temperature distribution in the momentum space is described (up to a constant coefficient) by the product of $\xi_\mathbf{q}$ and the heating kernel

$$K(\mathbf{q}) = \frac{q_\parallel^2 - q_\perp^2}{q_\parallel^2 + q_\perp^2}\frac{1}{iq_\parallel + 1/l_{\text{in}}}. \tag{34}$$

It is worth stressing that this equation is derived in the incompressible limit corresponding to very strong interactions and, therefore, does not contain any specific feature of the 2D system. One can easily show that this form of the heating kernel universally appears in other dimensions. In particular, for quasi-1D strips parallel to the x-axis, where $\xi = \xi(x)$, the transverse wave vector equals zero, $q_\| = 0$, and we arrive at equations derived previously in Ref. [46]:

$$K(q_x) = 1/(iq_x + 1/l_{\text{in}}),$$

see Eq. (27) in that work. By using the same calculations as presented above, one finds that for a 3D case, the heating kernel is also given by Eq. (34), with $q_\perp^2 = q_y^2 + q_z^2$ (for $\mathbf{v}_\infty$ parallel to x axis).

3. Landauer-dipole structure of the heating kernel

Direct calculation of the Fourier transform of (34) yields

$$K(\mathbf{r}) = \int_0^\infty ds e^{-s/l_{\text{in}}} Q_D(x_\parallel - s, \mathbf{r}_\perp), \tag{35}$$

where $Q_D(\mathbf{r})$ is equivalent to the “electric field” of a Landauer dipole [51], which can be written in the universal form for all dimensions:

$$Q_D(\mathbf{r}) = C_D\frac{\partial}{\partial x_\parallel}\left(\frac{x_\parallel}{r^D}\right), \qquad D = 1,2,3, \tag{36}$$

where $C_1 = 1/2$, $C_2 = 1/\pi$, $C_3 = 1/2\pi$ for 1D, 2D and 3D cases. Choosing x axis along the $\mathbf{v}_\infty$ direction, we get the explicit expressions for Landauer dipoles in the heating kernel:

$$Q_D(\mathbf{r}) = \begin{cases} \delta(x), & \text{for} \quad D = 1, \\ \dfrac{y^2 - x^2}{\pi(x^2+y^2)^2}, & \text{for} \quad D = 2, \\ \dfrac{y^2 + z^2 - 2x^2}{2\pi(x^2+y^2+z^2)^{5/2}}, & \text{for} \quad D = 3. \end{cases} \tag{37}$$

These functions obey the following property:

$$Q_1(x) = \int dy\, Q_2(x,y) = \int dy\, dz\, Q_3(x,y,z).$$

In order to illustrate the physics behind the Landauer-dipole-like temperature distributions, we present in Appendix B the discussion of the Joule heat distribution in a simple two-component model within the conventional theory of local Joule dissipation.

As seen from Eq. (35), the heating kernel is given by a Landauer dipole that is spatially-shifted at the distance $\sim l_{\text{in}}$. This shift induces the asymmetry of the Landauer dipole along the direction of the current. Physically, the asymmetry stems from "convection" described by the term $\text{div}(\mathbf{v}T)$ in the heat balance equation which is absent in the theory of local dissipation and in the weak-drive theory of non-local heat transport of Ref. [43]. Interestingly, the kernel Eq. (35) remains finite in the limit $l_{\text{in}} \to \infty$, which means that the temperature distribution in this case is fully determined by convection.

The correction to the electron temperature is given in the coordinate space by

$$\delta T(\mathbf{r}) = \frac{T_0}{l_{\text{in}}}\frac{\alpha}{1-\alpha}\int d^D\mathbf{r}'K(\mathbf{r}-\mathbf{r}')\xi(\mathbf{r}'). \tag{38}$$

The asymmetry of temperature distribution manifests itself at distances of the order of (or smaller than) the inelastic lengths. At larger distances, we get

$$K(\mathbf{r}) \approx l_{\text{in}} Q_D(\mathbf{r}), \quad \text{for} \quad r \gg l_{\text{in}}. \tag{39}$$

Hence, away from the inhomgeneity center, for $r \gg \max(L, l_{\text{in}})$, the temperature distribution can be considered as symmetric and given by the Landauer-dipole profile:

$$\delta T(\mathbf{r}) \approx \frac{\alpha T_0}{(1-\alpha)}\, Q_D(\mathbf{r}) \int \xi(\mathbf{r}')d^D\mathbf{r}', \quad \text{for} \quad r \to \infty. \tag{40}$$

For the case of the Gaussian constriction (13), equations for temperature distribution valid for arbitrary relation between r, L and l_{in} are derived in Appendix A. They can be written as follows.

$$\delta T(\mathbf{r}) = A_D \int_0^\infty \frac{ds}{l_{\text{in}}} e^{-s/l_{\text{in}}} \tilde{Q}_D(x-s, \mathbf{r}_\perp), \tag{41}$$

where

$$A_D = \pi^{D/2}L^D\frac{\alpha}{1-\alpha}\left(\frac{\tau_\infty}{\tau_0}-1\right)T_0, \tag{42}$$

and

$$\begin{aligned}
\tilde{Q}_1 &= \frac{1}{\sqrt{\pi}L}e^{-x^2/L^2}, \\
\tilde{Q}_2 &= \frac{y^2-x^2}{\pi r^4}\left[1-e^{-r^2/L^2}(1+r^2/L^2)\right], \\
\tilde{Q}_3 &= \frac{(y^2+z^2-2x^2)\,\text{Erf}(r/L)}{2\pi r^5} \\
&\quad - \frac{e^{-r^2/L^2}[(y^2+z^2)^2-x^4+L^2(z^2+y^2-2x^2)]}{\pi^{3/2}L^3r^4}.
\end{aligned} \tag{43}$$

Sending $L \to 0$ for fixed r, we reproduce Eqs. (37): $\tilde{Q}_D \to Q_D$ for $L \to 0$.

Above, we have assumed that the elastic drift length l is much smaller in comparison to the constriction size: $l \ll L$. The obtained results can be straightforwardly generalized to the case of arbitrary relation between l and L. As follows from Eqs. (20), (21), and (22), the general expression for the heating kernel, Eq. (34), becomes then

$$K(\mathbf{q}) = \left(\frac{q_\parallel^2-q_\perp^2}{q_\parallel^2+q_\perp^2}+iq_\parallel l\right)\frac{1}{iq_\parallel+1/l_{\text{in}}}\frac{1}{1+iq_\parallel l}. \tag{44}$$

In real space, the heating kernel is again expressed in terms of the Landauer-dipole field:

$$\begin{aligned}
K(\mathbf{r}) &= \frac{l_{\text{in}}}{l_{\text{in}}-l}\int_0^\infty d\rho\left(e^{-\rho/l_{\text{in}}}-e^{-\rho/l}\right)Q_D(x_\parallel-\rho,\mathbf{r}_\perp) \\
&\quad + \delta(\mathbf{r}_\perp)\theta(x_\parallel)\left(e^{-x_\parallel/l}-\frac{l}{l_{\text{in}}}e^{-x_\parallel/l_{\text{in}}}\right).
\end{aligned} \tag{45}$$

Clearly, this reproduces Eq. (35) in the limit $l \to 0$. As another feature, similar to Eq. (35), the kernel Eq. (45) remains finite in the limit of diverging l_{in}:

$$\begin{aligned}
K(\mathbf{r}) &\approx \int_0^\infty d\rho\left(1-e^{-\rho/l}\right)Q_D(x_\parallel-\rho,\mathbf{r}_\perp) \\
&\quad + \delta(\mathbf{r}_\perp)\theta(x_\parallel)e^{-x_\parallel/l}, \quad \text{for} \quad l_{\text{in}} \to \infty.
\end{aligned} \tag{46}$$

C. Spatial profiles of the electron temperature, velocity, and concentration

The spatial dependences of the temperature, as well as the velocity and concentration, are plotted in Fig. 2 for a 2D system with a Gaussian constriction [see Eq. (13)]. We assumed that the liquid is nearly incompressible ($s \to \infty$) and $l \ll l_{\text{in}}$, $l \ll L$. The temperature, velocity and concentration are measured, respectively, in the following units

$$T^* = \left(\frac{\tau_\infty}{\tau_0}-1\right)\frac{\alpha}{1-\alpha}T_0 = \left(\frac{\tau_\infty}{\tau_0}-1\right)(T_\infty-T_0), \tag{47}$$

$$v^* = \left(\frac{\tau_\infty}{\tau_0}-1\right)v_\infty, \tag{48}$$

$$n^* = \left(\frac{\tau_\infty}{\tau_0}-1\right)\frac{L}{l}\frac{v_\infty^2}{s^2}. \tag{49}$$

The profiles of $\delta T/T^*$, $\delta v/v^*$, and $\delta n/n^*$ depend only on the relation between characteristic lengths of the problem, L, l_{in}, and l. Figures 2(a),(b) show the temperature plots $\delta T/T_*$ for $l \ll (L, l_{\text{in}})$ for the two different ratios between L and l_{in}: $L = 2l_{\text{in}}$ for Fig. 2(a) and $L = 0.5 l_{\text{in}}$ for Fig. 2(b).

Panel (a) of Fig. 2 reproduces an almost symmetric Landauer-dipole structure, while panel (b) is much more asymmetric because of a larger value of l_{in}/L. In both panels, the electron temperature is reduced along the current

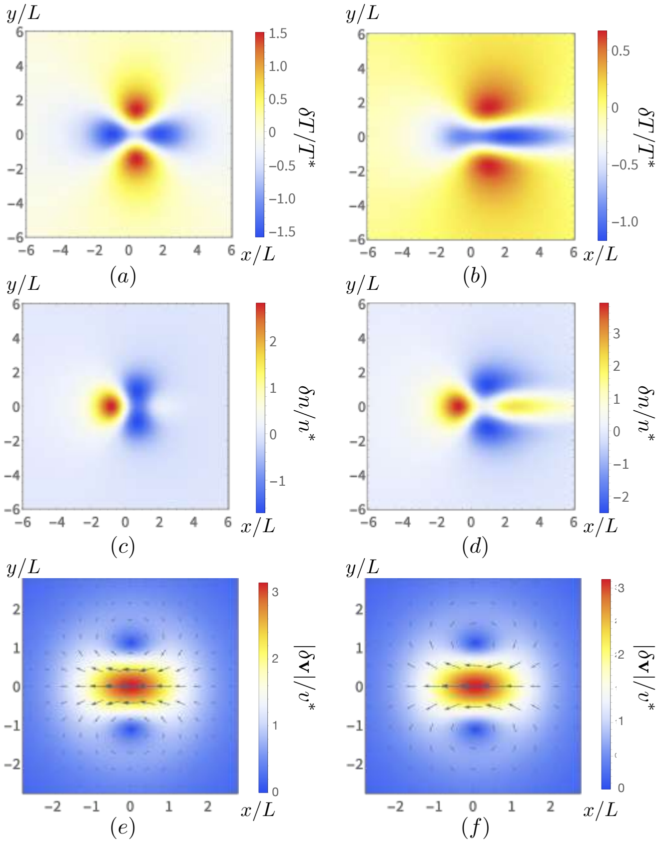


FIG. 2. Density plots of $\delta T/T_*$ in 2D systems with a Gaussian constriction, Eq. (13) in the incompressible limit for $L = 2l_{\rm in}$ (a) and $L = 0.5l_{\rm in}$ (b), respectively; (c) and (d): corresponding density plots of $\delta n/n^*$; (e) and (f): vector density plots of $\delta\mathbf{v}/v^*$.

injecting direction (the x direction), and enhanced in the direction perpendicular to the current (the y direction). The Landauer-dipole structures are shifted by the inelastic drift length $l_{\rm in}$, yielding the temperature asymmetry similar to the one predicted for the quasi-1D geometry [46].

The density plots of $\delta n/n^*$ for the same values of L and $l_{\rm in}$ are shown in panels (c) and (d) of Fig. 2. We observe that among these distributions the most asymmetric is the density profile, which emerges in the limit of nearly incompressible fluid only for finite values of $1/s^2$, in contrast to the temperature and hydrodynamic velocity. The asymmetry of the temperature profiles corresponds to a similar feature of the density distributions: particles tend to gather in front of the constricted area, where a larger impurity scattering rate is present.

The appearance of the Landauer-dipole structure in the temperature distribution can be understood based on the vector density plots for the velocity distribution $\delta\mathbf{v}/v^*$ that are shown in panels (e) and (f) of Fig. 2. One sees that particles tend to detour around the constricted area. Heat current, carried by these detoured particles, thus leads to the hot spots on the y axis, which are clearly seen in panels (a) and (b).

IV. THE VISCOUS CASE

So far, we have considered hydrodynamics of an ideal electron fluid characterized by zero viscosity (taking the limit $\tau_{\rm ee} \to 0$). In the presence of a finite viscosity η, expressions for the velocity and temperature corrections are modified as follows:

$$\delta\mathbf{v}_\mathbf{q} = -\frac{\xi_\mathbf{q}\mathbf{t}_\mathbf{q}(\mathbf{t}_\mathbf{q}\mathbf{v}_\infty)}{1+iq_\parallel l+\kappa q^2} \tag{50}$$

$$\delta T_\mathbf{q} = \frac{M\xi_\mathbf{q}v_\infty}{\tau_\infty}\,\frac{1}{iq_\parallel+1/l_{\rm in}}\left(1-\frac{2q_\perp^2}{q^2}\,\frac{1}{1+iq_\parallel l+\kappa q^2}\right). \tag{51}$$

Here,

$$\kappa = \eta\tau_\infty = l_\kappa^2, \tag{52}$$

where $l_\kappa = \sqrt{\eta\tau_\infty}$ is a new length scale, arising in the presence of viscosity. It is worth noting that viscosity by itself gives a direct contribution to the dissipation [52], so that the term

$$P_{\rm vis} = \eta\iint dxdy\left[(\partial_x v_x-\partial_y v_y)^2+(\partial_x v_y+\partial_y v_x)^2\right], \tag{53}$$

appears on the r.h.s. of the heat balance equation, Eq. (7). Here v_x and v_y are the spatially dependent velocities along the x and y directions, respectively. As seen from Eq. (53), the direct viscous dissipation is present only in the second order in the weak inhomogeneity ξ. Hence, in the linear approximation with respect to ξ, the viscosity affects the temperature distribution indirectly by changing velocity distribution, thus modifying the impurity dissipation mv^2/τ. Depending on the relation between four characteristic lengths: L, l, $l_{\rm in}$, and l_κ, various regimes of dissipation are possible. However, the main effect of viscosity is quite simple: with increasing l_κ, the Landauer-dipole structure is suppressed. It makes sense, therefore, to start with the discussion of the high-viscosity limit.

A. Strong viscosity, $l_\kappa \to \infty$

Let us assume that l_κ is larger than all other length scales in the problem. Interestingly, even in the limit $l_\kappa \to \infty$, there is a finite temperature correction, which can be found by neglecting the last term of Eq. (51). Performing the Fourier transformation, we get

$$\delta T = \frac{Mv_\infty^2}{l}\int_{-\infty}^{x_\parallel} dx'_\parallel\,\xi(\mathbf{r}')e^{-(x_\parallel-x'_\parallel)/l_{\rm in}}. \tag{54}$$

Let us comment on the physics behind this expression. For infinitely large viscosity, all velocity gradients are suppressed, so that $\mathbf{v} \equiv \mathbf{v}_\infty$. In this limit, the heat balance equation is dramatically simplified

$$\frac{d\delta T}{dx_\parallel} = \frac{Mv_\infty^2\,\xi(\mathbf{r})}{l} - \frac{\delta T}{l_{\rm in}}. \tag{55}$$

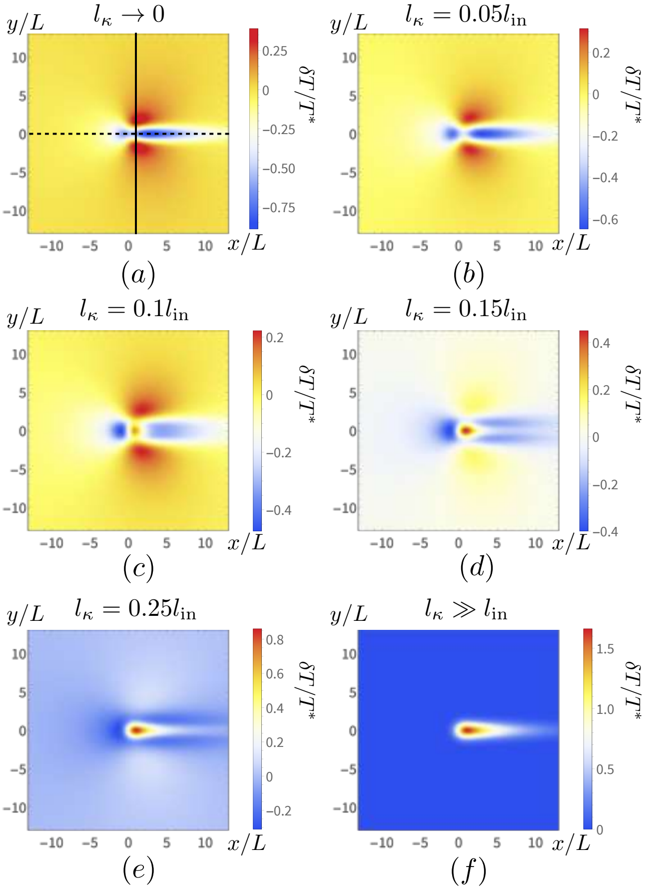


FIG. 3. Evolution of the 2D temperature profiles for fixed L and $l_{\rm in}$, $L = 0.2 l_{\rm in}$, with increasing l_κ: (a) the limit of negligible viscosity, $l_\kappa = 0.001 l_{\rm in}$; (b) small viscosity, $l_\kappa = 0.05 l_{\rm in}$, which is insufficient to suppress the Landauer-dipole feature; (c) "critical" viscosity, $l_\kappa = 0.1 l_{\rm in}$, at which a hot spot appears in the center of the constriction, which increases with further increasing viscosity, as shown in (d) for $l_\kappa = 0.15 l_{\rm in}$ and (e) for $l_\kappa = 0.25 l_{\rm in}$, gradually suppressing the Landauer-dipole structure; (f) strong viscosity, $l_\kappa = 10 l_{\rm in}$: the Landauer-dipole feature is suppressed.

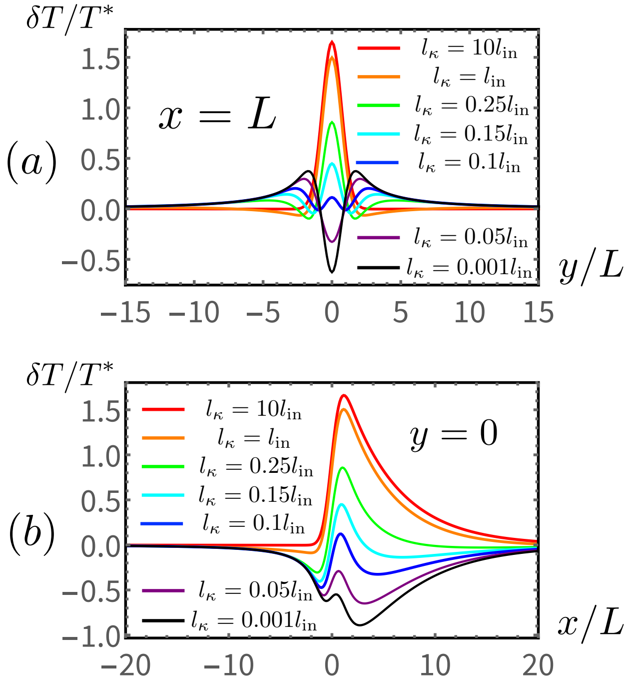


FIG. 4. Evolution of the temperature profiles of the cross-sections at (a) $x = L$ and (b) $y = 0$, corresponding to the solid and dashed lines in Fig. 3(a). Parameters are the same as in Fig. 3.

Physically, this equation means that the heating is effectively one dimensional, regardless of the system dimension. Integrating Eq. (55), we reproduce Eq. (54). We also notice that Eqs. (54) and (55) do not require linearization and are valid for an arbitrary relation between δT and T_∞.

B. Evolution of the temperature profile with viscosity

In Figs. 3 and 4, we illustrate the evolution of the temperature distribution with increasing viscosity. We fix L and $l_{\rm in}$ such that the inelastic drift length is larger than the size of the constriction, $L = 0.2 l_{\rm in}$, and, thus, asymmetry in the temperature distribution is sufficiently pronounced. At the same time, the Landauer-dipole structure remains apparent for zero and very small viscosity, as shown in Figs. 3(a) and (b). When l_κ becomes the order of L, a hot spot appears in the center of the constriction, Figs. 3(c), whose amplitude increases with increasing viscosity, see Figs. 3(d), (e), and (f), while, simultaneously, the Landauer-dipole structure becomes suppressed. At very large values of l_κ, Figs. 3(f), the Landauer-dipole feature is fully suppressed, and the temperature distribution is very well described by Eq. (54).

We also plot in Figs. 4(a), (b) the cross-sections of the temperature distribution along the lines $x = L$ and $y = 0$, respectively. These plots clearly demonstrate the suppression of the Landauer dipoles and the formation of the hot spot, which is symmetric in the y direction but asymmetric in the x direction. Evolution of the velocity profiles with viscosity is shown in Fig. 5 for the same values of parameters as in Fig. 3

We reiterate that, to the leading order in ξ, viscosity does not introduce extra dissipation, but instead redistributes it. To see this point, in Fig. 6 we present the temperature profile ($l_\kappa = l_{\rm in} \gg L$) obtained when we keep only the second term of Eq. (51) in the parentheses [i.e, after removing the contribution of Eq. (54)]. Apparently, the Landauer-dipole structure survives in the strong-viscosity limit, although it is strongly reduced and obscured by the "hot spot" at the constriction.

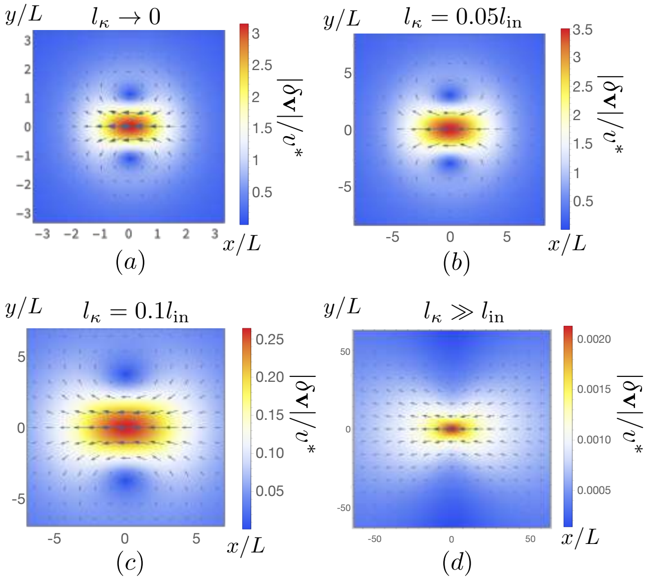


FIG. 5. Evolution of the velocity profiles with increasing viscosity. Panels (a), (b), (c) and (d) correspond to panels (a), (b), (c) and (d) in Fig. 3, respectively.

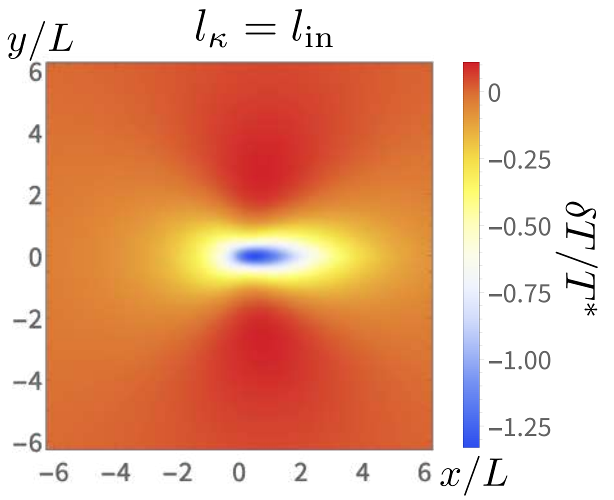


FIG. 6. The temperature profile after the removal of the "hot spot" at the constriction position, see Eq. (54).

V. PHONON TEMPERATURE DISTRIBUTION

In previous sections, we have discussed the distribution of the electron temperature. However, experimentally, the electron temperature profiles are hard to measure directly. On the other hand, the phonon temperature can be measured at a very high resolution by means of the tSOT technique [5, 6, 8–10]. Using the results obtained above, we obtain the phonon temperature, restricting ourselves to the discussion of the 2D case.

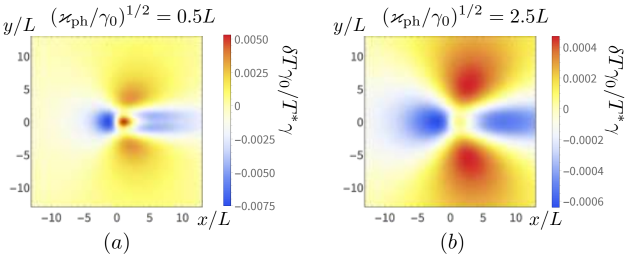


FIG. 7. The phonon temperature near the constriction that corresponds to the electron temperature of Fig. 3(d). The phonon temperature profile reproduces the major features of the electron one when $\varkappa_{\text{ph}}$ is small (a). On the contrary, the "hot spot" disappears when $\varkappa_{\text{ph}}$ becomes large enough (b).

First, we note that the heat exchange rate between phonons and the substrate, γ_0, is typically several orders of magnitude larger than the electron-phonon heat transfer rate γ (see experimental measurements [53–55] and estimates in Ref. [47]). In the homogeneous case we have

$$\gamma(T_{\text{el}} - T_{\text{ph}}) = \gamma_0(T_{\text{ph}} - T_0), \tag{56}$$

where T_0 is the substrate temperature. For $\gamma_0 \gg \gamma$, we have $T_{\text{ph}} \approx T_0$. It is worth noting that this estimate justifies the approach used in the previous sections, where, for determining the electron temperature, we assumed that the lattice temperature is a constant.

Small deviations of the lattice temperature from T_0 can be found from the phonon heat balance equation

$$-\varkappa_{\text{ph}} \Delta T_{\text{ph}} = \gamma(T_{\text{e}} - T_{\text{ph}}) - \gamma_0(T_{\text{ph}} - T_0), \tag{57}$$

where $\varkappa_{\text{ph}}$ is the phonon thermal conductivity. Introducing

$$\delta T_{\text{ph}} = T_{\text{ph}} - T_0,$$

and making use of the inequality $\gamma \ll \gamma_0$, we find

$$\delta T_{\text{ph}}^{\mathbf{q}} = \frac{\gamma}{\gamma_0 + \varkappa_{\text{ph}} q^2} \delta T_{\mathbf{q}}, \tag{58}$$

where $\delta T_{\mathbf{q}}$ is the inhomogeneity-induced correction to the electronic temperature calculated above. We see that for

$$q < q_0 = \sqrt{\gamma_0/\varkappa_{\text{ph}}},$$

the phonon temperature profile coincides with the electronic one up to a small factor γ/γ_0. Very sharp gradiends of the electronic temperature with $q > q_0$ are suppressed by the phonon heat conductivity. In the coordinate representation, Eq. (58) becomes

$$\delta T_{\text{ph}}(\mathbf{r}) = \int d\mathbf{r}' K(\mathbf{r} - \mathbf{r}') \delta T(\mathbf{r}'), \tag{59}$$

where

$$K(\mathbf{r}) = \int \frac{d^2\mathbf{q}\, e^{i\mathbf{q}\mathbf{r}}}{(2\pi)^2} \frac{\gamma}{\gamma_0 + \varkappa_{\text{ph}} q^2} = \frac{\gamma}{2\pi \varkappa_{\text{ph}}} K_0(q_0 r), \tag{60}$$

and K_0 is the MacDonald function. Evolution of the phonon temperature distribution with increasing phonon heat conductivity is illustrated in Fig. 7.

VI. SUMMARY

To summarize, we have studied the heat balance in a weakly disordered system with local inhomogeneities. Specifically, we have considered a spherical local constriction with increased impurity scattering rate as compared to the scattering in the uniform background. We assumed that electron-electron interaction is strong in two senses: firstly, fast electron-electron collisions drive the system into the hydrodynamic regime; secondly, it guarantees the electrical neutrality, thus making the electron liquid nearly incommpressible.

In the absence of viscosity, the electron temperature distribution induced by the inhomogeneity is described by a Landauer-dipole-like structure that is shifted along the current by the amount of the order of the inelastic drift length (Fig. 2). The Landauer-dipole distribution stems from the tendency of particles to travel around the inhomogeneity (in the 2D and 3D cases). Remarkably, the thermal Landauer dipole and its asymmetry, both induced by the current flow, exist in systems of arbitrary dimensionality, and of genuine constriction geometries. Thus, we have found that the heating of inhomogeneous ideal electron fluid is universally described by a Landauer dipole, which is deformed by the flow beyond the linear-response regime.

Further, focusing on the 2D case, we have explored the evolution of the electron temperature and velocity profiles with increasing viscosity of the electron fluid (Fig. 3). Our main conclusion is that the viscosity dramatically changes the heat balance in the system. Specifically, we have found that viscosity suppresses the Landauer-dipole structure and—for sufficiently high viscosity—essentially reduces the heat-balance problem in all dimensions to a quasi-1D problem. Most importantly, viscosity leads to the formation of the "hot spot" exactly in the position of the inhomogeneity. We have also derived a relation between electron temperature and the temperature of the phonon system (Fig. 7), which can be directly measured in experiment.

ACKNOWLEDGMENTS

We thank A. Dmitriev, A. Mirlin, B. Narozhny, and E. Zeldov for discussions. The work of GZ and IG is supported by the German-Israeli Foundation (GIF grant No. I-1505-303.10/2019), the DFG grant (No. GO 1405/5) within the FLAG-ERA Joint Transnational Call (project GRANSPORT), and European Commission under the EU Horizon 2020 MSCA RISE-2019 program (project 873028 HYDROTRONICS). The work of VK and IG was supported by the Russian Science Foundation (Grant No. 20-12-00147) . The work of VK was also supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS".

Appendix A: Gaussian constriction

In this Appendix, we derive the expressions for Landauer-dipole structures in $D = 1, 2, 3$ dimensions. To this end, one needs to calculate the Fourier transform of $K_{\mathbf{q}}\xi_{\mathbf{q}}$, where $K_{\mathbf{q}}$ and $\xi_{\mathbf{q}}$ are given by Eqs. (34) and (13) of the main text, respectively. Introducing auxiliary integrals,

$$\frac{1}{q^2} = \int\limits_0^\infty dt\, e^{-tq^2}, \quad \frac{1}{iq_\parallel + 1/l_{\text{in}}} = \int\limits_0^\infty dse^{-s(iq_\parallel + 1/l_{\text{in}})},$$

we find

$$\begin{aligned} &\int \frac{d^D\mathbf{q}}{(2\pi)^D}\frac{q_\parallel^2 - q_\perp^2}{q_\parallel^2 + q_\perp^2}\frac{e^{i\mathbf{q}\mathbf{r} - q^2L^2/4}}{iq_\parallel + 1/l_{\text{in}}} \\ &= \int\limits_0^\infty dse^{-s/l_{\text{in}}}\int\limits_0^\infty dt(\partial_\perp^2 - \partial_\parallel^2)\int\frac{d^D\mathbf{q}}{(2\pi)^D}e^{-q^2(t+L^2/4)+i\mathbf{q}\mathbf{r}-iq_\parallel s} \\ &= \int\limits_0^\infty dse^{-s/l_{\text{in}}}\tilde{Q}_D(x-s, \mathbf{r}_\perp), \end{aligned} \tag{A1}$$

where

$$\tilde{Q}_D(\mathbf{r}) = \frac{1}{\pi^{D/2}}(\partial_\perp^2 - \partial_\parallel^2)\int\limits_0^\infty dt\frac{e^{-r^2/(4t+L^2)}}{4t + L^2}. \tag{A2}$$

Applying the spatial derivatives $(\partial_\perp^2 - \partial_\parallel^2)$, we get

$$\begin{aligned} \tilde{Q}_D(\mathbf{r}) = &\int\limits_{L^2/4}^\infty dte^{-r^2/4t} \\ &\times\begin{cases} \dfrac{2t - x_\parallel^2}{8\sqrt{\pi}t^{5/2}}, & \text{for} \quad D = 1, \\ \dfrac{r_\perp^2 - x_\parallel^2}{16\pi t^3}, & \text{for} \quad D = 2, \\ \dfrac{r_\perp^2 - x_\parallel^2 - 2t}{32\pi^{3/2}t^{7/2}}, & \text{for} \quad D = 3. \end{cases} \end{aligned} \tag{A3}$$

Calculating the remaining integral over t and assuming that electric field is parallel to the x-axis, we arrive at Eq. (43) of the main text.

Appendix B: Temperature Landauer dipole in 3D and 2D system—a local heating approximation

In this Appendix, we employ the local heating approximation (Fig. 8), and study the constriction-induced local heating of 2D and 3D systems. The calculation here generalizes the one in Ref. [47].

1. Landauer dipole in 3D

Within the local heating approximation, the constriction of radius L has the uniform conductivity σ_{in} that is smaller

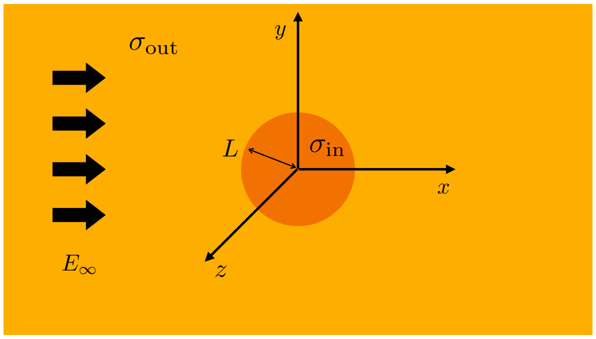


FIG. 8. Phenomenological local-conductivity model adopted for the local Joule approximation. A constriction of radius L, located at the origin, is characterized by the conductivity σ_{in} that is smaller than the uniform background conductivity: $\sigma_{\text{in}} < \sigma_{\text{out}}$. The external field E_∞ is directed along the x-axis.

than the conductivity σ_{out} outside the constricted area. The electrical current is driven by the field E_∞ directed along the x-axis. Under the continuity boundary condition, the electric potential becomes

$$\varphi(\mathbf{r}) = \begin{cases} -\dfrac{3\sigma_{\text{out}}}{\sigma_{\text{in}} + 2\sigma_{\text{out}}} rE_\infty \cos\vartheta, & r < L, \\[2ex] -\left[r + \dfrac{(\sigma_{\text{out}} - \sigma_{\text{in}})L^3}{(2\sigma_{\text{out}} + \sigma_{\text{in}})r^2}\right] E_\infty \cos\vartheta, & r > L, \end{cases} \tag{B1}$$

where ϑ is the angle between the x axis and the direction of the vector $\mathbf{r}$ in real space. Apparently, $\varphi(\mathbf{r})$ reduces to the uniform value $-E_\infty r\cos\vartheta$ in the $r \equiv |\mathbf{r}| \gg L$ limit.

With the electric potential from Eq. (B1), we find the local electric field and calculate the local Joule heating $\sigma(\mathbf{r})E(\mathbf{r})^2$

$$\begin{aligned} \frac{P(\mathbf{r})}{\sigma_{\text{out}}E_\infty^2} &= \frac{\sigma(\mathbf{r})E(\mathbf{r})^2}{\sigma_{\text{out}}E_\infty^2} \\ &= \begin{cases} (1+\lambda-2\lambda^2)\sigma_{\text{out}}E_\infty^2, & r < L, \\[2ex] \left[1 - \dfrac{L^3}{r^3}\lambda(1-3\cos\vartheta)\right]^2 + \dfrac{9L^6}{4r^6}\lambda^2\sin^2(2\vartheta), & r > L, \end{cases} \end{aligned} \tag{B2}$$

where

$$\lambda \equiv \frac{\sigma_{\text{in}} - \sigma_{\text{out}}}{2\sigma_{\text{out}} + \sigma_{\text{in}}} \tag{B3}$$

is the parameter characterizing the inhomogeneity. It vanishes in the homogeneous situation ($\sigma_{\text{in}} = \sigma_{\text{out}}$), and approaches $1/2$ when the constriction is insulating ($\sigma_{\text{in}} \to 0$).

To describe the effect of the constriction, we define

$$\delta P = P - \sigma_{\text{out}}E_\infty^2,$$

as the constriction-induced extra Joule heating and express it in terms of λ:

$$\begin{aligned} &\frac{\delta P}{\sigma_{\text{out}}E_\infty^2} \\ &= \begin{cases} \lambda - 2\lambda^2, & r < L, \\[2ex] \dfrac{\lambda L^3[1+3\cos(2\vartheta)]}{r^3} + \dfrac{\lambda^2 L^6[5+3\cos(2\vartheta)]}{2r^6}, & r > L. \end{cases} \end{aligned} \tag{B4}$$

Clearly, δP vanishes when $\lambda = 0$. It is also non-monotonous and approaches the peak value when $\lambda = 1/4$. Indeed, all particles detour around the constriction in the insulating limit, where the Joule heating vanishes.

For a more intuitive understanding, we provide the profile of the local Joule heating (when $\lambda = 1/3$) in Fig. 1(d) of the main text. Apparently, a Landauer-dipole heating pattern emerges around the constriction. This heating pattern naturally induces a shifted Landauer dipole of the electron temperature profile. Noteworthy, this Landauer-dipole feature is produced by the residual charges at the boundary of the constriction. These charges also guarantee the constant electric field in the constricted area.

2. Landauer dipole in 2D

Let us move to the 2D case with a circular-shape constriction. To begin with, following the same method as that of Appendix B 1, we arrive at the electric field of 2D

$$\mathbf{E}(\mathbf{r}) = \begin{cases} \mathbf{E}_\infty + \delta\mathbf{E}_{\text{in}}, & r < L, \\[1ex] \mathbf{E}_\infty + \delta\mathbf{E}_{\text{out}}, & r > L, \end{cases} \tag{B5}$$

where

$$(\delta\mathbf{E}_{\text{in}})_x = \delta E, \quad (\delta\mathbf{E}_{\text{in}})_y = 0 \tag{B6}$$

is the electric field in the constricted area, and

$$\begin{aligned} (\delta\mathbf{E}_{\text{out}})_x &= \delta E \frac{y^2 - x^2}{(x^2+y^2)^2} L^2, \\ (\delta\mathbf{E}_{\text{out}})_y &= \delta E \frac{-2xy}{(x^2+y^2)^2} L^2, \end{aligned} \tag{B7}$$

refers to that outside of the constriction. In Eqs. (B6) and (B7), the inhomogeneity-induced electric field is given by

$$\delta E = E_\infty \frac{\sigma_{\text{out}} - \sigma_{\text{in}}}{\sigma_{\text{out}} + \sigma_{\text{in}}} = E_\infty \lambda_2 > 0, \tag{B8}$$

where the parameter λ_2 characterizes the strength of the inhomogeneity in the 2D case.

In the local approximation, the drift velocity is proportional to the electric field $\mathbf{v}(\mathbf{r}) = \sigma(\mathbf{r})\mathbf{E}(\mathbf{r})/N_\infty e$. Following Eq. (B5), we obtain the drift velocity profile

$$\mathbf{v}(\mathbf{r}) = \begin{cases} \mathbf{v}_\infty + \delta\mathbf{v}_{\text{in}}, & r < L, \\[1ex] \mathbf{v}_\infty + \delta\mathbf{v}_{\text{out}}, & r > L, \end{cases} \tag{B9}$$

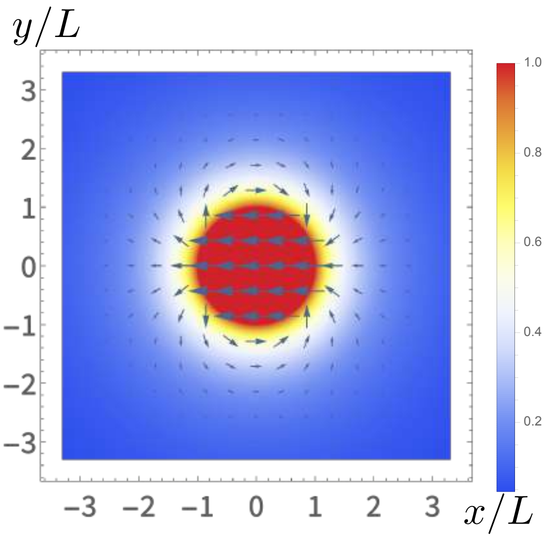


FIG. 9. The vector velocity profile of the 2D local heating model.

where

$$(\delta\mathbf{v}_{\text{in}})_x = -\delta v, \quad (\delta\mathbf{v}_{\text{in}})_y = 0 \tag{B10}$$

in the constricted area, and

$$\begin{aligned}(\delta\mathbf{v}_{\text{out}})_x &= \delta v \frac{y^2 - x^2}{(x^2+y^2)^2} L^2 \\ (\delta\mathbf{v}_{\text{out}})_y &= \delta v \frac{-2xy}{(x^2+y^2)^2} L^2\end{aligned} \tag{B11}$$

outside of the constriction. In Eqs. (B11) and (B10), the homogeneous velocity $\mathbf{v}_\infty = \sigma_{\text{out}}\mathbf{E}_\infty/Ne$, and

$$\delta v = v_\infty \frac{\sigma_{\text{out}} - \sigma_{\text{in}}}{\sigma_{\text{out}} + \sigma_{\text{in}}} \equiv v_\infty \lambda_2 \tag{B12}$$

is the inhomogeneity-induced velocity variation, defined in the same manner as Eq. (B8). In the momentum space, the variation of velocity becomes

$$\begin{aligned}\delta v_x(\mathbf{q}) &= \frac{2\pi L}{q}[J_1(Lq)\cos(2\varphi) - J_1(Lq)]\lambda_2 \\ \delta v_y(\mathbf{q}) &= \frac{2\pi L}{q} J_1(Lq)\sin(2\varphi)\lambda_2,\end{aligned} \tag{B13}$$

where $q = |\mathbf{q}|$, the angle φ refers to the momentum direction, and $J_1(Lq)$ is the Bessel function.

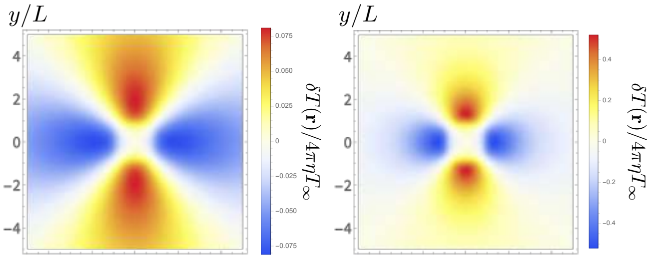
To better understand Eq. (B13), we Fourier transform the circular-shape step function, $\xi(\mathbf{r}) = \xi_0\Theta(x^2+y^2-L^2)$, into the expression in the momentum space

$$\xi(\mathbf{q}) = \xi_0 \frac{2\pi}{q} L J_1(Lq). \tag{B14}$$

The combination of Eqs. (B13) and (B14) gives us the velocity kernel of the local-heating model

$$\begin{aligned}K_{u_x} &= \frac{-2q_\perp^2}{q_\parallel^2 + q_\perp^2}, \\ K_{u_y} &= \frac{2q_\parallel q_\perp}{q_\parallel^2 + q_\perp^2},\end{aligned} \tag{B15}$$

where $q_\parallel$ and $q_\perp$ refer to two momentum components. Equation (B15) coincides with Eq. (31) of the main text, except for an overall extra factor of two. This extra factor is related to the definition of λ_2: at the weak inhomogeneity limit, $\lambda_2 \approx (\sigma_{\text{out}} - \sigma_{\text{in}})/2\sigma_{\text{out}}$. Following Eq. (B12), we plot the 2D velocity profile in Fig. 9. It captures the major feature of that in the main text: particles tend to detour around the constriction.



(a) (b)

FIG. 10. The temperature profiles of the phenomenological model when (a) $l_* = 2L$; (b) $l_* = 0.5L$. The temperature profiles of the diffusive phenomenological model are completely symmetric.

With the velocity profile known, we can calculate the temperature profile. Since this local-heating model is diffusive, $\varkappa_e \gg v_\infty$ (where $\varkappa_e$ is the electronic heat conductivity), we keep only the diffusive term in the heat diffusion equation

$$-\varkappa_e \bigtriangledown^2 \delta T = ne\mathbf{E}\mathbf{v} - \gamma\delta T. \tag{B16}$$

As a simple check, we focus on the weak-inhomogeneity limit $\lambda_2 \ll 1$. In this limit, the Fourier-transformed temperature profile becomes

$$\delta T_{\mathbf{q}} = T_\infty \frac{4\pi}{q} \frac{LJ_1(Lq)\cos(2\varphi)}{l_*^2 q^2 + 1}\lambda_2, \tag{B17}$$

where $l_* = \sqrt{\varkappa_e/\gamma}$. In Eq. (B17), the prefactor T_∞ is the homogeneous temperature when $\lambda_2 = 0$. We present the temperature profile of Eq. (B17) in the real space, in Fig. 10. In contrast to the hydrodynamic model of the main text Eq. (32), the temperature kernel of Eq. (B17) strictly corresponds to a Landauer dipole that sits exactly at the center of the constricted area: the asymmetry thus completely disappears in the diffusive situation.

---

[1] T. Ihn, *Semiconductor Nanostructures* (Oxford University Press, Oxford, 2010).

[2] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature **438**, 197 (2005).

[3] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, Nature **391**, 156 (1998).

[4] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven, Science **281**, 540 (1998).

[5] A. Finkler, Y. Segev, Y. Myasoedov, M. L. Rappaport, L. Ne'eman, D. Vasyukov, E. Zeldov, M. E. Huber, J. Martin, and A. Yacoby, Nano Letters **10**, 1046 (2010).

[6] D. Vasyukov, Y. Anahory, L. Embon, D. Halbertal, J. Cuppens, L. Neeman, A. Finkler, Y. Segev, Y. Myasoedov, M. L. Rappaport, M. E. Huber, and E. Zeldov, Nature Nanotechnology **8**, 639 EP (2013).

[7] U. Vool, A. Hamo, G. Varnavides, Y. Wang, T. X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C. A. C. Garcia, A. T. Pierce, J. Gooth, P. Anikeeva, C. Felser, P. Narang, and A. Yacoby, "Imaging phonon-mediated hydrodynamic flow in WTe_2 with cryogenic quantum magnetometry," (2020), arXiv:2009.04477 [cond-mat.mes-hall].

[8] D. Halbertal, J. Cuppens, M. B. Shalom, L. Embon, N. Shadmi, Y. Anahory, H. R. Naren, J. Sarkar, A. Uri, Y. Ronen, Y. Myasoedov, L. S. Levitov, E. Joselevich, A. K. Geim, and E. Zeldov, Nature **539**, 407 (2016).

[9] D. Halbertal, M. Ben Shalom, A. Uri, K. Bagani, A. Y. Meltzer, I. Marcus, Y. Myasoedov, J. Birkbeck, L. S. Levitov, A. K. Geim, and E. Zeldov, Science **358**, 1303 (2017).

[10] A. Marguerite, J. Birkbeck, A. Aharon-Steinberg, D. Halbertal, K. Bagani, I. Marcus, Y. Myasoedov, A. K. Geim, D. J. Perello, and E. Zeldov, Nature **575**, 628 (2019).

[11] A. Uri, Y. Kim, K. Bagani, C. K. Lewandowski, S. Grover, N. Auerbach, E. O. Lachman, Y. Myasoedov, T. Taniguchi, K. Watanabe, J. Smet, and E. Zeldov, Nature Physics **16**, 164 (2020).

[12] A. Aharon-Steinberg, A. Marguerite, D. J. Perello, K. Bagani, T. Holder, Y. Myasoedov, L. S. Levitov, A. K. Geim, and E. Zeldov, "Long-range nontopological edge currents in charge-neutral graphene," (2020), arXiv:2012.02842 [cond-mat.mes-hall].

[13] R. N. Gurzhi, Zh. Eksp. Teor. Fiz. **44**, 771 (1963), [Sov. Phys. JETP **17**, 521 (1963)].

[14] R. N. Gurzhi, Zh. Eksp. Teor. Fiz. **47**, 1415 (1964), [Sov. Phys. JETP **20**, 953 (1965)].

[15] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968).

[16] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[17] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[18] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).

[19] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[20] M. Polini and A. K. Geim, Physics Today **73**, 28 (2020).

[21] D. A. Bandurin, I. Torre, R. K. Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, L. A. Ponomarenko, A. K. Geim, and M. Polini, Science **351**, 1055 (2016).

[22] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, Science **351**, 1058 (2016).

[23] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[24] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[25] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, and A. K. Geim, Nature Physics **13**, 1182 (2017).

[26] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nature Communications **9**, 4533 (2018).

[27] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, M. Polini, T. Ihn, and K. Ensslin, Phys. Rev. B **98**, 241304 (2018).

[28] A. Jaoui, B. Fauqué, C. W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D. L. Maslov, C. Felser, and K. Behnia, npj Quantum Materials **3**, 64 (2018).

[29] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[30] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin, Science **364**, 162 (2019).

[31] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[32] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A. K. Geim, and S. Ilani, Nature **576**, 75 (2019).

[33] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, and J. A. Sulpizio, Nature Nanotechnology **14**, 480 (2019).

[34] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M. M. Fogler, P. Kim, A. Yacoby, and R. L. Walsworth, Nature **583**, 537 (2020).

[35] O. E. Raichev, G. M. Gusev, A. D. Levin, and A. K. Bakarov, Phys. Rev. B **101**, 235314 (2020).

[36] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Scientific Reports **10**, 7860 (2020).

[37] J. Geurs, Y. Kim, K. Watanabe, T. Taniguchi, P. Moon, and J. H. Smet, "Rectification by hydrodynamic flow in an encapsulated graphene Tesla valve," (2020), arXiv:2008.04862 [cond-mat.mes-hall].

[38] M. Kim, S. G. Xu, A. I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, R. V. Gorbachev, K. Watanabe, T. Taniguchi, I. V. Grigorieva, V. I. Fal'ko, M. Polini, and A. K. Geim, Nature Communications **11**, 2339 (2020).

[39] A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021).

[40] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Phys. Rev. B **103**, 075303 (2021).

[41] Y. Zhang and M. S. Shur, Journal of Applied Physics **129**, 053102 (2021).

[42] A. Jaoui, B. Fauqué, and K. Behnia, Nature Communications **12**, 195 (2021).
[43] M. Rokni and Y. Levinson, Phys. Rev. B **52**, 1882 (1995).
[44] V. L. Gurevich, Phys. Rev. B **55**, 4522 (1997).
[45] R. G. Gatiyatov, V. N. Lisin, and A. A. Bukharaev, JETP Letters **91**, 425 (2010).
[46] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **100**, 205430 (2019).
[47] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **97**, 085415 (2018).
[48] J. F. Kong, L. Levitov, D. Halbertal, and E. Zeldov, Phys. Rev. B **97**, 245416 (2018).
[49] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).
[50] J. C. W. Song and L. S. Levitov, Journal of Physics: Condensed Matter **27**, 164201 (2015).
[51] R. Landauer, IBM Journal of Research and Development **1**, 223 (1957).
[52] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1959).
[53] Z. Chen, W. Jang, W. Bao, C. N. Lau, and C. Dames, Applied Physics Letters **95**, 161910 (2009).
[54] B. N. J. Persson and H. Ueba, EPL (Europhysics Letters) **91**, 56001 (2010).
[55] X. Li, Y. Yan, L. Dong, J. Guo, A. Aiyiti, X. Xu, and B. Li, Journal of Physics D: Applied Physics **50**, 104002 (2017).

Hydrodynamic approach to electronic transport in graphene: energy relaxation

B.N. Narozhny[1, 2] and I.V. Gornyi[3, 4]
[1]*Institut for Theoretical Condensed Matter Physics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[2]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[3]*Institut for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[4]*Ioffe Institute, 194021 St. Petersburg, Russia*
(Dated: April 30, 2021)

In nearly compensated graphene, disorder-assisted electron-phonon scattering or "supercollisions" are responsible for both quasiparticle recombination and energy relaxation. Within the hydrodynamic approach, these processes contribute weak decay terms to the continuity equations at local equilibrium, i.e., at the level of "ideal" hydrodynamics. Here we report the derivation of the decay term due to weak violation of energy conservation. Such terms have to be considered on equal footing with the well-known recombination terms due to nonconservation of the number of particles in each band. At high enough temperatures in the "hydrodynamic regime" supercollisions dominate both types of the decay terms (as compared to the leading-order electron-phonon interaction). We also discuss the contribution of supercollisions to the heat transfer equation (generalizing the continuity equation for the energy density in viscous hydrodynamics).

Electronic hydrodynamics is quickly growing into a mature field of condensed matter physics [1–3]. Similarly to the usual hydrodynamics [4, 5], this approach offers a universal, long-wavelength description of collective flows in interacting many-electron systems. As a macroscopic theory of strongly interacting systems, hydrodynamics should appear to be extremely attractive for condensed matter theorists dealing with problems where strong correlations invalidate simple theoretical approaches. However, electrons in solids exist in the environment created by a crystal lattice and typically experience collisions with lattice imperfections (or "disorder") and lattice vibrations (phonons). The former typically dominate electronic transport at low temperatures, while at high temperatures the electron-phonon interaction takes over. In both cases the electron motion is diffusive (unless the sample size is smaller than the mean free path in which case the motion is ballistic) since in both types of scattering the electronic momentum is not conserved. On the other hand, if a material would exist where the momentum-conserving electron-electron interaction would dominate at least in some non-negligible temperature range, then one could be justified in neglecting the momentum non-conserving processes and applying the hydrodynamic theory. In recent years, several extremely pure materials became available with graphene being the most studied [1, 3].

As a manifestation of macroscopic conservation laws, hydrodynamics is universal. Most conventional fluids are assumed Galilean invariant and are described by the same set of hydrodynamic equations [4]. Similar approach was at the heart of the early theoretical work on electronic hydrodynamics [6, 7]. Another well-known case is the relativistic hydrodynamics [4] relevant to neutron stars and interstellar matter. Since its low-energy quasiparticles are characterized by the Dirac spectrum, graphene attracted significant theoretical attention as a possible condensed matter realization of relativistic hydrodynamics [8–13]. However, due to the classical, three-dimensional nature of the Coulomb interaction between electrons, the emergent hydrodynamics in graphene is neither Galilean- nor Lorentz-invariant [2].

In nearly neutral (or compensated) graphene the electron system is non-degenerate (at least at relatively high temperatures where the hydrodynamic approach is justified) with both the conductance and valence bands contributing on equal footing. Although the electron system is not Lorenz-invariant, the linearity of the Dirac spectrum plays an important role. Firstly, the Auger processes are kinematically suppressed leading to the near-conservation of the number of particles in each band [2, 3, 14, 15]. Secondly, the so-called collinear scattering singularity [10–12, 15–19] allows for a non-perturbative solution to the kinetic (Boltzmann) equation focusing on the three hydrodynamic modes [18, 20, 21]. As a result, one can determine the general form of the hydrodynamic equations and to evaluate the kinetic coefficients [21–23]. To be of any practical value, the latter calculation has to be combined with the renormalization group approach [24] since the effective coupling constant in real graphene (either encapsulated or put on a dielectric substrate) is not too small, $\alpha_g \approx 0.2 - 0.3$ [25, 26].

Next to the conservation laws, the main assumption of the hydrodynamic approach is local equilibrium [4, 27] established by means of interparticle collisions. Neglecting all dissipative processes, this allows (together with the conservation laws) for a phenomenological derivation of hydrodynamic equations [4, 5] that can be further supported by the kinetic theory, where the local equilibrium distribution function nullifies the collision integral in the Boltzmann equation [27]. The resulting ideal hydrodynamics is described by the Euler equation and the continuity equations. This is where the electronic fluid in graphene differs from conventional fluids (both Galilean- and Lorentz-invariant): as in any solid, conservation laws in graphene are only approximate, leaving the collision integrals describing scattering processes other than the electron-electron interaction to be nonzero even in local equilibrium. This leads to the appearance of weak decay terms in the continuity equations.

Two such terms have already been discussed in literature. Firstly, even if the electron-electron interaction is the dominant scattering process in the system, no solid is absolutely pure. Consequently, even ultra-pure graphene samples possess some degree of weak disorder. Disorder scattering violates momentum conservation and hence a weak decay term must appear in the generalized Euler equation [2, 3, 20, 21]. Secondly, conservation of the number of particles in each band is violated by a number of processes (e.g., the Auger and three-particle scattering). While commonly assumed to be weak, they are manifested in the decay – or recombination – term in the corresponding continuity equation. This was first established in [14] in the context of thermoelectric phenomena (for the most recent discussion see [28]). Later, quasiparticle recombination was shown to lead to linear magnetoresistance in compensated semimetals [29–32] as well as giant magnetodrag [33, 34].

In this paper, we report the derivation of the third weak decay term in the hydrodynamic theory in graphene due to weak violation of energy conservation. Indeed, the electron-phonon interaction may lead not only to the loss of electronic momentum (responsible for electrical resistivity in most metals at high temperatures), but also to the loss of energy. Although subdominant in the hydrodynamic regime, the electron-phonon interaction should be taken into account as one of the dissipative processes. In graphene, the linearity of the Dirac spectrum once again plays an important role: as the speed of sound is much smaller than the electron velocity v_g, leading-order scattering on acoustic phonons is kinematically suppressed. Consequently, scattering off the optical branch is usually considered [35, 36]. In contrast, we argue that there is another process, the disorder-assisted electron-phonon scattering [37] or "supercollisions" [38–41], that is responsible for both quasiparticle recombination and energy relaxation. In the high-temperature hydrodynamic regime, the supercollisions are expected to dominate both decay contributions [37]. Moreover, this process contributes weak decay terms to the continuity equations already at local equilibrium, i.e., at the level of "ideal" hydrodynamics.

Our arguments are based on the kinetic theory approach to electronic transport. In the spirit of Ref. 24, we assume the possibility of deriving the hydrodynamic equations from the kinetic equation in the weak coupling limit [21], $\alpha_g \ll 1$, with the subsequent renormalization of the kinetic coefficients to the realistic parameter regime [22]. Under these assumptions, we start with the kinetic equation

$$\mathcal{L} f_{\lambda\boldsymbol{k}} = \mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}] + \mathrm{St}_{R}[f_{\lambda\boldsymbol{k}}] + \mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}], \tag{1a}$$

with the Liouville's operator (in the left-hand side)

$$\mathcal{L} = \partial_t + \boldsymbol{v}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}} + \left(e\boldsymbol{E} + \frac{e}{c}\boldsymbol{v}\times\boldsymbol{B}\right)\cdot\boldsymbol{\nabla}_{\boldsymbol{k}}, \tag{1b}$$

and the collision integrals describing the electron-electron interaction (St_{ee}), disorder scattering ($\mathrm{St}_{\mathrm{dis}}$), and quasiparticle recombination (St_{R}), where in this paper we focus on "supercollisions". We employ the following notations for the Dirac spectrum (the chirality $\lambda = \pm 1$ distinguishes the conduction and valence bands)

$$\epsilon_{\lambda\boldsymbol{k}} = \lambda v_g k, \tag{2a}$$

and velocities

$$\boldsymbol{v}_{\lambda\boldsymbol{k}} = \lambda v_g \frac{\boldsymbol{k}}{k}, \qquad \boldsymbol{k} = \frac{\lambda k}{v_g}\boldsymbol{v}_{\lambda\boldsymbol{k}} = \frac{\epsilon_{\lambda\boldsymbol{k}}\boldsymbol{v}_{\lambda\boldsymbol{k}}}{v_g^2}. \tag{2b}$$

Hydrodynamics is the macroscopic manifestation of the conservation of energy, momentum, and the number of particles. In a two-band system, the latter comprises excitations in both bands. In the conductance band these are electron-like quasiparticles with the number density ($N = 4$ reflects spin and valley degeneracy in graphene)

$$n_+ = N\int\frac{d^2k}{(2\pi)^2} f_{+,\boldsymbol{k}}, \tag{3a}$$

while in the valence band the quasiparticles are hole-like

$$n_- = N\int\frac{d^2k}{(2\pi)^2}\left(1 - f_{-,\boldsymbol{k}}\right), \tag{3b}$$

with the total "charge" (or "carrier") density being

$$n = n_+ - n_-. \tag{3c}$$

Assuming the numbers of particles in the conduction and valence bands are conserved independently, we can also define the total quasiparticle ("imbalance" [14]) density

$$n_I = n_+ + n_-. \tag{3d}$$

Global charge conservation (or gauge symmetry) can be expressed in terms of the usual continuity equation. This can be obtained from Eq. (1) by performing a summation over all quasiparticle states upon which all three collision integrals vanish [27]

$$N\int\frac{d^2k}{(2\pi)^2}\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}] = N\int\frac{d^2k}{(2\pi)^2}\mathrm{St}_{R}[f_{\lambda\boldsymbol{k}}] = \tag{4a}$$

$$= N\int\frac{d^2k}{(2\pi)^2}\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}] = 0.$$

As a result, the continuity equation has the usual form

$$\partial_t n + \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{j} = 0, \tag{4b}$$

where the corresponding current is defined as

$$\boldsymbol{j} = \boldsymbol{j}_+ - \boldsymbol{j}_- = N\int\frac{d^2k}{(2\pi)^2}\left[\boldsymbol{v}_{+,\boldsymbol{k}} f_{+,\boldsymbol{k}} - \boldsymbol{v}_{-,\boldsymbol{k}}\left(1 - f_{-,\boldsymbol{k}}\right)\right], \tag{5}$$

The rest of the conservation laws in graphene are approximate as manifested in the collision integrals not vanishing upon corresponding summations. The continuity equation expressing momentum conservation (i.e. the Euler equation) is obtained by multiplying the kinetic equation by the quasiparticle momentum $\boldsymbol{k}$ and summing over all states. Since the electron-electron interaction conserves momentum, the corresponding collision integral vanishes

$$N\int\frac{d^2k}{(2\pi)^2}\boldsymbol{k}\,\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=0. \tag{6}$$

Weak disorder scattering is typically described within the simplest τ-approximation [27]

$$N\int\frac{d^2k}{(2\pi)^2}\boldsymbol{k}\,\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=\frac{\boldsymbol{n}_{\boldsymbol{k}}}{\tau_{\mathrm{dis}}}, \tag{7}$$

where the momentum density is defined as

$$\boldsymbol{n}_{\boldsymbol{k}}=N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\boldsymbol{k}f_{\lambda\boldsymbol{k}}=v_g^{-2}\boldsymbol{j}_E. \tag{8}$$

The last equality represents the fact that in graphene the momentum density is proportional to the energy density [due to the properties of the Dirac spectrum Eq. (2)].

Supercollisions contributing to the recombination collision integral also violate momentum conservation, however, in comparison to the above weak disorder scattering, this is a second-order process. Moreover, the disorder mean free time τ_{dis} is typically determined from experimental data (see e.g., Ref. 26) and hence can be assumed to include the contribution of supercollisions as well.

The remaining two continuity equations – energy and quasiparticle imbalance – are unaffected by the electron-electron interaction and weak disorder scattering. Indeed, the electron-electron interaction conserves energy and – neglecting the Auger processes – particle number in each band:

$$N\int\frac{d^2k}{(2\pi)^2}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=N\int\frac{d^2k}{(2\pi)^2}\lambda\,\mathrm{St}_{ee}[f_{\lambda\boldsymbol{k}}]=0. \tag{9}$$

Same applies to the (elastic) disorder scattering

$$N\int\frac{d^2k}{(2\pi)^2}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=N\int\frac{d^2k}{(2\pi)^2}\lambda\,\mathrm{St}_{\mathrm{dis}}[f_{\lambda\boldsymbol{k}}]=0. \tag{10}$$

However, supercollisions do not conserve both quantities and hence lead to weak decay terms in the two continuity equations.

Let us now specify the collision integral describing the disorder-assisted electron-phonon scattering. An electron in the upper (conductance) band may scatter into an empty state in the lower (valence) band – effectively recombining with a hole – emitting a phonon (which carries away the energy) and losing its momentum to the impurity. Within the standard approach to the electron-phonon interaction, this process is described by the collision integral

$$\mathrm{St}_R[f_{+\boldsymbol{k}_2}]=2\pi\sum_{\boldsymbol{k}_1,\boldsymbol{q}}W_q\delta(\epsilon_{+\boldsymbol{k}_2}-\epsilon_{-\boldsymbol{k}_1}-\omega_q)\times \tag{11a}$$

$$\times\left[f_{-\boldsymbol{k}_1}(1-f_{+\boldsymbol{k}_2})n_q-f_{+\boldsymbol{k}_2}(1-f_{-\boldsymbol{k}_1})(1+n_q)\right],$$

where n_q is the phonon (Plank's) distribution function (the phonons are assumed to be at equilibrium and play the role of a "bath"), W_q is the effective scattering rate that includes the Dirac factors and is averaged over the angles [37].

Similarly, an electron in the lower band may absorb a phonon and scatter into the upper band – effectively creating an electron-hole pair – while still losing its momentum to the impurity

$$\mathrm{St}_R[f_{-\boldsymbol{k}_2}]=2\pi\sum_{\boldsymbol{k}_1,\boldsymbol{q}}W_q\delta(\epsilon_{+\boldsymbol{k}_1}-\epsilon_{-\boldsymbol{k}_2}-\omega_q)\times \tag{11b}$$

$$\times\left[f_{+\boldsymbol{k}_1}(1-f_{-\boldsymbol{k}_2})(1+n_q)-f_{-\boldsymbol{k}_2}(1-f_{+\boldsymbol{k}_1})n_q\right].$$

The collision integral (11) conserves the total charge

$$N\sum_{\boldsymbol{k}}\mathrm{St}_R[f_{\lambda\boldsymbol{k}}]=N\sum_{\boldsymbol{k}_2}\left(\mathrm{St}_R[f_{+\boldsymbol{k}_2}]+\mathrm{St}_R[f_{-\boldsymbol{k}_2}]\right)=0, \tag{11c}$$

[see Eq. (4)] and vanishes in global equilibrium

$$\mathrm{St}_R[f^{(0)}]=0, \tag{11d}$$

where the quasiparticle distribution is described by the Fermi function. This should be contrasted with *local equilibrium* described by

$$f^{(le)}_{\lambda\boldsymbol{k}}=\left\{1+\exp\left[\frac{\epsilon_{\lambda\boldsymbol{k}}-\mu_\lambda(\boldsymbol{r})-\boldsymbol{u}(\boldsymbol{r})\cdot\boldsymbol{k}}{T(\boldsymbol{r})}\right]\right\}^{-1}, \tag{12}$$

where $\mu_\lambda(\boldsymbol{r})$ is the local chemical potential and $\boldsymbol{u}(\boldsymbol{r})$ is the hydrodynamic (or "drift") velocity. The local equilibrium distribution function (12) allows for independent chemical potentials in the two bands, which can be expressed in terms of the "thermodynamic" and "imbalance" chemical potentials

$$\mu_\lambda=\mu+\lambda\mu_I. \tag{13}$$

In global equilibrium

$$f^{(0)}=f^{(le)}_{\lambda\boldsymbol{k}}(\mu_I=0,\boldsymbol{u}=0). \tag{14}$$

Now we show, that in local equilibrium, i.e. for nonzero μ_I, the recombination collision integral remains finite [unlike Eq. (11d)]. As a scalar quantity, the collision integral (11) cannot depend on the hydrodynamic velocity $\boldsymbol{u}$ in the first (linear) order. Consequently, to the leading order the integrated collision integral yielding the decay terms in the continuity equations is proportional to μ_I.

To the leading order, we can describe the difference between the local equilibrium distribution function $f^{(le)}_{\lambda\boldsymbol{k}}$ and the Fermi function $f^{(0)}$ similarly to the leading nonequilibrium correction in the standard derivation of hydrodynamics [27]

$$\delta f=f_{\lambda\boldsymbol{k}}-f^{(0)}=-T\frac{\partial f^{(0)}}{\partial\epsilon}h=f^{(0)}\left(1-f^{(0)}\right)h. \tag{15}$$

Now we re-write the collision integral (11) with the help of the relations

$$f_1(1-f_2)(1+n_q) - f_2(1-f_1)n_q =$$

$$= (1-f_1)(1-f_2)(1+n_q)\left[\frac{f_1}{1-f_1}\frac{n_q}{1+n_q} - \frac{f_2}{1-f_2}\right],$$

and

$$f_2^{(0)}\left(1-f_1^{(0)}\right)(1+n_q) = -\frac{\partial n_q}{\partial\omega}\left(f_2^{(0)}-f_1^{(0)}\right),$$

and find (to the leading order in h)

$$\mathrm{St}_R[f_{+\boldsymbol{k}_2}] = -2\pi\sum_{\boldsymbol{k}_1,\boldsymbol{q}} W_q\delta(\epsilon_{+\boldsymbol{k}_2}-\epsilon_{-\boldsymbol{k}_1}-\omega_q)\frac{\partial n_q}{\partial\omega}$$

$$\times\left(f^{(0)}_{+\boldsymbol{k}_2}-f^{(0)}_{-\boldsymbol{k}_1}\right)(h_{-\boldsymbol{k}_1}-h_{+\boldsymbol{k}_2}), \tag{16a}$$

$$\mathrm{St}_R[f_{-\boldsymbol{k}_2}] = 2\pi\sum_{\boldsymbol{k}_1,\boldsymbol{q}} W_q\delta(\epsilon_{-\boldsymbol{k}_2}-\epsilon_{+\boldsymbol{k}_1}+\omega_q)\frac{\partial n_q}{\partial\omega}$$

$$\times\left(f^{(0)}_{+\boldsymbol{k}_1}-f^{(0)}_{-\boldsymbol{k}_2}\right)(h_{-\boldsymbol{k}_2}-h_{+\boldsymbol{k}_1}). \tag{16b}$$

Consider now the contribution of the recombination collision integral to the continuity equation for the quasiparticle imbalance

$$N\sum_{\boldsymbol{k}}\lambda\,\mathrm{St}_R[f_{\lambda\boldsymbol{k}}] = N\sum_{\boldsymbol{k}_2}\left(\mathrm{St}_R[f_{+\boldsymbol{k}_2}]-\mathrm{St}_R[f_{-\boldsymbol{k}_2}]\right) =$$

$$= -4\pi N\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q}} W_q\delta(\epsilon_{+\boldsymbol{k}_2}-\epsilon_{-\boldsymbol{k}_1}-\omega_q)\frac{\partial n_q}{\partial\omega}$$

$$\times\left(f^{(0)}_{+\boldsymbol{k}_2}-f^{(0)}_{-\boldsymbol{k}_1}\right)(h_{-\boldsymbol{k}_1}-h_{+\boldsymbol{k}_2}). \tag{17}$$

To the leading order, the deviation $h_{\lambda\boldsymbol{k}}$ is proportional to μ_I

$$h_{\lambda\boldsymbol{k}} \approx \frac{\lambda\mu_I}{T}. \tag{18}$$

The remaining integral has dimensions of particle density divided by time and therefore the result can be written in two equivalent forms

$$N\sum_{\boldsymbol{k}}\lambda\,\mathrm{St}_R[f_{\lambda\boldsymbol{k}}] \approx -\mu_I n_{I,0}\lambda_Q \approx -\frac{n_I-n_{I,0}}{\tau_R}. \tag{19}$$

Here n_I is the imbalance density (3d) in local equilibrium, while $n_{I,0}$ is the same quantity evaluated with the Fermi distribution function (14), i.e. for $\mu_I = 0$ and $\boldsymbol{u} = 0$. The first equality in Eq. (19) coincides with the expression used in Ref. 14 and serves as the definition of the dimensionless coefficient λ_Q, while the second (valid to the leading order) was suggested in Refs. 21 and 29 and provides the definition of the "recombination time" τ_R (see also Ref. 28). The two expressions are equivalent since $n_I - n_{I,0} \propto \mu_I$.

The same scattering process contributes a weak decay term to the continuity equation for the energy density. Indeed, multiplying the collision integral (11) by the quasiparticle energy and summing over all states, we find after similar algebra

$$N\sum_{\boldsymbol{k}}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_R[f_{\lambda\boldsymbol{k}}] = \tag{20}$$

$$= N\sum_{\boldsymbol{k}_2}\left(\epsilon_{+\boldsymbol{k}_2}\mathrm{St}_R[f_{+\boldsymbol{k}_2}]+\epsilon_{-\boldsymbol{k}_2}\mathrm{St}_R[f_{-\boldsymbol{k}_2}]\right) =$$

$$= -2\pi N\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{q}} W_q\delta(\epsilon_{+\boldsymbol{k}_2}-\epsilon_{-\boldsymbol{k}_1}-\omega_q)\frac{\partial n_q^{(0)}}{\partial\omega}\omega_q$$

$$\times\left(f^{(0)}_{+\boldsymbol{k}_2}-f^{(0)}_{-\boldsymbol{k}_1}\right)(h_{-\boldsymbol{k}_1}-h_{+\boldsymbol{k}_2}).$$

Defining the decay coefficients similarly to Eq. (19) above, we may present the result in the form

$$N\sum_{\boldsymbol{k}}\epsilon_{\lambda\boldsymbol{k}}\mathrm{St}_R[f_{\lambda\boldsymbol{k}}] = -\mu_I n_{E,0}\lambda_{QE} \approx -\frac{n_E-n_{E,0}}{\tau_{RE}}. \tag{21}$$

Here the equivalence of the two forms of the decay term stems from the fact that $n_E - n_{E,0} \propto \mu_I$ assuming the electrons and holes are characterized by the same temperature.

Supercollisions are not the only scattering process contributing to both quasiparticle recombination and energy relaxation. Clearly, direct (not impurity-assisted) electron-phonon interaction contributes to energy relaxation as well as to quasiparticle recombination (in the case of intervalley scattering) [14, 18, 19, 34, 37]. In addition, optical phonons may also contribute [35, 36], although within the hydrodynamic approach these contributions were discussed only at the level of dissipative (viscous) hydrodynamics [36]. The contribution of the direct [42, 43] and impurity assisted electron-phonon scattering to energy relaxation was compared in [37], where it was argued that at high enough temperatures, $T \gtrsim T_{BG}$ (where T_{BG} is the Bloch-Grüneisen temperature) the supercollisions do dominate. In the degenerate regime, where $T_{BG} \propto \sqrt{n}$, Ref. [37] estimates T_{BG} as "few tens of Kelvin". At charge neutrality, we estimate $T_{BG} = (s/v_g)T \ll T$ (where s is the speed of sound), such that supercollisions always dominate over direct electron-phonon coupling. Moreover, taking into account the additional scattering processes will not change the form of the decay terms (19) and (21), but rather change the numerical values of the parameters λ_Q and λ_{QE}, which may have to be considered phenomenological while interpreting experimental data [34].

At charge neutrality and in the hydrodynamic regime, the coefficients λ_Q and λ_{QE} are of the same order of magnitude (both are dominated by the same supercollisions), but quantitatively different. Indeed, the continuity equation for the energy density should contain one more contribution of the similar form. "Quasiparticle recombination" typically refers to scattering between the quasiparticle states in different bands only. This is the only type of supercollisions contributing to λ_Q. Similar supercollisions may also take place within a single band [37]. While this process does not change the number of particles in the band, it does describe the energy loss as the electron may scatter from the higher energy state into the lower energy state (and losing its momentum to the impurity along the way). Consequently, this additional process does contribute to energy relaxation. Given that the form of the corresponding collision integral is very similar to Eq. (11) – one only has to change to band indices to be the same – the algebra remains the same and thus we can treat Eq. (21) as the final result that takes this additional intraband supercollisions into account making the numerical values of λ_Q and λ_{QE} substantially different – we do not expect any accidental cancellation or smallness should the difference $\lambda_Q - \lambda_{QE}$ appear in a particular solution of hydrodynamic equations. At the same time, at low temperatures – i.e., below the hydrodynamic range – we expect the coefficients λ_Q and λ_{QE} to be parametrically different: energy relaxation is now dominated by the direct electron-phonon interaction [37], while the recombination is still governed by supercollisions (together with the phonon-induced intervalley scattering).

The order of magnitude of τ_R could be estimated based on the calculations of Ref. [37]. Adapting the latter to charge neutrality, we find $\tau_R^{-1} \sim D^2T^2/(\rho s^2 v_g^2)$ (where $D \approx 20\,\mathrm{eV}$ is the deformation potential [42, 43] and ρ is the mass density per unit area) yielding the corresponding length scale $\ell_R^{-1} \approx 10\,\mu\mathrm{m}$ at the top of the hydrodynamic temperature range, $T \approx 250\,\mathrm{K}$. This should be further compared to the contribution of three-particle collisions [3, 36], $\tau_3^{-1} \sim \alpha_g^4 T$. Assuming the common sample design where graphene is encapsulated in hexagonal boron nitride (with the dielectric constant $\epsilon \approx 4$), the effective coupling constant (taking into account renormalizations) is $\alpha_g \approx 0.3 - 0.4$ leading to the similar estimate at high temperatures. On the other hand, at the low end of the hydrodynamic range [44], $T \approx 50\,\mathrm{K}$, the contribution of the three-body collisions should dominate (accounting for the empirical value $\ell_R \approx 1.2\,\mu\mathrm{m}$ reported in [34]), however preserving the functional form of the weak decay terms in the continuity equations.

Finally, once the dissipative processes due to electron-electron interaction are taken into account, one usually replaces the continuity equation for the energy density by the equivalent equation for the entropy density, the so-called "heat transfer equation" [4]. The decay terms discussed in this paper appear in that equation as well. Let us briefly discuss their form.

Recall the derivation of the continuity equation for the entropy from the kinetic equation [21]. The entropy density of a system of fermions is defined in terms of the distribution function as

$$s = -N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\left[f_{\lambda\boldsymbol{k}}\ln f_{\lambda\boldsymbol{k}} + (1-f_{\lambda\boldsymbol{k}})\ln(1-f_{\lambda\boldsymbol{k}})\right]. \quad (22)$$

Treating this integral as

$$s = N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\mathcal{S}[f_{\lambda\boldsymbol{k}}],$$

any derivative of s can be represented in the form

$$\frac{\partial s}{\partial z} = N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\frac{\partial f_{\lambda\boldsymbol{k}}}{\partial z}.$$

Consider now each term of the kinetic equation multiplied by the derivative $\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]/\partial f_{\lambda\boldsymbol{k}}$ and summed over all states. Using the above relation with $z \to t$, one finds for the time derivative term

$$N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\frac{\partial f_{\lambda\boldsymbol{k}}}{\partial t} = \frac{\partial s}{\partial t}.$$

The gradient term yields similarly

$$N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\boldsymbol{v}_{\lambda\boldsymbol{k}}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}f_{\lambda\boldsymbol{k}} =$$

$$= \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\boldsymbol{v}_{\lambda\boldsymbol{k}}\mathcal{S}[f_{\lambda\boldsymbol{k}}] = \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{j}_S,$$

where the quantity

$$\boldsymbol{j}_S = N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\boldsymbol{v}_{\lambda\boldsymbol{k}}\mathcal{S}[f_{\lambda\boldsymbol{k}}], \quad (23)$$

can be interpreted as the entropy current.

The electric field term vanishes as the total derivative

$$e\boldsymbol{E}\cdot N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\boldsymbol{\nabla}_{\boldsymbol{k}}f_{\lambda\boldsymbol{k}} =$$

$$= e\boldsymbol{E}\cdot N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\boldsymbol{\nabla}_{\boldsymbol{k}}\mathcal{S}[f_{\lambda\boldsymbol{k}}] = 0,$$

while the Lorentz term vanishes for rotationally invariant systems upon integrating by parts [justified by the fact that $\mathcal{S}(k \to \infty) \to 0$]

$$\frac{e}{c}N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\left[\boldsymbol{v}_{\lambda\boldsymbol{k}}\times\boldsymbol{B}\right]\cdot\boldsymbol{\nabla}_{\boldsymbol{k}}f_{\lambda\boldsymbol{k}} =$$

$$= \frac{e}{c}N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\left[\boldsymbol{v}_{\lambda\boldsymbol{k}}\times\boldsymbol{B}\right]\cdot\boldsymbol{\nabla}_{\boldsymbol{k}}\mathcal{S}[f_{\lambda\boldsymbol{k}}]$$

$$= -\frac{e}{c}N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\mathcal{S}[f_{\lambda\boldsymbol{k}}]\boldsymbol{\nabla}_{\boldsymbol{k}}\cdot\left[\boldsymbol{v}_{\lambda\boldsymbol{k}}\times\boldsymbol{B}\right] = 0.$$

The last equality follows from

$$\frac{\partial v^{\alpha}_{\lambda\boldsymbol{k}}}{\partial k^{\beta}} = \frac{v_g}{\lambda k}\left(\delta_{\alpha\beta} - \frac{k^{\alpha}k^{\beta}}{k^2}\right).$$

Similar approach was used in Ref. 21 to derive the continuity equations (as outlined above).

Combining all four terms, we conclude that integration with the factor $\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]/\partial f_{\lambda\boldsymbol{k}}$ turns the left-hand side of the kinetic equation to the familiar form

$$N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\mathcal{L}f_{\lambda\boldsymbol{k}}=\frac{\partial s}{\partial t}+\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{j}_S. \tag{24}$$

Eq. (24) is valid for an arbitrary distribution function.

Denoting the integral of the right-hand side of the kinetic equation by

$$\mathcal{I}=N\sum_{\lambda}\int\frac{d^2k}{(2\pi)^2}\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}\left(\mathrm{St}_{ee}[f]+\mathrm{St}_R[f]+\mathrm{St}_{\mathrm{dis}}[f]\right), \tag{25}$$

we arrive at the "continuity equation for the entropy"

$$\frac{\partial s}{\partial t}+\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{j}_S=\mathcal{I}. \tag{26}$$

In the usual hydrodynamics [4] the only contribution to the collision integral is given by particle-particle scattering, i.e. the processes assumed to be responsible for establishing local equilibrium such that at $\mathcal{I}=0$ the ideal (Euler) hydrodynamic is isentropic. In the present case, local equilibrium is assumed to be achieved by means of the electron-electron interaction.

Evaluating the derivative of $\mathcal{S}$ explicitly, we find

$$\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}=-\ln\frac{f_{\lambda\boldsymbol{k}}}{1-f_{\lambda\boldsymbol{k}}}=\ln\left[\frac{1}{f_{\lambda\boldsymbol{k}}}-1\right].$$

For the local equilibrium distribution function

$$\frac{\partial\mathcal{S}[f_{\lambda\boldsymbol{k}}]}{\partial f_{\lambda\boldsymbol{k}}}=\frac{\epsilon_{\lambda\boldsymbol{k}}-\mu_\lambda-\boldsymbol{u}\cdot\boldsymbol{k}}{T}.$$

Substituting this expression into Eq. (25), we find that the remaining integration is very similar to the above derivation of the continuity equations.

The integral with the quasiparticle energy yields exactly the above Eq. (21). The integral with $\lambda\mu_I$ yields Eq. (19) multiplied by μ_I. Finally, the term $\boldsymbol{u}\cdot\boldsymbol{k}$ yields Eq. (7) multiplied by the hydrodynamic velocity. The integral of this term with the recombination collision integral is assumed to be included into the definition of the mean free time, see the corresponding discussion above. As a result, we arrive at the following form of the integrated collision integral

$$\mathcal{I}=-\frac{1}{T}\frac{n_E-n_{E,0}}{\tau_{RE}}+\frac{\mu_I}{T}\frac{n_I-n_{I,0}}{\tau_R}+\frac{\boldsymbol{u}\cdot\boldsymbol{n}_{\boldsymbol{k}}}{T\tau_{\mathrm{dis}}}.$$

The decay terms (27) appear already at local equilibrium. To complete the heat transfer equation one has to take into account dissipation. In graphene, this is most conveniently done by considering the classical limit of relativistic hydrodynamics since the form of dissipative corrections is determined by the symmetries of the quasiparticle spectrum. The result has been already reported in literature, therefore we combine the dissipative corrections with Eq. (27) and write the heat transfer equation in graphene in the form (here $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ are the dissipative corrections to the electric and imbalance currents, respectively).

$$\begin{aligned}T\left[\frac{\partial s}{\partial t}+\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left(s\boldsymbol{u}-\delta\boldsymbol{j}\frac{\mu}{T}-\delta\boldsymbol{j}_I\frac{\mu_I}{T}\right)\right]&= \\ =\delta\boldsymbol{j}\cdot\left[e\boldsymbol{E}+\frac{e}{c}\boldsymbol{u}\times\boldsymbol{B}-T\boldsymbol{\nabla}\frac{\mu}{T}\right]&-T\delta\boldsymbol{j}_I\cdot\boldsymbol{\nabla}\frac{\mu_I}{T} \\ +\frac{\eta}{2}\left(\nabla_\alpha u_\beta+\nabla_\beta u_\alpha-\delta_{\alpha\beta}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right)^2& \\ -\frac{n_E-n_{E,0}}{\tau_{RE}}+\mu_I\frac{n_I-n_{I,0}}{\tau_R}&+\frac{\boldsymbol{u}\cdot\boldsymbol{n}_{\boldsymbol{k}}}{\tau_{\mathrm{dis}}}.\end{aligned} \tag{27}$$

The obtained equation (27) should be compared to the corresponding equations in Refs. 3, 14, and 36, where energy relaxation due to supercollisions were not taken into account. All other terms are present in all four equations with the following exceptions. The equation (54) of Ref. 3 is written in the relativistic notation omitting the imbalance mode, quasiparticle recombination, and disorder scattering, all of which are discussed separately elsewhere in Ref. 3. Ref. 14 was the first to focus on the imbalance mode with the equation (2.6) containing all the terms of Eq. (27) except for the viscous term (and energy relaxation). Finally, the equation (1c) of Ref. 36 contains all of the terms in Eq. (27) except for energy relaxation and in addition contains a term describing energy relaxation due to electrons scattering on the optical phonon branch that is neglected here (generalization of the resulting theory is straightforward).

To summarize, we have considered supercollisions as a mechanism of quasiparticle recombination and energy relaxation in graphene and derived the corresponding decay terms in the hydrodynamic continuity equations. Since the same scattering mechanism is responsible for both effects, one has to take into account energy relaxation while considering quasiparticle recombination. The latter is an indispensable feature of electronic hydrodynamics in graphene in constrained geometries, where homogeneous solutions violate the boundary conditions [29].

ACKNOWLEDGMENTS

The authors are grateful to P. Alekseev, A. Andreev, U. Briskot, A. Dmitriev, L. Golub, V. Kachorovskii, E. Kiselev, A. Mirlin, J. Schmalian, M. Schütt, A. Shnirman, K. Tikhonov, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 program (Project 873028 HYDROTRONICS), by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant no. I-1505-303.10/2019 (IVG) and by the Russian Science Foundation, Grant No. 20-12-00147 (IVG). BNN acknowledges the support by the MEPhI Academic Excellence Project, Contract No. 02.a03.21.0005.

---

[1] M. Polini and A. K. Geim, Physics Today **73**, 28 (2020).
[2] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).
[3] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).
[4] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1959).
[5] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, 1995).
[6] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968).
[7] A. V. Andreev, S. A. Kivelson, and B. Spivak, Physical Review Letters **106**, 256804 (2011).
[8] S. A. Hartnoll, P. K. Kovtun, M. Müller, and S. Sachdev, Phys. Rev. B **76**, 144502 (2007).
[9] M. Müller, L. Fritz, and S. Sachdev, Phys. Rev. B **78**, 115406 (2008).
[10] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).
[11] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).
[12] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).
[13] M. J. Bhaseen, A. G. Green, and S. L. Sondhi, Phys. Rev. B **79**, 094502 (2009).
[14] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).
[15] A. Tomadin, D. Brida, G. Cerullo, A. C. Ferrari, and M. Polini, Phys. Rev. B **88**, 035430 (2013).
[16] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).
[17] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).
[18] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).
[19] M. Schütt, P. M. Ostrovsky, M. Titov, I. V. Gornyi, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. Lett. **110**, 026601 (2013).
[20] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).
[21] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).
[22] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).
[23] B. N. Narozhny, Phys. Rev. B **100**, 115434 (2019).
[24] D. E. Sheehy and J. Schmalian, Phys. Rev. Lett. **99**, 226803 (2007).
[25] A. A. Kozikov, A. K. Savchenko, B. N. Narozhny, and A. V. Shytov, Phys. Rev. B **82**, 075424 (2010).
[26] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).
[27] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).
[28] M. Zarenia, A. Principi, and G. Vignale, Phys. Rev. B **102**, 214304 (2020).
[29] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).
[30] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).
[31] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).
[32] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).
[33] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, et al., Nat. Phys. **8**, 896 (2012).
[34] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).
[35] H.-Y. Xie and M. S. Foster, Phys. Rev. B **93**, 195103 (2016).
[36] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).
[37] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).
[38] M. W. Graham, S.-F. Shi, D. C. Ralph, J. Park, and P. L. McEuen, Nature Physics **9**, 103 (2013).
[39] A. C. Betz, S. H. Jhang, E. Pallecchi, R. Ferreira, G. Fève, J.-M. Berroir, and B. Plaçais, Nature Physics **9**, 109 (2013).
[40] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **97**, 085415 (2018).
[41] J. F. Kong, L. Levitov, D. Halbertal, and E. Zeldov, Phys. Rev. B **97**, 245416 (2018).
[42] R. Bistritzer and A. H. MacDonald, Phys. Rev. Lett. **102**, 206410 (2009).
[43] W.-K. Tse and S. Das Sarma, Phys. Rev. B **79**, 235406 (2009).
[44] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. Krishna Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

Electronic viscosity and energy relaxation in neutral graphene

Vanessa Gall,[1, 2] Boris N. Narozhny,[2, 3] and Igor V. Gornyi[1, 2, 4]
[1] *Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2] *Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[3] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[4] *Ioffe Institute, 194021 St. Petersburg, Russia*

(Dated: August 30, 2022)

We explore hydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. In the absence of magnetic field, the bulk Ohmic charge flow and the hydrodynamic energy flow are decoupled. However, the energy flow does affect the overall resistance of the system through viscous dissipation and energy relaxation that has to be compensated by the work done by the current source. Solving the hydrodynamic equations, we find that local temperature and electric potential are discontinuous at the interfaces with the leads as well as the device resistance and argue that this makes Corbino geometry a feasible choice for an experimental observation of the Dirac fluid.

Quantum dynamics of charge carriers is one of the most important research directions in condensed matter physics. In many materials transport properties can be successfully described under the assumption of weak electron-electron interaction allowing for free-electron theories [1]. An extension of this approach to strongly-correlated systems remains a major unsolved problem. The advent of "ultra-clean" materials poses new challenges, especially if the electronic system is nondegenerate. At high temperatures such systems may exhibit signatures of a collective motion of charge carriers resembling the hydrodynamic flow of a viscous fluid [2–14].

Electronic viscosity has been discussed theoretically for a long time [15–20], but became the subject of dedicated experiments [2, 9] only recently, after ultra-clean materials became available. Up until now, most experimental efforts were focusing on graphene [2–11] where the hydrodynamic regime is apparently easier to achieve [21, 22]. Viscous effects manifest themselves in nonuniform flows. In the common "linear" geometry (channels, wires, Hall bars, etc.) this occurs in "narrow" samples where the typical length scale associated with viscosity is of the same order as the channel width [23–27]. In contrast, in the "circular" Corbino geometry, see Fig. 1, the electric current is nonuniform even in the simplest Drude picture (in the absence of magnetic field, $\boldsymbol{j} \propto \boldsymbol{e_r}/|\boldsymbol{r}|$, where $\boldsymbol{e_r} = \boldsymbol{r}/|\boldsymbol{r}|$) making it an excellent platform to measure electronic viscosity [28–31]. In the last year, electronic hydrodynamics in the Corbino geometry has been studied both experimentally [32] and theoretically [33–36].

In this paper we address the "Dirac fluid" [3, 9] (the hydrodynamic flow of charge carriers in neutral graphene) in the Corbino geometry. Unlike doped graphene where degenerate, Fermi-liquid-like electrons may be described by the Navier-Stokes equation with a weak damping term due to disorder [16, 21, 23], the two-band physics of neutral graphene leads to unconventional hydrodynamics [22, 37]. In the hydrodynamic approach any macroscopic current can be expressed as a product of the corresponding density and hydrodynamic velocity $\boldsymbol{u}$ (up to dissipative corrections), e.g., the electric and energy current densities are $\boldsymbol{j} = n\boldsymbol{u}$ and $\boldsymbol{j}_E = n_E\boldsymbol{u}$, respectively. In the degenerate regime the charge and energy densities are proportional to each other (to the leading approximation in thermal equilibrium $n_E = 2\mu n/3$, where μ is the chemical potential) and the two currents are equivalent [38]. In contrast, the equilibrium charge density vanishes at charge neutrality, $n(\mu = 0) = 0$, while the energy density remains finite. The two currents "decouple": the energy current remains "hydrodynamic", the charge current is completely determined by the dissipative correction $\delta\boldsymbol{j}$.

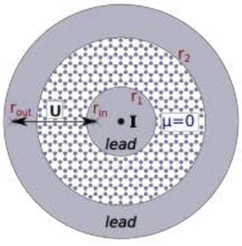


FIG. 1. Corbino geometry: the annulus-shaped sample of neutral graphene ($\mu = 0$) is placed between the the two leads: the inner circle of the radius r_1 and the outer shell with the inner radius r_2. A current I is injected through at the center point and a voltage U is measured between electrodes placed at the inner and outer radius $r_{\rm in}$ and $r_{\rm out}$.

Electronic transport at charge neutrality has been a subject of intensive research [9, 24–27, 38–46] leading to general consensus on the basic result: in the absence of magnetic field, $\boldsymbol{B} = 0$, resistivity of neutral graphene is determined by the electron-electron interaction

$$R_0 = \frac{\pi}{2e^2 T \ln 2}\left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\rm dis}}\right) \xrightarrow[\tau_{\rm dis}\to\infty]{} \frac{1}{\sigma_Q}. \qquad (1)$$

Here $\tau_{11} \propto \alpha_g^{-2} T^{-1}$ describes the appropriate electron-electron collision integral and σ_Q is the "intrinsic" or "quantum" conductivity of graphene. Disorder scattering is characterized by the mean free time $\tau_{\rm dis}$, which is large under the assumptions of the hydrodynamic regime,

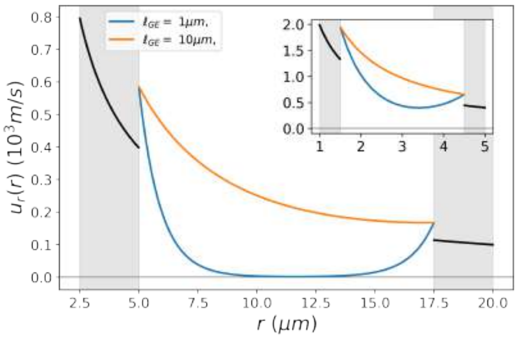


FIG. 2. Radial component of the hydrodynamic velocity u_r. Black lines show the drift velocity in the leads, $u_r^{\rm in(out)} \propto 1/r$. Colored curves correspond to the solution Eq. (4) for the two indicated values of ℓ_{GE}. The results are plotted for the two cases of a large (main panel) and small (inset) device.

$\tau_{\rm dis} \gg \tau_{11}$ and yields a negligible contribution to Eq. (1). Equation (1) describes the uniform bulk current and is independent of viscosity (i.e., in a channel [21, 24, 44, 46]). In contrast, in the Corbino geometry the current flow is necessarily inhomogeneous and hence viscous dissipation must be taken into account.

We envision the following experiment: a graphene sample (at charge neutrality) in the shape of an annulus is placed between the inner (a disk of radius r_1) and outer (a ring with the inner radius r_2) metallic contacts (leads). For simplicity, we assume both leads to be of the same material, e.g., highly doped graphene with the same doping level. The electric current I is injected into the center of the inner lead preserving the rotational invariance (e.g., through a thin vertical wire attached to the center point) and spreads towards the outer lead, which for concreteness we assume to be grounded. The overall voltage drop U is measured between two points in the two leads (at the radii $r_{\rm in} < r_1$ and $r_{\rm out} > r_2$) yielding the device resistance, $R = U/I$. In most traditional measurements, the leads' resistance is minimal, while the contact resistance is important only in ballistic systems, see e.g., Ref. [10]. Hence, one may interpret the measured voltage drop in terms of resistivity of the sample material. Here we focus on the device resistance and show that in the hydrodynamic regime there is an additional contribution due to electronic viscosity and energy relaxation.

Charge flow through the Corbino disk can be described as follows. The injected current spreads through the inner lead according to the Ohm's law and continuity equation. In the stationary case, the latter determines the radial component of the current density, $j_r^{\rm in} = I/(2\pi e r)$. This defines the drift velocity $\boldsymbol{u}^{\rm in} = \boldsymbol{j}^{\rm in}/n^{\rm in}$ ($n^{\rm in}$ is the carrier density in the inner lead) and the energy current $\boldsymbol{j}_E^{\rm in} = n_E^{\rm in}\boldsymbol{u}^{\rm in}$. Reaching the interface, both currents continue to flow into the graphene sample. Here (at $n = 0$ and $\boldsymbol{B} = 0$) the energy current $\boldsymbol{j}_E = n_E\boldsymbol{u}$ is decoupled from the electric current $\boldsymbol{j} = \delta\boldsymbol{j}$. Charge conservation requires the radial component of the electric current to be continuous at the interface, $\delta\boldsymbol{j}(r_1) = \boldsymbol{j}^{\rm in}(r_1)$. Due to the continuity equation, the current density in graphene has the same functional form, $\delta j_r = I/(2\pi e r)$. Does this mean that the device resistance trivially follows if one knows the resistivity of graphene? The answer is "no", since the electrochemical potential is discontinuous at the interface! There are two mechanisms for the "jump" of the potential: (i) the usual Schottky contact resistance [42, 47], and (ii) dissipation due to viscosity [31] and energy relaxation [48]. Since the lost energy must come from the current source, both contribute to R.

The energy flow in neutral graphene is described by the set of hydrodynamic equations developed in Refs. [37, 44, 48] and most recently solved in Ref. [46] in the channel geometry. Within linear response, the equations are

$$\boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{2a}$$

$$n_I\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -(12\ln 2/\pi^2)n_I\mu_I/(T\tau_R), \tag{2b}$$

$$\boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} - 3P\boldsymbol{u}/(v_g^2\tau_{\rm dis}), \tag{2c}$$

$$3P\boldsymbol{\nabla}\cdot\boldsymbol{u} = -2\delta P/\tau_{RE}. \tag{2d}$$

Here Eq. (2a) is the continuity equation; Eq. (2b) is the "imbalance" continuity equation [37, 42] (μ_I is the imbalance chemical potential, $n_I = \pi T^2/(3v_g^2)$ is the equilibrium imbalance density, v_g is the band velocity in graphene, and τ_R is the recombination time); Eq. (2c) is the linearized Navier-Stokes equation [37, 46, 49, 50]; and Eq. (2d) is the linearized "thermal transport" equation (τ_{RE} is the energy relaxation time [48]). Equilibrium thermodynamic quantities (the pressure $P = 3\zeta(3)T^3/(\pi v_g^2)$, enthalpy density $\mathcal{W}$, and energy density are related by the "equation of state", $\mathcal{W} = 3P = 3n_E/2$. The dissipative corrections to the macroscopic currents are given by

$$\delta\boldsymbol{j} = \boldsymbol{E}/(eR_0), \tag{3a}$$

$$\delta\boldsymbol{j}_I = -\frac{2\gamma\ln 2}{\pi}T\tau_{\rm dis}\boldsymbol{\nabla}\mu_I, \quad \gamma = \frac{\delta_I}{1+\tau_{\rm dis}/(\delta_I\tau_{22})}, \tag{3b}$$

where $\tau_{22} \propto \alpha_g^{-2}T^{-1}$ describes a component of the collision integral that is qualitatively similar, but quantitatively distinct from τ_{11} and $\delta_I \approx 0.28$. The equations (2) and (3) should be solved for $\boldsymbol{u}$, $\delta\boldsymbol{j}$, $\delta\boldsymbol{j}_I$, $\boldsymbol{E}$, μ_I, and δP.

Excluding δP from Eqs. (2c) and (2d) we find a second-order differential equation for $\boldsymbol{u}$

$$\eta'\Delta\boldsymbol{u} = 3P\boldsymbol{u}/(v_g^2\tau_{\rm dis}), \quad \eta' = \eta + 3P\tau_{RE}/2. \tag{4a}$$

In the Corbino disk, the general solution for the radial component of the velocity has the form

$$u_r = a_1 I_1\left(\frac{r}{\ell_{GE}}\right) + a_2 K_1\left(\frac{r}{\ell_{GE}}\right), \quad \ell_{GE}^2 = \frac{v_g^2\eta'\tau_{\rm dis}}{3P}, \tag{4b}$$

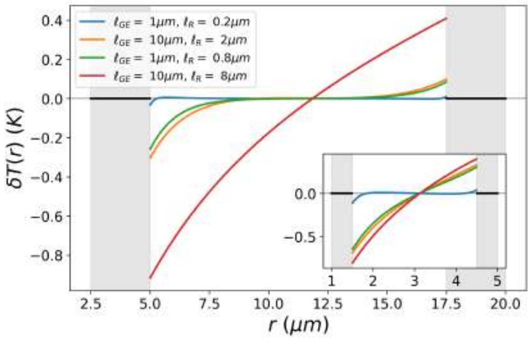


FIG. 3. Temperature distribution in the device. Colored curves correspond to the solution of the hydrodynamic equations for the indicated values of ℓ_{GE} and ℓ_R. The results are plotted for the two cases of a large (main panel) and small (inset) device. In the leads $\delta T = 0$, shown by black lines.

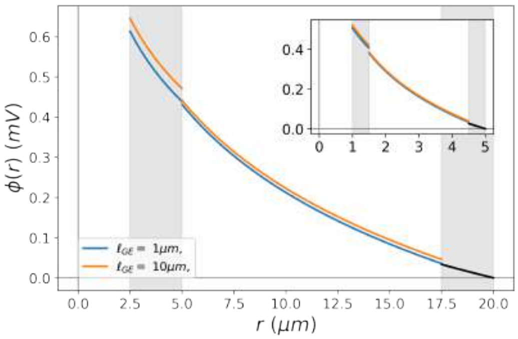


FIG. 4. Electrochemical potential (voltage drop) throughout the device. The black line shows the Ohmic behavior in the outer lead relative to the ground. The jumps at the interfaces are due to dissipative effects (viscosity and energy relaxation) in the bulk of the sample.

where $I_1(z)$ and $K_1(z)$ are the Bessel functions. The coefficients a_1 and a_2 can be found using the continuity of the entropy current at the two interfaces (within linear response). The resulting behavior in shown in Fig. 2 (here we choose to show our results in graphical form since the analytic expressions are somewhat cumbersome [51]; quantitative calculations were performed for $T = 100\,\mathrm{K}$ and experimentally relevant values of the parameters taken from Refs. [8–10, 48]).

In the hydrodynamic regime, the electron-electron scattering time is the shortest scale in the problem, hence the spatial variation of $\boldsymbol{u}$ is determined by energy relaxation. If $\ell_{GE} \ll r_{\rm out} - r_{\rm in}$, then the energy current injected from the leads decays in a (relatively small) boundary region while in the bulk of the sample $\boldsymbol{u} \to 0$. On the other hand, if ℓ_{GE} is of the same order as (or larger than) the system size, then u_r does not vanish and approaches the standard Corbino profile, $u_r \propto 1/r$. At each interface, u_r exhibits a jump due to the mismatch of the entropy densities in the sample and leads.

The nonequilibrium quantities δP and μ_I can now be found straightforwardly. The former follows directly from Eq. (2d) using the solution (4), while the differential equation for the latter can be found by substituting Eq. (3b) into Eq. (2b) and using the solution (4). The boundary conditions for δP and μ_I follow from the continuity equations for the charge and imbalance. The two quantities can be combined to determine the nonequilibrium temperature variation, δT, shown in Fig. 3. For a large sample ($\ell_{GE}, \ell_R \ll r_{\rm out} - r_{\rm in}$, $\ell_R^2 = \gamma v_g^2 \tau_{\rm dis} \tau_R / 2$), δT exhibits fast decay and vanishes in the bulk of the sample. For larger values of ℓ_{GE}, ℓ_R energy relaxation is less effective and the system exhibits an inhomogeneous temperature profile.

The obtained solutions completely describe the hydrodynamic energy flow in neutral graphene. Our remaining task is to find the behavior of the electrochemical potential at the two interfaces enabling us to determine R.

The standard description of interfaces between metals or semiconductors [47] can be carried over to neutral graphene [42] in terms of the contact resistance. Typically, this is a manifestation of the difference of work functions of the two materials across the interface. In graphene, the contact resistance was recently measured in Ref. [10], see also Refs. [32, 52, 53]. In the standard diffusive (or Ohmic) case, the contact resistance leads to a voltage drop that is small compared to the voltage drop in the bulk of the sample and can be ignored. In contrast, in the ballistic case there is almost no voltage drop in the bulk, such that most energy is dissipated at the contacts. Both scenarios neglect electron-electron interactions.

In the diffusive case interactions lead to corrections to the bulk resistivity [54, 55] and the contact resistance can still be ignored. In the ballistic case electron-electron interaction may give rise to a "Knudsen-Poiseuille" crossover [16] and drive the electronic system to the hydrodynamic regime. While the Ohmic resistivity of the electronic fluid may remain small, the hydrodynamic flow possesses another channel for dissipation through viscosity [31]. At charge neutrality, this effect is subtle, since the electric current is decoupled from the hydrodynamic energy flow, see Eq. (3a). At the same time, both are induced by the current source providing the energy dissipated not only by Ohmic effects, but also by viscosity [31] and energy relaxation processes [48] that should be taken into account in the form of an additional voltage drop. Since the voltage drop in the bulk of the sample is completely determined by Eq. (3a), the additional contribution takes the form of a jump in ϕ at the interface corresponding to an excess electric field induced in the thin Knudsen layer around the interface [31].

The magnitude of the jump in ϕ can be established by considering the flow of energy through the interface. Following the standard route [31, 56], we consider the time derivative of the kinetic energy, $\mathcal{A} = \dot{\mathcal{E}}$, where $\mathcal{E}$ is obtained by integrating the energy density $n_E(\boldsymbol{u}) - n_E(0)$

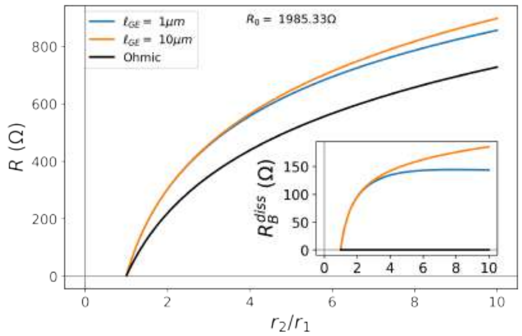


FIG. 5. Total resistance of the Corbino device for different values of ℓ_{GE} (here $r_1 = 0.5\,\mu$m). Inset: additional contribution to the resistance due to viscous dissipation.

over the volume. Working within linear response, we expand the latter to the leading order in the hydrodynamic velocity. Finding time derivatives from the equations of motion and using the continuity equation and partial integration, we then separate the "bulk" and "boundary" contributions, $\mathcal{A} = \mathcal{A}_{\rm bulk} + \mathcal{A}_{\rm edge}$. We interpret the former as the bulk dissipation, while $\mathcal{A}_{\rm edge}$ includes the energy brought in (carried away) through the boundary by the incoming (outgoing) flow. In the stationary state $\dot{\mathcal{E}} = 0$, dissipation is balanced by the work done by the source. Assuming that no energy is accumulated at the interface, we find the corresponding boundary condition.

The specific form of the equations of motion depends on the choice of the material. Assuming the leads' material is highly doped graphene, the equation of motion is the usual Ohm's law with the diffusion term [57] coming from the gradient of the stress-energy tensor [38], here we include a viscous contribution due to disorder [58] and find [31] (omitting the continuous entropy flux)

$$\mathcal{A}^{\rm lead}_{\rm edge} = \int dS_\beta \left(u^L_\alpha \sigma'_{L;\alpha\beta} - u^L_\beta \delta P_L - e\boldsymbol{j}^L_\beta \phi \right), \qquad (5a)$$

where $\boldsymbol{j}^L = n_L \boldsymbol{u}^L$ is the current density, $\boldsymbol{u}^L$ is the drift velocity, δP_L is the nonequilibrium pressure, and σ'_L is the viscous stress tensor in the lead. The first two terms are the usual dissipative contributions to the energy flow across the boundary [56], the last term is the Joule heat.

In neutral graphene, we obtain similar results from the Navier-Stokes equation, except that the Joule heat is now determined by $\delta\boldsymbol{j}$

$$\mathcal{A}^{\rm sample}_{\rm edge} = \int dS_\beta \left(u_\alpha \sigma'_{\alpha\beta} - u_\beta \delta P - e\delta j_\beta \phi \right). \qquad (5b)$$

Equating the two contributions (5) and using the above solutions for the velocity and pressure, we find the jumps of the potential ϕ at the two interfaces. This allows us to determine ϕ everywhere in the device, see Fig. 4, as well as the device resistance.

The total resistance of the Corbino device is shown in Fig. 5. Neglecting hydrodynamic effects, we find the usual logarithmic dependence of R on the system size. Viscosity and energy relaxation provide an additional dissipation channel and hence increase R. Energy relaxation contributes to this increase since it dominates the hydrodynamic energy flow, see Eq. (4). At the same time, the boundary condition for the electric potential, Eqs. (5), is determined by viscosity.

In this paper we have solved the hydrodynamic equations in neutral graphene. We have shown, that despite the known decoupling of the Ohmic charge flow and hydrodynamic energy flow, in Corbino geometry the latter does affect the observable behavior leading to jumps in temperature (shown in Fig. 3) and the electric potential, see Fig. 4. The potential jump is distinct from the usual contact resistance insofar it is a function of the system size. Both effects are observable using the modern imaging techniques (the local temperature variation can be measured using the approach of Refs. [59–61], while measurements of the local potential are at the heart of the technique proposed in Refs. [10, 62]). Hydrodynamics also affects the more conventional transport measurements through the size-dependent contribution to the device resistance, see Fig. 5.

Our results highlight several particular features of the Dirac fluid in neutral graphene. Firstly, the "linear response" currents (3) are independent of the temperature gradient due to exact particle-hole symmetry [42]. Secondly, in contrast to the case of doped graphene [31] *the Dirac fluid is compressible* even within linear response (due to energy relaxation, see Eq. (2d). Finally, the hydrodynamic flow in neutral graphene is the energy flow. Hence, energy relaxation effectively dominates over viscous effects, see Eqs. (4), complicating experimental determination of η.

External magnetic field is also known to couple the charge and energy flows in neutral graphene [37]. We expect that our theory will yield interesting results on Corbino magnetoresistance [52]. Another extension of our theory is the study of thermoelectric phenomena, which is more interesting if one moves away from the neutrality point [34] (where the thermopower must vanish due to the exact particle-hole symmetry). Our results on both issues will be reported elsewhere.

The authors are grateful to P. Hakonen, V. Kachorovskii, A. Levchenko, A. Mirlin, J. Schmalian, A. Shnirman, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 Program (Project 873028 HYDROTRONICS), by the German Research Foundation DFG project NA 1114/5-1 (BNN), and by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant No. I-1505-303.10/2019 (IVG).

Supplemental material

Starting with the general form of the hydrodynamic equations in graphene, we obtain the analytical results presented graphically in the main text. In Sec. we summarize the hydrodynamic equations for graphene. In Sec. we specify these equations within linear response in polar coordinates at charge neutrality and $B = 0$. In Sec. we formulate a description of the leads followed by the relevant boundary conditions at the lead-graphene interfaces in Sec. . Next, in Sec. we present the full analytical solution for the hydrodynamic equations in the Corbino geometry with the above boundary conditions. In Sec. we discuss the dissipation in the system and corroborate the argument used in the main text to obtain the device resistance. Finally, we conclude with a brief analysis in Sec. .

Electronic hydrodynamics in graphene

Following Ref. [63] we combine the chemical potentials of the two bands in graphene $\mu_\pm$ into

$$\mu = (\mu_+ + \mu_-)/2, \quad \mu_I = (\mu_+ - \mu_-)/2 \tag{6}$$

and introduce their conjugate charge and imbalance densities

$$n = n_+ - n_-, \quad n_I = n_+ + n_-. \tag{7}$$

Taking into account dissipative corrections due to electron-electron collisions we then obtain the electric ($\vec{j}$) and imbalance ($\vec{j}_I$) currents as

$$\vec{j} = n\vec{u} + \delta\vec{j}, \quad \vec{j}_I = n_I\vec{u} + \delta\vec{j}_I, \tag{8}$$

where $\vec{u}$ is the drift velocity. The energy current $\vec{j}_E = n_E\vec{u}$ is proportional to the momentum density and is not relaxed by electron-electron collisions. The currents $\vec{j}$ and $\vec{j}_I$ satisfy the continuity equations

$$\partial_t n + \vec{\nabla}\cdot\vec{j} = 0, \tag{9a}$$

which describes the exact conservation of charge and

$$\partial_t n_I + \vec{\nabla}\cdot\vec{j}_I = -\frac{n_I - n_{I,0}}{\tau_R} = -\frac{12\ln 2}{\pi^2}\frac{n_{I,0}\mu_I}{T\tau_R}, \tag{9b}$$

where $n_{I,0} = \pi T^2/(3v_g^2)$ is the equilibrium value of the total quasiparticle density (at $\mu_I = 0$) and τ_R is the recombination time.

A similar equation can be formulated for the energy density

$$\partial_t n_E + \vec{\nabla}\cdot\vec{j}_E = e\vec{j}\vec{E} - \frac{n_E - n_{E,0}}{\tau_{\rm RE}}, \tag{9c}$$

where $\tau_{\rm RE}$ is the energy relaxation time. Typically this is replaced by the thermal transport equation

$$\begin{aligned} T\left[\frac{\partial s}{\partial t} + \vec{\nabla}_{\vec{r}}\cdot\left(s\vec{u} - \delta\vec{j}\frac{\mu}{T} - \delta\vec{j}_I\frac{\mu_I}{T}\right)\right] &= \delta\vec{j}\cdot\left[e\vec{E} + \frac{e}{c}\vec{u}\times\vec{B} - T\vec{\nabla}\frac{\mu}{T}\right] - T\delta\vec{j}_I\cdot\vec{\nabla}\frac{\mu_I}{T} \\ &+ \frac{\eta}{2}\left(\nabla_\alpha u_\beta + \nabla_\beta u_\alpha - \delta_{\alpha\beta}\vec{\nabla}\cdot\vec{u}\right)^2 - \frac{n_E - n_{E,0}}{\tau_{RE}} + \mu_I\frac{n_I - n_{I,0}}{\tau_R} + \frac{\mathcal{W}\vec{u}^2}{v_g^2\tau_{\rm dis}}. \end{aligned} \tag{9d}$$

Within linear response the two equations coincide. Finally, the generalized Navier-Stokes equation is given by

$$\mathcal{W}(\partial_t + \vec{u}\cdot\vec{\nabla})\vec{u} + v_g^2\vec{\nabla}P + \vec{u}\partial_t P + e(\vec{E}\cdot\vec{j})\vec{u} = v_g^2\left[\eta\Delta\vec{u} - \eta_H\Delta\vec{u}\times\vec{e}_B + en\vec{E} + \frac{e}{c}\vec{j}\times\vec{B}\right] - \frac{\vec{j}_E}{\tau_{\rm dis}}. \tag{9e}$$

Here η and η_H are the shear and Hall viscosity coefficients, respectively.

The expressions for the dissipative corrections can be found in the Appendix of Ref. [63].

Charge neutral Corbino disk at $B = 0$

Taking into account the rotational symmetry of the Corbino disk, we express the hydrodynamic theory in polar coordinates (r, ϑ). All quantities can only depend on the radial component r. Within linear response and at $B = 0$, the hydrodynamic equations (9) can be transformed to

$$\frac{1}{r}\frac{\partial(r\delta j_r)}{\partial r} = 0, \tag{10a}$$

$$n_{I,0}\frac{1}{r}\frac{\partial(ru_r)}{\partial r} + \frac{1}{r}\frac{\partial(r\delta j_{Ir})}{\partial r} = -\frac{12\ln 2}{\pi^2}\frac{n_{I,0}\mu_I(r)}{T\tau_R}, \tag{10b}$$

$$u_\vartheta = 0, \tag{10c}$$

$$\frac{\partial\delta P}{\partial r} = \eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_r)}{\partial r}\right) - \frac{3Pu_r}{v_g^2\tau_{\text{dis}}}, \tag{10d}$$

$$3P\frac{1}{r}\frac{\partial(ru_r)}{\partial r} = -\frac{2\delta P(r)}{\tau_{RE}}. \tag{10e}$$

The electric field $\vec{E}$ does not appear in Eqs. (10) due to charge neutrality. It does however determine the dissipative correction $\delta\vec{j}$ which at charge neutrality is the whole current. In the absence of the magnetic field, all currents are radial. In polar coordinates, the dissipative corrections take the form

$$\delta j_r = \frac{E_r(r)}{eR_0}, \tag{11a}$$

$$\delta j_\vartheta = 0, \tag{11b}$$

$$\delta j_{Ir} = -\frac{\delta_I}{\tau_{\text{dis}}^{-1} + \delta_I^{-1}\tau_{22}^{-1}}\frac{2T\ln 2}{\pi}\frac{\partial\mu_I}{\partial r}, \tag{11c}$$

$$\delta j_{I\vartheta} = 0. \tag{11d}$$

Equations (10) and (11) have to be solved together taking into account the corresponding boundary conditions see below.

For the purposes of establishing the boundary conditions we also need to specify the stress tensor. At $B = 0$ (and within linear response, meaning neglecting terms that are higher than the leading order in velocity or its derivatives), the stress tensor is

$$\Pi_E^{\alpha\beta} = P\delta^{\alpha\beta} - \sigma^{\alpha\beta}. \tag{12}$$

Since the Hall viscosity vanishes at charge neutrality (as well as at $B = 0$), the viscous stress tensor in Cartesian coordinates is given by

$$\sigma^{\alpha\beta} = \eta\left(\nabla^\alpha u^\beta + \nabla^\beta u^\alpha - \delta^{\alpha\beta}\vec{\nabla}\cdot\vec{u}\right), \tag{13}$$

which in polar coordinates becomes

$$\sigma_{rr} = -\sigma_{\vartheta\vartheta} = \eta\left(\partial_r - \frac{1}{r}\right)u_r, \quad \sigma_{r\vartheta} = \sigma_{\vartheta r} = \eta\left(\partial_r - \frac{1}{r}\right)u_\vartheta. \tag{14}$$

Description of leads

The leads, which are attached at the inner and outer radius of the Corbino disk, are assumed to be a normal metal in the degenerate regime ($\mu_L \gg T$), where transport is dominated by disorder scattering characterized by the relaxation time τ_L. In this case we may restrict ourselves to a single band, such that there is a single macroscopic current satisfying the continuity equation

$$\partial_t n_L + \vec{\nabla}\vec{j} = 0. \tag{15}$$

Within linear response, one can obtain the macroscopic equation of motion (or generalized Ohm's law) integrating the kinetic equation [64]. This way one finds

$$m\partial_t\vec{j} + \vec{\nabla}\check{\Pi}_E - en_L\vec{E} - \frac{e}{c}\vec{j}\times\vec{B} = -\frac{m}{\tau_L}\vec{j}, \tag{16}$$

where the stress tensor may me expressed in terms of thermodynamic pressure and disorder-induced viscosity

$$\Pi_E^{\alpha\beta} = P\delta^{\alpha\beta} - \sigma^{\alpha\beta}, \qquad \eta_L = \frac{\mu^3\tau_L}{4\pi v_g^2\hbar^2}. \tag{17}$$

To be concrete, we assume that the leads' material is doped graphene. In that case we may introduce the "effective mass" $m = \mu_L/v_g^2$ and the drift velocity $\vec{u}_L$, such that $\vec{j} = n_L\vec{u}_L$. Expressing the carrier density in terms of pressure, we find

$$m\vec{j} = \frac{3P_L}{v_g^2}\vec{u}_L, \tag{18}$$

where to lowest order in temperature we find $P_L = \mu^3/(3\pi v_g^2\hbar^2)$. In the stationary state and at $B = 0$, the equations of motion become

$$\vec{\nabla}\vec{u}_L = 0, \tag{19}$$

$$\vec{\nabla}\check{\Pi}_E + en_L\vec{\nabla}\phi = -\frac{3P_L}{v_g^2\tau_L}\vec{u}_L. \tag{20}$$

Experimentally, the density n_L and the chemical potential μ are fixed by the gate voltage. Moreover, under the common assumption of fast equilibration in the leads, we may assume a uniform temperature T as well. The general variation of P_L is found to be

$$\delta P_L = \left(\frac{2\pi\mu T\delta T}{3v_g^2} + \frac{\pi T^2\delta\mu}{3v_g^2} + \frac{\mu^2\delta\mu}{\pi v_g^2}\right) \tag{21}$$

and thus vanishes under the condition we consider. Since the leads are highly doped, we find $n_L = n_+ = n_I$, such that the imbalance chemical potential μ_I vanishes.

Boundary conditions

The differential equations (10) and (11) should be supplemented by a suitable set of boundary conditions. The only boundaries present in the Corbino are boundaries between the sample and the leads. Since charge conservation is exact and also holds in the leads, we find

$$j_r(r_1 - \epsilon) = \delta j_r(r_1 + \epsilon), \quad \delta j_r(r_2 - \epsilon) = j_r(r_2 + \epsilon). \tag{22}$$

Fixing the total current I in a radially symmetric system completely determines the current density

$$I = e\int \mathrm{d}\vec{A}\cdot\vec{j} = 2\pi e r j_r. \tag{23}$$

In contrast, the total quasiparticle number (imbalance) and entropy are not conserved due to recombination and energy relaxation processes. However, assuming that the corresponding relaxation rates are not singular at the interface, the continuity equations (9b) and (9d) yield the following boundary conditions for the radial components of the current densities.

The resulting boundary conditions at the two interfaces can be summarized as follows

$$\begin{aligned}
j_r(r_1-\epsilon) &= n_L u_r(r_1-\epsilon) = \delta j_r(r_1+\epsilon), && (24)\\
j_{I,r}(r_1-\epsilon) &= n_L u_{L,r}(r_1-\epsilon) = n_{I,0}u_r(r_1+\epsilon)+\delta j_{I,r}(r_1+\epsilon) = \delta j_r(r_1+\epsilon), && (25)\\
s_L u_{L,r}(r_1-\epsilon) &= s_B u_r(r_1+\epsilon) && (26)
\end{aligned}$$

$$\begin{aligned}
j_r(r_2+\epsilon) &= n_L u_{L,r}(r_2+\epsilon) = \delta j_r(r_2-\epsilon), && (27)\\
j_{I,r}(r_2+\epsilon) &= n_L u_{L,r}(r_2+\epsilon) = n_{I,0}u_r(r_2-\epsilon)+\delta j_{I,r}(r_2-\epsilon) = \delta j_r(r_2-\epsilon), && (28)\\
s_L u_{L,r}(r_2+\epsilon) &= s_B u_r(r_2-\epsilon). && (29)
\end{aligned}$$

Full solution

Solving the equations of motion in the leads, we find

$$\begin{aligned}
u_{L,r} &= \frac{I}{2\pi e n_L r}, \quad u_{L,\vartheta}=0, && (30)\\
\sigma_{rr} &= \frac{-I\eta_L}{\pi e n_L r^2}, \quad \sigma_{r\vartheta}=0, && (31)\\
E_r &= \frac{2P_L}{e n_L v_g^2 \tau_L}\frac{I}{2\pi e n_L r}, && (32)\\
\phi(r) &= -\frac{I}{2\pi}\frac{2P_L}{e^2 n_L^2 v_g^2 \tau_L}\log\left(\frac{r}{r_0}\right). && (33)
\end{aligned}$$

Here the drift velocity follows from the continuity equation and the relation to the current which in turn is given by Eq. (23). After that, the assumption $\delta P = 0$ leads to the simple $1/r$ behavior for the electrical field E_r as well. Consequently, the charge density (from the Poisson equation) is indeed constant. On the other hand, the constant r_0 in the potential is not fixed by the boundary conditions we have imposed so far. Finally, neither the electric field nor the current depend on the disorder dominated viscosity η_L. However, the viscous stress tensor itself is not zero, which will be used below later.

The above expressions can be re-written in terms of the temperature T and the chemical potential μ_L in the leads. Under our assumptions, the leads' material is graphene, where the entropy density is defined as

$$Ts = 3P - \mu n - \mu_I n_I. \tag{34}$$

For $\mu \gg T$ in the leads we then find

$$\begin{aligned}
P_L &= \frac{\pi T^2 \mu}{3v_g^2} + \frac{\mu^3}{3\pi v_g^2} = P_L^T + P_L^{T=0}, && (35)\\
n_L &= \frac{\pi T^2}{3v_g^2} + \frac{\mu^2}{\pi v_g^2} && (36)\\
s_L T &= 3P_L - n_L\mu = \frac{\pi T^2\mu}{v_g^2} + \frac{\mu^3}{\pi v_g^2} - \frac{\pi T^2 \mu}{3v_g^2} - \frac{\mu^3}{\pi v_g^2} = \frac{2}{3}\frac{\pi T^2\mu}{v_g^2} = 2P_L^T, && (37)
\end{aligned}$$

so we need to keep finite temperature corrections in the leads as well.

In our sample, the situation is more involved since in neutral graphene the electric current is not related to the hydrodynamic velocity. As a manifestation of this fact, the differential equations (10) and (11) decouple into two disjunct sets. The first one consists of equations (10a) and(11a) with the solution

$$\begin{aligned}
\delta j_r &= \frac{I}{2\pi e r}, && (38)\\
E_r &= \frac{IR_0}{2\pi r}, \quad \phi = -\frac{IR_0}{2\pi}\log\left(\frac{r}{r_0}\right). && (39)
\end{aligned}$$

The constant r_0 (not necessarily the same as in the corresponding solution for the leads) is not fixed by the boundary conditions we have imposed so far.

The second set of equations consists of (10b), (10d), (10e) and (11c). Expressing δP through u_r, we find

$$0 = \partial_r \left(\frac{1}{r} \frac{\partial (r u_r)}{\partial r} \right) - \frac{u_r}{\ell_{\mathrm{GE}}^2} \tag{40}$$

$$\frac{1}{\ell_{\mathrm{GE}}^2} = \left(\eta + \frac{3P\tau_{RE}}{2} \right)^{-1} \frac{3P}{v_g^2 \tau_{\mathrm{dis}}}, \tag{41}$$

where the Gurzhi length is renormalized by energy relaxation through the combination $\eta' = \eta + 3P\tau_{RE}/2$. The other two equations can be combined to form

$$\partial_r \left(\frac{1}{r} \frac{\partial (r u_r)}{\partial r} \right) - M \partial_r \left(\frac{1}{r} \frac{\partial (r \frac{\partial \mu_I}{\partial r})}{\partial r} \right) = - \frac{M}{\ell_R^2} \frac{\partial \mu_I(r)}{\partial r} \tag{42}$$

$$M = \frac{2T \ln 2}{n_{I,0}\pi} \frac{\delta_I}{\tau_{\mathrm{dis}}^{-1} + \delta_I^{-1} \tau_{22}^{-1}}, \quad \ell_R^2 = \frac{\delta_I}{\tau_{\mathrm{dis}}^{-1} + \delta_I^{-1} \tau_{22}^{-1}} \frac{\pi T^2 \tau_R}{6 n_{I,0}}. \tag{43}$$

The two coupled Bessel differential equations for u_r and $\partial_r \mu_I$ can be expressed using the differential operator $\mathbb{D} = \partial_r (1/r) \partial_r r$. This way we can write the system of equations in the matrix form

$$\mathbb{D} \begin{pmatrix} 1 & 0 \\ 1 & -M \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} = \begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ 0 & -\frac{M}{\ell_R^2} \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} \Leftrightarrow \mathbb{D} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix} = \begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ \frac{1}{M \ell_{\mathrm{GE}}^2} & \frac{1}{\ell_R^2} \end{pmatrix} \begin{pmatrix} u_r \\ \frac{\partial \mu_I}{\partial r} \end{pmatrix}. \tag{44}$$

This can be formally solved by diagonalizing the matrix

$$\begin{pmatrix} \frac{1}{\ell_{\mathrm{GE}}^2} & 0 \\ \frac{1}{M \ell_{\mathrm{GE}}^2} & \frac{1}{\ell_R^2} \end{pmatrix} = \hat{U}^{-1} \hat{D} \hat{U}, \tag{45}$$

where $\hat{D}$ is a diagonal matrix with the eigenvalues d_1 and d_2 (in units of inverse length squared) and then transforming back to the u_r and $\partial_r \mu_I$ basis. Then this coupled Bessel differential equation has the general solution

$$u_r = M \left(1 - \frac{\ell_{\mathrm{GE}}^2}{\ell_R^2} \right) \left[f_1 I_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + f_2 K_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) \right] \tag{46}$$

$$\frac{\partial \mu_I}{\partial r} = f_1 I_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + f_2 K_1 \left(\frac{r}{\ell_{\mathrm{GE}}} \right) + g_1 I_1 \left(\frac{r}{\ell_R} \right) + g_2 K_1 \left(\frac{r}{\ell_R} \right), \tag{47}$$

where the coefficients f_1, f_2, g_1 and g_2 should be determined from the boundary conditions. These involve the entropy density

$$T s_B = 3P = 3 \frac{3T^3 \zeta(3)}{\pi v_g^2}. \tag{48}$$

From the conservation of entropy current Eqs. (26) and (29) we find f_1 and f_2 so that

$$u_r = \frac{I s_L}{2\pi e n_L s_B} \left\{ \frac{I_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right] - K_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}{r_1 r_2 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 r_2 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)} \right\}. \tag{49}$$

This leads to the stress tensor elements

$$\sigma_{rr} = \frac{\eta I s_L}{2\pi e \ell_{\mathrm{GE}} n_L s_B} \frac{I_2\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right] + K_2\left(\frac{r}{\ell_{\mathrm{GE}}}\right) \left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}{r_1 r_2 \left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) \right]}, \tag{50}$$

$$\sigma_{r\vartheta} = 0 \tag{51}$$

and

$$\delta P = -\frac{3P\tau_{\mathrm{RE}}}{2}\frac{1}{r}\frac{\partial(ru_r)}{\partial r} = \tag{52}$$

$$= -\frac{3P\tau_{\mathrm{RE}}}{2}\frac{Is_L}{2\pi e\ell_{\mathrm{GE}}n_L s_B}\left[\frac{K_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right] + I_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]}{r_1 r_2 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 r_2 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right],$$

Using the conservation of the imbalance current Eqs. (25) and (28) we find the imbalance chemical potential

$$\begin{aligned}
\mu_I(r) = \frac{Is_L\ell_R}{2\pi eMn_L r_1 r_2 s_B(\ell_{\mathrm{GE}}^2 - \ell_R^2)}&\left[\frac{K_0\left(\frac{r}{\ell_R}\right)\left[r_1 I_1\left(\frac{r_1}{\ell_R}\right) - r_2 I_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2 + \left(\ell_R^2 - \ell_{\mathrm{GE}}^2\right)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right) I_1\left(\frac{r_2}{\ell_R}\right) - I_1\left(\frac{r_1}{\ell_R}\right) K_1\left(\frac{r_2}{\ell_R}\right)}\right.\\
&+\frac{I_0\left(\frac{r}{\ell_R}\right)\left[r_1 K_1\left(\frac{r_1}{\ell_R}\right) - r_2 K_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2 + \left(\ell_R^2 - \ell_{\mathrm{GE}}^2\right)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right) I_1\left(\frac{r_2}{\ell_R}\right) - I_1\left(\frac{r_1}{\ell_R}\right) K_1\left(\frac{r_2}{\ell_R}\right)}\\
&\left.+\frac{\ell_{\mathrm{GE}}\ell_R K_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)} + \frac{\ell_{\mathrm{GE}}\ell_R I_0\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right]
\end{aligned}\tag{53}$$

and the dissipative correction to the imbalance current

$$\begin{aligned}
\delta j_{Ir}(r) = \frac{In_{I,0}s_L}{2\pi en_L r_1 r_2 s_B\left(\ell_{\mathrm{GE}}^2 - \ell_R^2\right)}&\left[\frac{K_1\left(\frac{r}{\ell_R}\right)\left[r_1 I_1\left(\frac{r_1}{\ell_R}\right) - r_2 I_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2 + \left(\ell_R^2 - \ell_{\mathrm{GE}}^2\right)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right) I_1\left(\frac{r_2}{\ell_R}\right) - I_1\left(\frac{r_1}{\ell_R}\right) K_1\left(\frac{r_2}{\ell_R}\right)}\right.\\
&-\frac{I_1\left(\frac{r}{\ell_R}\right)\left[r_1 K_1\left(\frac{r_1}{\ell_R}\right) - r_2 K_1\left(\frac{r_2}{\ell_R}\right)\right]\left[\ell_{\mathrm{GE}}^2 + \left(\ell_R^2 - \ell_{\mathrm{GE}}^2\right)\frac{n_L}{n_{I,0}}\frac{s_B}{s_L}\right]}{K_1\left(\frac{r_1}{\ell_R}\right) I_1\left(\frac{r_2}{\ell_R}\right) - I_1\left(\frac{r_1}{\ell_R}\right) K_1\left(\frac{r_2}{\ell_R}\right)}\\
&\left.+\frac{\ell_R^2 K_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_1 I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) - r_2 I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)} + \frac{\ell_R^2 I_1\left(\frac{r}{\ell_{\mathrm{GE}}}\right)\left[r_2 K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1 K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)\right]}{K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right) K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}\right].
\end{aligned}\tag{54}$$

From δP and μ_I we find δT according to

$$\delta T = \frac{\pi v_g^2}{9T^2\zeta(3)}\delta P - \frac{\pi^2}{27\zeta(3)}\mu_I. \tag{55}$$

Our hydrodynamic system is not characterized by a local thermal conductivity κ. In other words, the heat current

$$\vec{j}_Q(r) = 3P\vec{u} - \mu\vec{j} - \mu_I\vec{j}_I \tag{56}$$

is related to the temperature gradient $\nabla\delta T(r)$ at the same point r non locally. The non-local (integral) relation between $\vec{j}_Q(r)$ and $\nabla\delta T(r')$ characterized by a non-local kernel $\kappa(r,r')$ follows from the fact that the equation for $\vec{u}(r)$ is now a second-order differential equation with a non-local Green's function. Expressing $\delta P(r)$ there in terms of $\delta T(r)$ and $\mu_I(r)$, we have a non-local relation between $\vec{u}(r)$, $\delta T(r')$ and $\nabla\mu_I(r')$. Substituting this $\vec{u}(r)$ into the definition of $\vec{j}_Q(r)$, we obtain a non-local thermal conductivity. As a result one can only introduce the thermal conductance for the device, relating the temperature difference between the contacts with the total heat current through the system. This will be done in a subsequent publication.

Dissipation and total resistance

The above solution is not sufficient to determine the drop in electrochemical potential between the points r_{in} and r_{out} (in the inner and outer lead, respectively) since it contains the undefined constant r_0 that has to be determined from a boundary condition for the electric potential. Although microscopically the potential has to be continuous, several effects might contribute to an apparent discontinuity on the hydrodynamic scale. The most obvious contribution is the contact resistance that is a manifestation of the different work functions in the two materials across the interface as well as the mismatch in their chemical potentials [65]. A more subtle effect due to electron-electron interaction giving rise to viscosity and hence an additional dissipation channel [66]. At charge neutrality, this effect is subtle, since the electric current is decoupled from the hydrodynamic energy flow. However, both flows are induced by the same current source providing the energy dissipated by means of both the Ohmic and viscous effects [66] as well as energy relaxation [48]. The latter processes should be taken into account in the form of an additional voltage drop at the interface.

Under the assumption that energy is not being accumulated at the interface, we generalize the idea proposed in Ref. [66] and consider viscous dissipation in the sample. Since the electric field in bulk of the sample is completely determined by the Ohmic resistance R_0, additional dissipation due to viscosity and energy relaxation corresponds to a jump in the electric potential (on the hydrodynamic scale) at the interface. Microscopically, the voltage jump is due to an excess electric field in the Knudsen layer around the sample-lead boundary.

Consider the kinetic energy associated with the hydrodynamic flow that can be found from the energy density

$$\mathcal{E} = \int \mathrm{d}V \left(n_E - n_E(\vec{u}=0)\right) \approx \int \mathrm{d}V \frac{6P}{v_g^2}\vec{u}^2. \tag{57}$$

Working within linear response, here we only keep terms up to the second order in $\vec{u}$ (and thus the drive I). Dissipation is then describe by

$$\mathcal{A} = \dot{\mathcal{E}} = 2\frac{6P}{v_g^2}\int \mathrm{dV}\vec{u}\partial_t\vec{u} = 0, \tag{58}$$

vanishing in the steady state. This expression can now be simplified using the generalized Navier-Stokes equation.

In the leads (still assuming graphene at finite carrier density) we find

$$\begin{aligned}\frac{3P_L}{v_g^2}\vec{u}_L\partial_t\vec{u}_L &= \vec{u}_L\left(-\frac{3P_L}{v_g^2}\frac{\vec{u}_L}{\tau_L} - \vec{\nabla}\check{\Pi}_E + n_L e\vec{E}\right)\\ &= -\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} - \vec{\nabla}\delta P\vec{u}_L + \vec{u}_L\vec{\nabla}\check{\sigma} - e\vec{j}\vec{\nabla}\phi.\\ &= -\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} - \frac{\partial u_{L,i}}{\partial x_j}\sigma_{ij} + \vec{\nabla}\left(\vec{u}_L\check{\sigma} - e\vec{j}\phi - \vec{u}_L\delta P\right).\end{aligned} \tag{59}$$

The term $en_L\vec{u}_L\vec{E} = e\vec{j}\vec{E}$ is the Joule heating. Using the divergence theorem we can divide this into a boundary and a bulk term

$$0 = \mathcal{A} = \mathcal{A}_{\mathrm{boundary}} - \mathcal{A}_{\mathrm{bulk}}, \tag{60}$$

$$\mathcal{A}_{\mathrm{boundary}} = 4\int \mathrm{d}\vec{A}\left(\vec{u}_L\check{\sigma} - \vec{u}_L\delta P - e\vec{j}\phi\right), \tag{61}$$

$$\mathcal{A}_{\mathrm{bulk}} = 4\int \mathrm{d}V\left(\frac{3P_L}{v_g^2}\frac{\vec{u}_L^2}{\tau_L} + \frac{\partial u_{L,i}}{\partial x_j}\sigma_{ij}\right). \tag{62}$$

The boundary term includes the energy transmitted through the interface.

Since the current density is conserved at the interface, we can immediately write down the corresponding equation in the neutral graphene sample, where the Joule heating is given by $e\delta\vec{j}\vec{E}$. Then we find

$$0 = \mathcal{A} = \mathcal{A}_{\mathrm{boundary}} - \mathcal{A}_{\mathrm{bulk}}, \tag{63}$$

$$\mathcal{A}_{\mathrm{boundary}} = 4\int \mathrm{d}\vec{A}\left(\vec{u}\sigma - \vec{u}\delta P - e\delta\vec{j}\phi\right), \tag{64}$$

$$\mathcal{A}_{\mathrm{bulk}} = 4\int \mathrm{d}V\left(\frac{3P}{v_g^2}\frac{\vec{u}^2}{\tau_{\mathrm{dis}}} + \frac{\partial u_i}{\partial x_j}\sigma_{ij} - \delta P(\vec{\nabla}\cdot\vec{u})\right). \tag{65}$$

As stated above, under realistic experimental conditions the non-equilibrium part of the pressure at $u = 0$ on the lead side vanishes

$$\delta P = 0. \tag{66}$$

At the same time, in neutral graphene sample we find

$$\delta P = \left(\frac{9T^2\delta T\zeta(3)}{\pi v_g^2} + \frac{\pi\mu_I T^2}{3v_g^2}\right) = \frac{T^2}{v_g^2}\left(\frac{9\delta T\zeta(3)}{\pi} + \frac{\pi\mu_I}{3}\right). \tag{67}$$

Using the hydrodynamic equations, one may replace δP by $[-3P_B\tau_{\mathrm{RE}}/(2r)]\partial(ru_r)/\partial r$, thus determining δP for finite τ_{RE} without any additional boundary conditions. The same goes for μ_I. Thus we may use the viscous part of the dissipation to find the difference in the electrochemical potential across the interface.

In addition, we may include the contact resistance described by

$$I^2R_c = \vec{I}^T\check{R}\vec{I}, \tag{68}$$

where $\vec{I}$ includes charge and entropy current and $\check{R}$ includes the thermoelectric coefficients of the interface.

In the absence of magnetic field, both u_ϑ and $\sigma_{r\vartheta}$ vanish. In the leads $\delta P = 0$ and hence we find the condition

$$\begin{aligned} &4\pi\left[r\left(u_r\sigma_{rr}\right)\right]_{r_1-\epsilon} - 2I\phi(r_1-\epsilon) = 4\pi\left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_1+\varepsilon} - 2I\phi(r_1+\epsilon) - 2\vec{I}^T\check{R}\vec{I} \\ &\Leftrightarrow \phi(r_1-\epsilon) - \phi(r_1+\epsilon) = \frac{2\pi}{I}\left\{\left[r\left(u_r\sigma_{rr}\right)\right]_{r_1-\epsilon} - \left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_1+\varepsilon}\right\} + IR_c \end{aligned} \tag{69}$$

at the first interface and similarly for the second interface

$$\begin{aligned} &4\pi\left[r\left(u_r\sigma_{rr} - \delta P u_r\right)\right]_{r_2-\epsilon} - 2I\phi(r_2-\epsilon) = 4\pi\left[r\left(u_r\sigma_{rr}\right)\right]_{r_2+\varepsilon} - 2I\phi(r_2+\epsilon) - 2\vec{I}^T\check{R}\vec{I} \\ &\Leftrightarrow \phi(r_2-\epsilon) - \phi(r_2+\epsilon) = -\frac{2\pi}{I}\left\{\left[r\left(u_r\sigma_{rr}\right)\right]_{r_2+\epsilon} - \left[r\left(u_r\sigma_{rr} - u_r\delta P\right)\right]_{r_2-\varepsilon}\right\} + IR_c. \end{aligned} \tag{70}$$

Combining the above general solution with these conditions, we find at r_1

$$ru_r(\sigma_{rr} - \delta P) = \frac{I^2s_L^2\eta\ell_{\mathrm{GE}}}{4\pi^2e^2n_L^2s_B^2\ell_G^2}\frac{\ell_{\mathrm{GE}} - r_2K_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_2I_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}{r_1r_2\left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]} - \frac{\eta I^2s_L^2}{2\pi^2e^2n_L^2r_1^2s_B^2}$$

and at r_2

$$ru_r(\sigma_{rr} - \delta P) = -\frac{I^2s_L^2\eta\ell_{\mathrm{GE}}}{4\pi^2e^2\ell_G^2n_L^2s_B^2}\frac{\ell_{\mathrm{GE}} - r_1I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - r_1K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)}{r_1r_2\left[K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right]} - \frac{\eta I^2s_L^2}{2\pi^2e^2n_L^2r_2^2s_B^2}.$$

As a result, we find the total resistance R of the system in the form

$$IR = \phi(r_{\mathrm{in}}) - \phi(r_{\mathrm{out}}) = I(R_L + R_B + 2R_C + R_L^{\mathrm{diss}} + R_B^{\mathrm{diss}}), \tag{71}$$

$$R_L = \frac{3P_L}{2\pi e^2n_L^2v_g^2\tau_L}\log\left(\frac{r_1r_{\mathrm{out}}}{r_{\mathrm{in}}r_2}\right), \quad R_B = \frac{R_0}{2\pi}\log\left(\frac{r_2}{r_1}\right), \tag{72}$$

$$R_C = \frac{\vec{I}^T\check{R}\vec{I}}{I^2}, \tag{73}$$

$$R_L^{\mathrm{diss}} = \frac{\eta_L}{\pi e^2n_L^2}\left(\frac{1}{r_2^2} - \frac{1}{r_1^2}\right) \tag{74}$$

$$\begin{aligned} R_B^{\mathrm{diss}} &= \frac{\eta s_L^2}{\pi e^2n_L^2s_B^2}\left\{\frac{1}{r_1^2} - \frac{1}{r_2^2} + \frac{\ell_{\mathrm{GE}}}{2\ell_G^2}\right. \\ &\times \left.\frac{r_2\left[K_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) + I_0\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right] + r_1\left[I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) + K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_0\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right] - 2\ell_{\mathrm{GE}}}{r_1r_2\left(K_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)I_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right) - I_1\left(\frac{r_1}{\ell_{\mathrm{GE}}}\right)K_1\left(\frac{r_2}{\ell_{\mathrm{GE}}}\right)\right)}\right\}. \end{aligned} \tag{75}$$

Analysis of results

The behavior of the obtained resistance depends on the hierarchy of length scales r_1, r_2, $r_2 - r_1$, ℓ_{GE} and ℓ_R. In this Section, we specify the quantitative values of the parameters used to produce the plots shown in the main text. For clarity, here we restore the constants $\hbar$ and k_B.

We perform our quantitative analysis assuming the carrier density in the leads to be $n_L = 5 \times 10^{14}$ m^{-2}. The equilibrium temperature in the device (including both leads and the sample) is fixed to $T = 100$ K. The current, that is supplied by the source is $I = 1$ μA and we assume that the effective interaction constant is screened to $\alpha = 0.2$. We further use $\tau_{\text{dis}} = 1.25 \times 10^{-12}$ s and $\tau_L = 0.189 \times 10^{-12}$ s [65], since the density is higher in the leads. This determines all other parameters, except for τ_{RE} and τ_R (or alternatively ℓ_{GE} and ℓ_R). Since these quantities are difficult to extract from the available experimental data, we show results for several different regimes.

The time scales related to electron-electron interaction are given by [63]

$$\tau_{ii} = \hbar \frac{4\pi t_{ii} \log 2}{\alpha^2 k_B T}, \qquad t_{11} = \frac{1}{33.13}, \quad t_{22} = \frac{1}{5.45}. \tag{76}$$

For the above parameter values, we find $\tau_{11} = 0.5 \times 10^{-12}$ s and $\tau_{22} = 3 \times 10^{-12}$ s. The viscosity can be estimated as

$$\eta = \frac{0.446 k_B^2 T^2}{\alpha^2 v_g^2 \hbar} \tag{77}$$

and amounts to $\nu = v_g^2 \eta/(3P) = 0.25$ m^2/s. In addition

$$R_0 = \frac{\pi}{2\log 2} \frac{\hbar^2}{e^2 k_B T} \left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\text{dis}}} \right) = 1985.33\Omega \tag{78}$$

When describing the hydrodynamic velocity u_r and the pressure δP one can consider three different limits. If $\ell_{\text{GE}} \ll r_1, r_2$, which is achieved for very small τ_{dis}, one finds

$$u_r \approx \frac{I s_L \left(\sqrt{r r_2} \sinh\left(\frac{r - r_2}{\ell_{\text{GE}}}\right) - \sqrt{r r_1} \sinh\left(\frac{r - r_1}{\ell_{\text{GE}}}\right) \right)}{2\pi e n_L r s_B \sqrt{r_1 r_2} \sinh\left(\frac{r_1 - r_2}{\ell_{\text{GE}}}\right)} \tag{79}$$

which means, that the velocity vanishes exponentially close to the interface and is very small in the bulk of the sample. In the opposite limit $\ell_{\text{GE}} \gg r_1, r_2$ u_r shows a behavior similar to the drift velocity in the leads with logarithmic corrections

$$u_r \approx \frac{I s_L}{2\pi e n_L r s_B} + \frac{I s_L r_1^2 r_2^2 \log\left(\frac{r_1}{r_2}\right)}{4\pi e \ell_{\text{GE}}^2 n_L r s_B (r_1^2 - r_2^2)} + \frac{I r s_L \left[r_1^2 \log\left(\frac{r}{r_1}\right) - r_2^2 \log\left(\frac{r}{r_2}\right) \right]}{4\pi e \ell_{\text{GE}}^2 n_L s_B \, (r_1^2 - r_2^2)}. \tag{80}$$

Finally, if $r_2 - r_1 \ll r_1, r_2, \ell_{\text{GE}}$ we find the same $1/r$ behavior as in the leads

$$u_r \approx \frac{I s_L}{2\pi e n_L r s_B} \tag{81}$$

The resulting velocity u_r is shown in Fig. 2 of the main text. In the leads, the drift velocity shows a simple $1/r$ behavior, while one finds a jump due to the mismatch of entropy directly at the interface. Inside the sample, the situation depends on the relative size of ℓ_{GE}. If $\ell_{\text{GE}} \ll r_1, r_2$ we indeed observe, that the velocity decreases rapidly close to the interface and exactly vanishes in the bulk of the sample. This behavior is generally only observable in rather large samples, since the quantity τ_{dis} cannot be arbitrarily small while still staying in the hydrodynamic regime. In all other cases, u_r resembles a $1/r$ behavior, that is slightly modified by logarithmic corrections.

The plot of δT are shown in Fig. 3 of the main text. In the limit of $\ell_{\text{GE}} \ll r_1, r_2$ the non-equilibrium part of the temperature δT vanishes in the bulk of the sample. In this limit energy relaxation processes transfer any heating, that may develop in the sample to the substrate and out of the device. There is only a small finite effect very close to the interface. Since this is an effect of τ_R it is in principle independent of ℓ_{GE} and τ_{RE}, however we need $\ell_R < \ell_{\text{GE}}$ to remain in the hydrodynamic regime. In all other scenarios, there is a finite temperature profile, which may amount to 0.5% of the equilibrium temperature.

Finally we take a look at the total resistance R of the system. In general one might place the measuring points r_{in} and r_{out} very close to the interface, in which case the bulk resistance of the leads R_L would not contribute to the total resistance R. We will further disregard the influence of the phenomenological contact resistance R_C, which only depends on the used materials and their relative chemical potential. Then one can consider again three limiting cases of the hydrodynamic, dissipative contribution to the resistance R_B^{diss}. The first limit is $\ell_{\text{GE}} \ll r_1, r_2$ in which case we find

$$\begin{aligned} R_B^{\text{diss}} \approx & \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left(\frac{1}{r_1^2}-\frac{1}{r_2^2}\right) \\ & -\frac{s_L^2(A+\eta)\left((r_1+r_2)\cosh\left(\frac{r_1-r_2}{\ell_{\text{GE}}}\right)-2\sqrt{r_1 r_2}\right)\operatorname{csch}\left(\frac{r_1-r_2}{\ell_{\text{GE}}}\right)}{2\pi e^2 \ell_{\text{GE}} n_L^2 r_1 r_2 s_B^2} \qquad (82) \\ \approx & \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left[\frac{1}{r_1^2}-\frac{1}{r_2^2}-\frac{\sqrt{\ell_G^2+\frac{v_g^2\tau_{\text{dis}}\tau_{\text{RE}}}{2}}}{2\ell_G^2}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right] \qquad (83) \end{aligned}$$

where the second approximation requires $r_1 - r_2 \gg \ell_{\text{GE}}$. The result of Ref. [66] corresponds to neglecting the term proportional to ℓ_{GE}. The second limit is the case $\ell_{\text{GE}} \gg r_1, r_2$ and we find

$$R_B^{\text{diss}} \approx \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left(\frac{1}{r_1^2}-\frac{1}{r_2^2}+\frac{1}{2\ell_G^2}\log\left(\frac{r_2}{r_1}\right)\right), \qquad (84)$$

which introduces a logarithmic correction of exactly the same form as the bulk resistance R_B of the sample. The final limit is $r_2 - r_1 \ll r_1, r_2, \ell_{\text{GE}}$ where we find the result

$$R_B^{\text{diss}} = \frac{s_L^2 \eta (r_2^2 - r_1^2)}{4\pi e^2 n_L^2 r_1 r_2 s_B^2 \ell_G^2}. \qquad (85)$$

If one instead directly takes the limit $\tau_{\text{RE}} \to 0$, and additionally $\ell_G \ll r_1, r_2, r_2 - r_1$ one would obtain

$$R_B^{\text{diss}} \approx \frac{\eta s_L^2}{\pi e^2 n_L^2 s_B^2}\left[\frac{1}{r_1^2}-\frac{1}{r_2^2}-\frac{1}{2\ell_G}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right]. \qquad (86)$$

This is the result for the viscous correction to the resistance at charge neutrality in the setup of Ref. [67].

The plots for $\phi(r)$ and $R = R_B + R_B^{\text{diss}} + R_L^{\text{diss}}$ are shown in Fig. 4 and 5 of the main text respectively. In the case of the potential ϕ we find a logarithmic dependence on the radial position r in both the leads and the sample, where the overall prefactor is however different. In all considered cases, the jump at the interface is in the same direction, which for the second interface is opposite to what Ref. [66] obtains. This is due to the fact, that in our case the contribution of δP is larger than the contributions due to η and η_L alone. The jump is larger, for larger ℓ_{GE}. As seen in Fig. 5 of the main text, the total measured resistance is only slightly changed. The correction shown in the inset of Fig. 5 of the main text is nearly logarithmic for the larger ℓ_{GE}, while is saturates for the smaller ℓ_{GE}.

---

[1] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).

[2] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[3] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, et al., Science **351**, 1058 (2016).

[4] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).

[5] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[6] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. Krishna Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

[7] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[8] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[9] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).

[10] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[11] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, T. T. K. Watanabe, A. Lucas, A. F. Young, and A. C. Bleszynski Jayich (2020), arXiv:2002.05065.

[12] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).
[13] B. A. Braem, F. M. D. Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V. Koski, M. Berl, W. Dietsche, W. Wegscheider, et al., Phys. Rev. B **98**, 241304(R) (2018).
[14] A. Jaoui, B. Fauqué, C. W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D. L. Maslov, C. Felser, et al., npj Quantum Materials **3**, 64 (2018).
[15] M. S. Steinberg, Phys. Rev. **109**, 1486 (1958).
[16] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968), [Usp. Fiz. Nauk **94**, 689 (1968)].
[17] B. Bradlyn, M. Goldstein, and N. Read, Phys. Rev. B **86**, 245309 (2012).
[18] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).
[19] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).
[20] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).
[21] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).
[22] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).
[23] L. S. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).
[24] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).
[25] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).
[26] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).
[27] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).
[28] O. M. Corbino, Nuovo Cimento **1**, 397 (1911).
[29] A. Tomadin, G. Vignale, and M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).
[30] T. Holder, R. Queiroz, and A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).
[31] M. Shavit, A. V. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).
[32] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, et al. (2021), arXiv:2111.06412.
[33] A. Hui, V. Oganesyan, and E.-A. Kim, Phys. Rev. B **103**, 235152 (2021).
[34] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022).
[35] A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, and S. Ilani (2021), arXiv:2110.15369.
[36] O. E. Raichev (2022), arXiv:2202.06623.
[37] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).
[38] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).
[39] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).
[40] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).
[41] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).
[42] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).
[43] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).
[44] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).
[45] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018).
[46] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021).
[47] M. P. Marder, *Condensed Matter Physics* (Wiley, 2010).
[48] B. N. Narozhny and I. V. Gornyi, Frontiers in Physics **9**, 108 (2021).
[49] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).
[50] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **103**, 115402 (2021).
[51] Gall, V. and Narozhny, B. N. and Gornyi, I. V., unpublished.
[52] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021).
[53] M. Kumar, A. Laitinen, and P. Hakonen, Nature Communications **9**, 2776 (2018).
[54] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).
[55] G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).
[56] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1987).
[57] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).
[58] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).
[59] D. Halbertal, J. Cuppens, M. Ben Shalom, L. Embon, N. Shadmi, Y. Anahory, H. R. Naren, J. Sarkar, A. Uri, Y. Ronen, et al., Nature **539**, 407 (2016).
[60] D. Halbertal, M. Ben Shalom, A. Uri, K. Bagani, A. Y. Meltzer, I. Markus, Y. Myasoedov, J. Birkbeck, L. S. Levitov, A. K. Geim, et al., Science **358**, 1303 (2017).
[61] A. Aharon-Steinberg, A. Marguerite, D. J. Perello, K. Bagani, T. Holder, Y. Myasoedov, L. S. Levitov, A. K. Geim, and E. Zeldov, Nature **593**, 528 (2021).
[62] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, et al., Nat. Nanotechnol. **14**, 480 (2019).

[63] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021), URL `https://link.aps.org/doi/10.1103/PhysRevB.104.075443`.

[64] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017), URL `https://link.aps.org/doi/10.1103/PhysRevB.96.195401`.

[65] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021), URL `https://link.aps.org/doi/10.1103/PhysRevB.104.115432`.

[66] M. Shavit, A. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019), URL `https://link.aps.org/doi/10.1103/PhysRevLett.123.026801`.

[67] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022), URL `https://link.aps.org/doi/10.1103/PhysRevB.105.125302`.

Corbino magnetoresistance in neutral graphene

Vanessa Gall,[1,2] Boris N. Narozhny,[2,3] and Igor V. Gornyi[1,2,4]
[1] *Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2] *Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[3] *National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*
[4] *Ioffe Institute, 194021 St. Petersburg, Russia*

(Dated: May 30, 2023)

We explore the magnetohydrodynamics of Dirac fermions in neutral graphene in the Corbino geometry. Based on the fully consistent hydrodynamic description derived from a microscopic framework and taking into account all peculiarities of graphene-specific hydrodynamics, we report the results of a comprehensive study of the interplay of viscosity, disorder-induced scattering, recombination, energy relaxation, and interface-induced dissipation. In the clean limit, magnetoresistance of a Corbino sample is determined by viscosity. Hence the Corbino geometry could be used to measure the viscosity coefficient in neutral graphene.

Transport measurements remain one of the most common experimental tools in condensed matter physics. Having dramatically evolved past the original task of establishing bulk material characteristics such as electrical and thermal conductivities, modern experiments often involve samples that are tailor-made to target particular properties or behavior.

In recent years considerable efforts have been devoted to uncovering the collective or hydrodynamic flows of charge carriers in ultraclean materials as predicted theoretically [1–4]. Several dedicated experiments focused on answering two major questions: is the observed electronic flow really hydrodynamic and how to measure electronic viscosity [5–10], the quantity that fascinates physicists beyond the traditional condensed matter physics [11–18]. The hydrodynamic regime is apparently easiest to achieve in graphene [2–4]. This material is especially interesting since it can host two drastically different types of hydrodynamic behavior: (i) "conventional" at relatively high carrier densities [3, 19, 20] and (ii) "unconventional" at charge neutrality [21, 22].

Linearity of the excitation spectrum in graphene leads to the fact that electronic momentum density defines the energy current, $\boldsymbol{j}_E$. In the intermediate temperature window where electron-electron interaction is the dominant scattering process in the system ($\ell_{ee} \ll \ell_{\rm dis}, \ell_{e-ph}, W$, in the self-evident notation) the energy flow becomes hydrodynamic. At high carrier densities (in "doped graphene") the energy current is essentially equivalent to the electric current, $\boldsymbol{j}$, allowing one to formulate a Navier-Stokes-like equation for $\boldsymbol{j}$ [20] as pioneered by Gurzhi [19].

At charge neutrality and in the absence of the external magnetic field ($\boldsymbol{B} = 0$) the energy and electric currents decouple [23]. In the hydrodynamic regime the electric current remains Ohmic [22] (with the "internal" or "quantum" conductivity σ_Q due to electron-electron interaction [24–27]), while the Navier-Stokes-like equation describes the energy current [22, 28, 29]. If external magnetic field is applied, the energy and charge flows become entangled [21–23] allowing for a possibility to detect the hydrodynamic flow in electronic transport experiments. In particular, a bulk (infinite) system is characterized by positive, parabolic magnetoresistance [23, 30] proportional to the disorder mean free time $\tau_{\rm dis}$ (disorder scattering is the only mechanism of momentum relaxation).

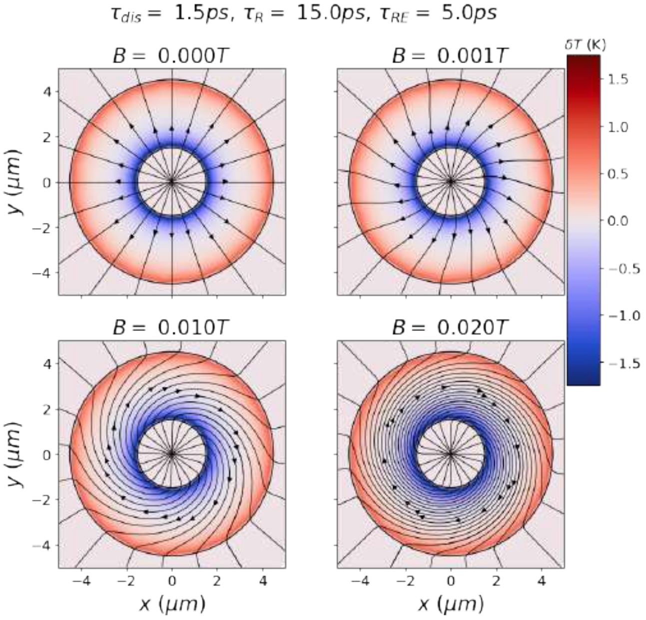


FIG. 1. Hydrodynamic velocity $\boldsymbol{u}$ and temperature δT distribution in the device obtained by solving the hydrodynamic equations at relatively high temperatures where energy relaxation is dominated by supercollisions. Arrows indicate $\boldsymbol{u}$ and the color map shows δT. The quantitative results were computed using the following values of the average temperature $T = 150\,$K, disorder scattering time $\tau_{\rm dis} = 1.5\,$ps (corresponding to the scattering rate $\tau_{\rm dis}^{-1} \approx 0.67\,THz\,\approx 5.1\,$K), recombination time $\tau_R = 15\,$ps, energy relaxation time $\tau_{RE} = 5\,$ps, dimensionless coupling constant in graphene $\alpha = 0.5$, carrier density in the leads $n_L = 5\times10^{12}\,{\rm cm}^{-2}$, and the current passing through the device $I = 1\,\mu$A. The four panels correspond to the indicated values of magnetic field.

The outcome of a given measurement is strongly influenced by the sample size and geometry. Early experiments focused on either the "strip" (or Hall bar) [5–8] or the point contact geometry [9, 10], while more recently data on Corbino disks became available [31].

The simplest viscous phenomenon one can look for in a long (striplike) sample [7–10, 12, 32–52] is the Poiseuille flow [53–55]. This flow is characterized by a parabolic velocity profile with the curvature determined by viscosity. In doped graphene the Poiseuille flow of charge can be detected by imaging the electric current density [8]. In contrast, neutral graphene exhibits the Poiseuille flow of the energy current [56]. Moreover, at relatively high temperatures where hydrodynamic behavior in graphene is observed the electron-phonon interaction (either direct [23, 57, 58] or via "supercollisions" [59–64]) cannot be neglected and hence electronic energy is not conserved. The resulting energy relaxation dwarfs the viscous contribution to the Navier-Stokes [65] equation.

Applying a perpendicular magnetic field to a neutral graphene strip leads to a coupled charge and energy flow with the two currents being orthogonal [23]. The electric current flowing along the strip is accompanied by a neutral quasiparticle flow in the lateral direction resulting in energy and quasiparticle accumulation near the strip boundaries [66, 67]. The accumulation is limited by quasiparticle recombination [67] and energy relaxation processes [59]. As a result, the boundary region's contribution to the resistance is linear in the applied magnetic field [23, 48, 67, 68], in contrast to the standard quadratic magnetoresistance of the bulk system [23, 30]. In classically strong fields the boundary contribution dominates making the linear magnetoresistance directly observable. This effect is not specific to Dirac fermions as shown by experiments in bilayer graphene [69].

The Corbino geometry presents an interesting alternative to the Hall bar experiments [31, 65, 70–78]. In a typical measurement the electric current is passed from the inner to the outer boundary of a Corbino disk. The specific feature of the stationary flow in this geometry is that the magnitude of the radial component of the current is determined by the continuity equation alone. In the absence of the magnetic field the whole current flows radially. Combining the solution of the continuity equation with the hydrodynamic Gurzhi equation (e.g., in doped graphene) leads to an apparent paradox [73]: the current flow appears unaffected by viscosity. However, the dissipated energy is still determined by viscosity leading to the jumps of electric potential at the contacts thus resolving the paradox. In a perpendicular magnetic field the system exhibits parabolic magnetoresistance inverse proportional to the viscosity and independent of the disorder scattering. Applied phenomenologically to neutral graphene (neglecting contact effects) [78] this conclusion stands in sharp contrast to the standard result [23, 30] raising the question of the fate of the disorder-limited bulk magnetoresistance in the Corbino geometry.

In this paper we investigate hydrodynamic flows in neutral graphene in the Corbino disk subjected to the perpendicular magnetic field based on the graphene-specific hydrodynamic theory [2, 22, 59] reporting the results of a careful study of the interplay of viscosity, disorder-induced scattering, recombination, energy relaxation, and interface-induced dissipation. Solving the hydrodynamic equations we find the spatial distribution of the hydrodynamic velocity $\boldsymbol{u}$, temperature (see Fig. 1), electric current, and potential φ (see Fig. 2). Furthermore, we calculate the field-dependent resistance of the whole Corbino sample including the leads. Keeping in mind recent and ongoing experiments, it appears logical to include the effect of the lead resistance in order to achieve a more realistic description of the Corbino device. However, the theoretical limit of "ideal" leads can be considered without any complications.

The main results of this paper are as follows. We show that magnetoresistance of the Corbino device exhibits a crossover from the "hydrodynamic" (viscosity-dominated) to the "bulk" (disorder-limited) behavior with the increasing system size as compared to the Gurzhi length $\ell_G = \sqrt{\nu\tau_{\rm dis}}$ [46–49, 52] (ν is the kinematic viscosity [3, 5, 6, 55, 79] and $\tau_{\rm dis}$ is the disorder mean free time). In the clean limit ($\tau_{\rm dis} \to \infty$) magnetoresistance remains finite and is determined by viscosity offering a way to measure the viscosity coefficient in neutral graphene. In classically strong fields magnetoresistance remain parabolic (in contrast to the linear magnetoresistance in the strip geometry). The "contact magnetoresistance" induced through the dissipation jump is present, but is typically weaker than the bulk contribution.

I. MAGNETOHYDRODYNAMICS IN GRAPHENE

Our arguments are based on the hydrodynamic theory of electronic transport in neutral graphene derived from the kinetic (Boltzmann) equation [21, 22, 59] or from the microscopic Keldysh technique [80]. At charge neutrality both bands contribute to transport on equal footing. A current-carrying state is characterized by the chemical potentials $\mu_\pm$ of each band or by their linear combinations [22, 81]

$$\mu = \frac{\mu_+ + \mu_-}{2}, \qquad \mu_I = \frac{\mu_+ - \mu_-}{2}, \tag{1a}$$

conjugate to the "charge" and "imbalance" (or "total quasiparticle") densities

$$n = n_+ - n_-, \qquad n_I = n_+ + n_-. \tag{1b}$$

In equilibrium $\mu_I = 0$. Any macroscopic current can be expressed as a product of the corresponding density and hydrodynamic velocity $\boldsymbol{u}$ (up to dissipative corrections). Due to the kinematic peculiarity of the Dirac fermions in graphene known as the "collinear scattering singularity" [21, 25] one has to consider the electric, energy, and imbalance, $\boldsymbol{j}_I$ currents defined as

$$\boldsymbol{j} = n\boldsymbol{u} + \delta\boldsymbol{j}, \quad \boldsymbol{j}_I = n_I\boldsymbol{u} + \delta\boldsymbol{j}_I, \quad \boldsymbol{j}_E = \mathcal{W}\boldsymbol{u}, \tag{2}$$

where $\mathcal{W}$ is the enthalpy density and $\delta\boldsymbol{j}$ and $\delta\boldsymbol{j}_I$ are the dissipative corrections. In the degenerate limit $\mu \gg T$ the dissipative corrections vanish [22, 28] justifying the applicability of the single-band picture to doped graphene. At charge neutrality $n = 0$, the electric and energy currents in Eq. (2) appear to be decoupled [22].

Within linear response, steady-state macroscopic currents obey the linearized hydrodynamic equations [82]. Assuming that the dominant mechanism of energy relaxation is supercollisions [59], the equations have the form

$$\boldsymbol{\nabla}\cdot\delta\boldsymbol{j} = 0, \tag{3a}$$

$$n_I\boldsymbol{\nabla}\cdot\boldsymbol{u} + \boldsymbol{\nabla}\cdot\delta\boldsymbol{j}_I = -\frac{12\ln 2}{\pi^2}\frac{n_I\mu_I}{T\tau_R}, \tag{3b}$$

$$\boldsymbol{\nabla}\delta P = \eta\Delta\boldsymbol{u} + \frac{e}{c}\delta\boldsymbol{j}\times\boldsymbol{B} - \frac{3P\boldsymbol{u}}{v_g^2\tau_{\rm dis}}, \tag{3c}$$

$$3P\boldsymbol{\nabla}\cdot\boldsymbol{u} = -\frac{2\delta P}{\tau_{RE}}. \tag{3d}$$

Here Eq. (3a) is the continuity equation; Eq. (3b) is the "imbalance" continuity equation [22, 81] (where v_g is the band velocity in graphene, c is the speed of light, e is the unit charge, and τ_R is the recombination time); Eq. (3c) is the linearized Navier-Stokes equation [22, 29, 82, 83] (with η being the shear viscosity); and Eq. (3d) is the linearized "thermal transport" equation (τ_{RE} is the energy relaxation time [59]). We follow the standard approach [55] where the thermodynamic quantities are replaced by the corresponding equilibrium functions of the hydrodynamic variables. Equilibrium thermodynamic quantities, i.e., the pressure $P = 3\zeta(3)T^3/(\pi v_g^2)$, enthalpy density $\mathcal{W}$, imbalance density, $n_I = \pi T^2/(3v_g^2)$, and energy density are related by the "equation of state", $\mathcal{W} = 3P = 3n_E/2$. Equations (3) should be solved for the unknowns $\boldsymbol{u}$, μ_I, and δP keeping the remaining (thermodynamic) quantities, e.g., n_I, P, and T, constant.

The dissipative corrections to the macroscopic currents can be determined from the underlying microscopic theory [22, 29, 82] and are expressed in terms of the same variables closing the set of hydrodynamic equations (3)

$$\delta\boldsymbol{j} = \frac{1}{e^2\tilde{R}}\left[e\boldsymbol{E} + \omega_B\boldsymbol{e}_B\times\left(\frac{\alpha_1\delta_I\boldsymbol{\nabla}\mu_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}} - \frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right)\right], \tag{4a}$$

$$\delta\boldsymbol{j}_I = -\frac{\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{1}{e^2\tilde{R}}\times \tag{4b}$$

$$\times\left[\alpha_1\omega_B\boldsymbol{e}_B\times\boldsymbol{E} + \frac{2T\ln 2}{\pi}e^2R_0\boldsymbol{\nabla}\mu_I + \alpha_1\omega_B^2\frac{2T\ln 2}{v_g^2}\boldsymbol{u}\right],$$

$$\tilde{R} = R_0 + \alpha_1^2\delta_I\tilde{R}_B. \tag{4c}$$

In Eqs. (4) the following notations are introduced. R_0 is the zero-field bulk resistivity in neutral graphene [23, 30]

$$R_0 = \frac{\pi}{2e^2T\ln 2}\left(\frac{1}{\tau_{11}} + \frac{1}{\tau_{\rm dis}}\right) \xrightarrow[\tau_{\rm dis}\to\infty]{} \frac{1}{\sigma_Q}, \tag{5}$$

where $\tau_{11} \propto \alpha_g^{-2}T^{-1}$ describes the appropriate electron-electron collision integral. $\tilde{R}_B$ denotes [65, 82]

$$\tilde{R}_B = \frac{\pi}{2e^2T\ln 2}\frac{\omega_B^2}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}, \tag{6}$$

where $\tau_{22} \propto \alpha_g^{-2}T^{-1}$ describes a component of the collision integral that is qualitatively similar, but quantitatively distinct from τ_{11} and $\delta_I \approx 0.28$. Another numerical factor in Eqs. (4) is $\alpha_1 \approx 2.08$ and $\omega_B = eBv_g^2/(2cT\ln 2)$ is the generalized cyclotron frequency at $\mu = 0$.

The shear viscosity at charge neutrality and in the absence of magnetic field was evaluated in Refs. [22, 79, 83] and has the form

$$\eta(\mu = 0, B = 0) = \mathcal{B}\frac{T^2}{\alpha_g^2v_g^2}, \qquad \mathcal{B} \approx 0.45. \tag{7}$$

Within the renormalization group (RG) approach, α_g is a running coupling constant [56, 83–86]. However, the product $\alpha_g v_g$ remains constant along the RG flow [24, 83]. Hence Eq. (7) gives the correct form of shear viscosity in neutral graphene [84]. Within the kinetic theory approach, the coefficient $\mathcal{B}$ can be expressed in terms of time scales characterizing the collision integral [22, 79]. At neutrality these time scales are qualitatively similar to, but quantitatively distinct from τ_{11} and τ_{22}. The similarity follows from the fact that in general all time scales are functions of the chemical potential and temperature [22, 28, 87]. At neutrality $\mu = 0$ and hence all time scales are inverse proportional to temperature.

As a function of the magnetic field, the viscosity coefficient in neutral graphene exhibit a weak decay until eventually saturating in classically strong fields [79]

$$\eta(\mu = 0, B) = \frac{\mathcal{B}+\mathcal{B}_1\gamma_B^2}{1+\mathcal{B}_2\gamma_B^2}\frac{T^2}{\alpha_g^2v_g^2}, \qquad \gamma_B = \frac{|e|v_g^2B}{\alpha_g^2cT^2}, \tag{8}$$

where

$$\mathcal{B}_1 \approx 0.0037, \qquad \mathcal{B}_2 \approx 0.0274.$$

This behavior should be contrasted with the more conventional Lorentzian decay of field-dependent shear viscosity in doped graphene [6, 45, 46, 79, 88]. However, in weak fields where most present-day experiments are performed this distinction is negligible. Moreover, due to the smallness of the coefficient $\mathcal{B}_1$ and $\mathcal{B}_2$ we disregard the field dependence of η in what follows.

Under the assumptions of the hydrodynamic regime, disorder scattering is characterized by the large mean free time, $\tau_{\rm dis} \gg \tau_{11}, \tau_{22}$, yielding a negligible contribution to Eqs. (5) and (6). Equation (5) describes the uniform bulk current (at $\boldsymbol{B} = 0$) and is independent of viscosity (i.e., in a channel [3, 21, 67, 82]). In contrast, in the Corbino geometry the current flow is necessarily inhomogeneous and hence viscous dissipation must be taken into account.

II. BOUNDARY CONDITIONS

Differential equations (3) should be supplemented by boundary conditions, which should be expressed in terms of the hydrodynamic velocity and macroscopic currents. The statement of the boundary conditions does not imply the validity of the hydrodynamic approximation at the sample edges and generally have to be derived from the underlying microscopic theory. However some of the boundary conditions can be derived based on the conservation laws alone. In the circular Corbino geometry conservation laws can be used to establish boundary conditions for radial components of the currents [65].

A. Radial components of macroscopic currents

A typical experimental setup involves a graphene sample (in our case, at charge neutrality) in the shape of an annulus placed between the inner (a disk of radius r_1) and outer (a ring with the inner radius r_2) metallic contacts (leads). The electric current I is injected into the center of the inner lead preserving the rotational invariance (e.g., through a thin vertical wire attached to the center point) and spreads towards the outer lead, which for concreteness we assume to be grounded. The overall voltage drop U is measured between two points in the two leads (at the radii $r_{\rm in} < r_1$ and $r_{\rm out} > r_2$) yielding the device resistance, $R = U/I$. The only boundaries in the system are between the sample and the external leads.

For simplicity, we assume both leads to be of the same material with a single-band electronic system, e.g., highly doped graphene with the same doping level. In that case, all macroscopic currents in the leads are proportional to the drift velocity and hence are determined by the injected current. In the stationary case, the continuity equation (3a) determines the radial component of the electric current density. In the inner lead this yields $j_r^{\rm in} = I/(2\pi e r)$, defining the radial component of the drift velocity, $u_r^{\rm in} = j_r^{\rm in}/n_L$ (n_L is the carrier density in the inner lead). Assuming charge conservation is not violated at the interface, we find the boundary condition between the inner lead and the sample

$$j_r(r_1-\epsilon) = n_L u_r(r_1-\epsilon) = \delta j_r(r_1+\epsilon), \tag{9a}$$

where $\epsilon > 0$ is infinitesimal and we took into account that in neutral ($n = 0$) graphene $\boldsymbol{j} = \delta\boldsymbol{j}$.

The second hydrodynamic equation, Eq. (3b), is the continuity equation for the imbalance density. Although the total quasiparticle number is not conserved, integrating this equation over an infinitesimally thin region encompassing the boundary yields a similar boundary condition for the imbalance current assuming that the relaxation rate due to quasiparticle recombination is not singular at the boundary

$$j_{I,r}(r_1-\epsilon) = n_L u_r(r_1-\epsilon) = n_I u_r(r_1+\epsilon) + \delta j_{I,r}(r_1+\epsilon). \tag{9b}$$

Here we took into account the fact that in a single-band system $\boldsymbol{j}_I$ is identical with $\boldsymbol{j}$.

Finally, Eq. (3d) is the linearized continuity equation for the entropy density (here we follow the standard practice [55] of replacing the continuity equation for the energy density by the entropy flow equation, also known as the thermal transport equation). Again, assuming the energy relaxation rate is not singular at the interface (i.e., the current flow is not accompanied by energy or excess heat accumulation at the boundary between the sample and the contact) we integrate Eq. (3d) over an infinitesimally thin region encompassing the boundary and arrive at the boundary condition for the entropy current

$$s^{\rm in} u_r(r_1-\epsilon) = s u_r(r_1+\epsilon), \tag{9c}$$

where s and $s^{\rm in}$ are the entropy densities in the sample and inner lead, respectively.

B. Tangential flows in external magnetic field

The above boundary conditions (and the corresponding conditions on the outer lead) are sufficient to solve the hydrodynamic equations in the absence of magnetic field where all currents are radial [65]. An external magnetic field induces the tangential components of the currents due to the classical Hall effect. The continuity equations do not determine the tangential components and hence the boundary conditions have to be derived from a microscopic theory. Generally speaking, the boundary conditions depend on the presence of tangential forces at the interface, usually associated with edge roughness. Typically [2–4, 55, 73], one considers the two limiting cases of either the "no-slip" or "no-stress" boundary conditions corresponding to either the presence or the absence of the drag-like friction across the interface.

For contact interfaces in the Corbino geometry, the boundary conditions corresponding to the above limiting cases differ from the well-known expression of conventional hydrodynamics. The no-slip boundary condition now means that the tangential component of the hydrodynamic velocity is continuous across the interface (written as above for the inner interface)

$$u_{L\vartheta}(r_1-\epsilon) = u_\vartheta(r_1+\epsilon), \tag{10a}$$

in contrast to the common condition of vanishing velocity at the channel boundary (the two are consistent, since in the latter case there is no flow beyond the edge).

The no-stress boundary condition means the absence of any forces along the interface in which case the tangential component of the stress tensor Π^{ij} is continuous. In polar coordinates appropriate for the Corbino geometry one finds

$$\Pi^{\vartheta r}_{L,E}(r_1-\epsilon) = \Pi^{\vartheta r}_{E}(r_1+\epsilon), \tag{10b}$$

The no-stress boundary condition is easy to derive starting from the kinetic equation. Multiplying the kinetic equation by the momentum and summing over all quasiparticle states, one finds an equation featuring the gradient of the stress tensor [22] as well as macroscopic forces in the system. Now the boundary condition can be obtained by integrating that equation over the small volume around the interface. Unless there is a force localized at the interface (with a δ-function-like coordinate dependence on the hydrodynamic scale), this procedure would yield Eq. (10b). Usually, the interfaces are microscopically rough with the roughness providing such a force. As a result, the no-slip boundary condition is more commonly used. In neutral graphene, however, the quasiparticle wavelength typically exceeds any length scale associated with edge roughness leading to specular scattering [82] and Eq. (10b).

In the case of the hard wall edges, the boundary conditions were previously studied theoretically in Ref. [89] and confirmed experimentally in Ref. [8] where a nonzero slip length was proposed indicating a more general Maxwell's boundary condition. However, the specific choice of the boudnary conditions does not lead to qualitatively different results [73]. Here we follow the hydrodynamic tradition and consider both the no-slip and no-stress boundary conditions.

C. Interface-induced dissipation and jumps of the electric potential

The hydrodynamic theory discussed so far completely describes the energy flow in neutral graphene. In order to establish the device resistance R we have to find the behavior of the electrochemical potential at the interfaces.

The standard description of interfaces between metals or semiconductors in terms of the contact resistance [90] can be carried over to neutral graphene [81]. In graphene, the contact resistance was recently measured in Ref. [8] (see also Refs. [31, 91, 92]). In the diffusive (or Ohmic) case, the contact resistance leads to a voltage drop that is small compared to that in the bulk of the sample and can be neglected. In contrast, in the ballistic case with almost no voltage drop in the bulk, most energy is dissipated at the contacts. Both scenarios neglect interactions.

In the diffusive regime interactions give rise to perturbative corrections to the bulk resistivity [93, 94] and the contact resistance can still be neglected. In ballistic samples electron-electron interaction may lead to the "Knudsen-Poiseuille" crossover [19] and drive the system to the hydrodynamic regime. In this case the Ohmic resistivity of the electronic fluid may remain small, but there exist other channels for dissipation due to viscosity [73] and energy relaxation processes [59]. In neutral graphene the effect is subtle [65], since the electric current is decoupled from the hydrodynamic energy flow. However, both are induced by the current source that provides the energy dissipated through all the above channels. The energy dissipated in the system corresponds to the overall voltage drop. In the bulk of the sample the voltage drop is Ohmic as determined by Eq. (4a), while the additional contribution takes the form of a potential jump at the interface between the sample and leads. At the same time, an excess electric field is induced in a thin Knudsen layer around the interface [73].

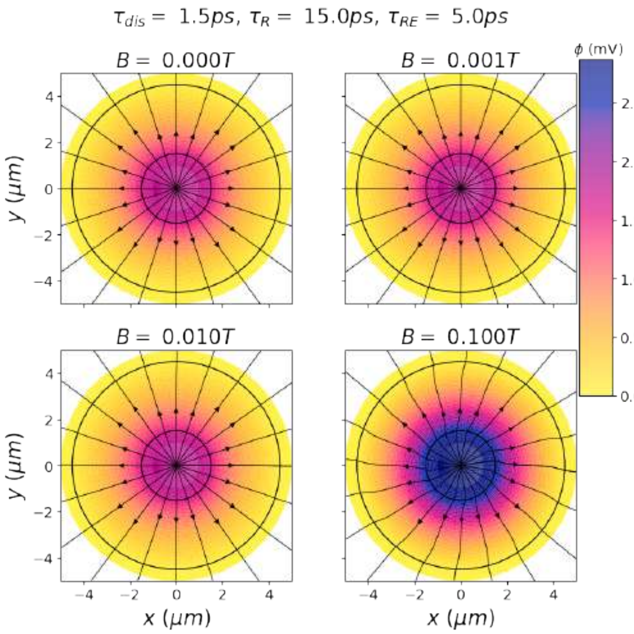


FIG. 2. Electric current density $\boldsymbol{j}$ and potential φ in the device obtained by solving the hydrodynamic equations at relatively high temperatures where energy relaxation is dominated by supercollisions. Arrows indicate $\boldsymbol{j}$ and the color map shows φ. The outer lead is chosen to be grounded. The four panels correspond to the indicated values of magnetic field. For the values of other parameters, see Fig 1.

The magnitude of the jump in ϕ can be established by considering the flow of energy through the interface as suggested in Ref. [73] and detailed in neutral graphene at $\boldsymbol{B} = 0$ in Ref. [65]. Consider the kinetic energy defined by integrating the energy density $n_E(\boldsymbol{u}) - n_E(0)$ over the volume

$$\mathcal{E} = \int dV\, [n_E(\boldsymbol{u}) - n_E(0)] \approx \int dV \frac{6P}{v_g^2} u^2, \tag{11}$$

which we have expanded to the leading order in $\boldsymbol{u}$ (and hence in I). In the stationary state, dissipation is balanced by the work done by the source, such that the time derivative of the kinetic energy vanishes, $\mathcal{A} = \dot{\mathcal{E}} = 0$. Using the equations of motion and continuity equations to find time derivatives, one may split $\mathcal{A}$ into the "bulk" and "boundary" contributions, $\mathcal{A} = \mathcal{A}_{\rm bulk} + \mathcal{A}_{\rm edge}$. The former may be interpreted as the bulk dissipation, while $\mathcal{A}_{\rm edge}$ must include the energy brought in (carried away) through the boundary by the incoming (outgoing) flow. The boundary condition is then found under the assumption that energy is not accumulated at the interface.

Assuming the leads' material is highly doped graphene, the equation of motion is the usual Ohm's law where we may combine the diffusion term [95] with a contribution of viscosity η_L due to disorder [96] into the gradient of the stress-energy tensor [23] and hence

$$\begin{aligned}
&\frac{3P_L}{v_g^2}\boldsymbol{u}_L\partial_t\boldsymbol{u}_L = \\
&\quad= u_L^i\left(-\frac{3P_L}{v_g^2}\frac{u_L^i}{\tau_L}-\nabla^j\Pi^{ij}_{L,E}+n_L eE^i+\frac{e}{c}\epsilon^{ijk}j^jB^k\right) \\
&\quad= -\frac{3P_L}{v_g^2}\frac{u_L^2}{\tau_L}+\frac{\partial u_{L,i}}{\partial x_j}\Pi^{ij}_{L,E}+\frac{e}{c}\boldsymbol{u}_L\cdot(\boldsymbol{j}\times\boldsymbol{B})+e\varphi\boldsymbol{\nabla}\cdot\boldsymbol{j} \\
&\qquad -\nabla^i\left(u_L^j\Pi^{ij}_{L,E}+ej^i\varphi\right).
\end{aligned}$$

The last term in this expression determines the boundary contribution. Given that the Lorentz force does not explicitly contribute, the only difference from the expression derived in Ref. [65] at $\boldsymbol{B}=0$ is the nonzero tangential components of the hydrodynamic velocity and the stress tensor (vanishing in the absence of magnetic field). In neutral graphene, we obtain a similar expression from the Navier-Stokes equation, while the Joule heat is determined by $\delta\boldsymbol{j}$. Equating the two contributions we find the jump of the potential in the form

$$\begin{aligned}
&\varphi(r_1-\varepsilon)-\varphi(r_1+\varepsilon) = IR_c + \\
&\quad +\frac{2\pi r_1}{I}\Bigg[\left(u_r\Pi^{rr}_E+u_\vartheta\Pi^{\vartheta r}_E\right)\Big|_{r_1+\varepsilon} \\
&\qquad -\left(u_r\Pi^{rr}_{L,E}+u_\vartheta\Pi^{\vartheta r}_{L,E}\right)\Big|_{r_1-\varepsilon}\Bigg],
\end{aligned} \tag{12}$$

where R_c is the usual contact resistance [81]. A similar condition holds at the boundary with the outer lead.

III. HYDRODYNAMIC FLOWS IN THE CORBINO GEOMETRY

In polar coordinates and taking into account radial symmetry, the hydrodynamic equations (3) and (4) form two disjoint sets of differential equations. The first one determines the tangential component of the velocity u_ϑ:

$$\frac{1}{r}\frac{\partial(r\delta j_r)}{\partial r}=0, \tag{13a}$$

$$\eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_\vartheta)}{\partial r}\right)-\frac{eB}{c}\delta j_r-\frac{3Pu_\vartheta}{v_g^2\tau_{\rm dis}}=0, \tag{13b}$$

$$\delta j_r=\frac{1}{e^2\tilde{R}}\left[eE_r(r)+\omega_B\frac{2T\ln 2}{v_g^2}u_\vartheta\right], \tag{13c}$$

$$\delta j_{I\vartheta}=-\frac{\alpha_1\delta_I\omega_B}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\delta j_r, \tag{13d}$$

while the second one involves the radial component u_r:

$$\frac{n_I}{r}\frac{\partial(ru_r)}{\partial r}+\frac{1}{r}\frac{\partial(r\delta j_{Ir})}{\partial r}=-\frac{12\ln 2}{\pi^2}\frac{n_I\mu_I(r)}{T\tau_R}, \tag{14a}$$

$$\frac{\partial\delta P}{\partial r}=\eta\partial_r\left(\frac{1}{r}\frac{\partial(ru_r)}{\partial r}\right)+\frac{eB}{c}\delta j_\vartheta-\frac{3Pu_r}{v_g^2\tau_{\rm dis}}, \tag{14b}$$

$$\frac{3P}{r}\frac{\partial(ru_r)}{\partial r}=-\frac{2\delta P(r)}{\tau_{RE}}. \tag{14c}$$

$$\delta j_\vartheta=\frac{\omega_B}{e^2\tilde{R}}\left(\frac{\alpha_1\delta_I}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\frac{\partial\mu_I}{\partial r}-\frac{2T\ln 2}{v_g^2}u_r\right), \tag{14d}$$

$$\delta j_{Ir}=-\frac{2\delta_I T\ln 2}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}\left[\frac{R_0}{\pi\tilde{R}}\frac{\partial\mu_I}{\partial r}+\frac{\alpha_1\omega_B^2}{e^2\tilde{R}}\frac{u_r}{v_g^2}\right]. \tag{14e}$$

A. Tangential component of the velocity and bulk voltage drop

The bulk magnetoresistance can be found by solving Eqs. (13) with the appropriate boundary conditions. Combining Eqs. (13a) and (13b) we find an inhomogeneous Bessel equation for the tangential component of the velocity u_ϑ with the characteristic length scale being the Gurzhi length $\ell_G^2=\eta v_g^2\tau_{\rm dis}/(3P)$. The boundary condition for u_ϑ is determined by microscopic details of viscous drag at the interface and hence is not universal. Here we follow the hydrodynamic tradition and consider both the no-slip and the no-stress boundary conditions, see Sec. II B. Moreover, one can distinguish two different setups where the external magnetic field is applied either to the sample only or to the whole device including the leads. In all these cases we can find an analytic expression for u_ϑ, which can be substituted into of Eq. (13c) to find the electric field in the sample, E_r (the radial component of the current is determined by the continuity equation alone). Similarly, Eq. (13d) determines $\delta j_{I\vartheta}$. Using the obtained electric field we can determine the voltage drop through the bulk of the sample as

$$U=\int\limits_{r_1}^{r_2}E_r dr=\int\limits_{r_1}^{r_2}dr\left(\frac{\tilde{R}I}{2\pi r}-\frac{B}{c}u_\vartheta\right). \tag{15}$$

For the no-slip boundary condition for u_ϑ and allowing the external magnetic field to penetrate the leads, the tangential component of the velocity is given by

$$\begin{aligned}
u_\vartheta = &-\frac{BI\ell_G^2}{2\pi c\eta r}+\frac{BI\left(\eta\ell_L^2-\eta_L\ell_G^2\right)}{2\pi c\eta\eta_L r_1 r_2}\times \\
&\times\left[K_1\left(\frac{r}{\ell_G}\right)\frac{r_1 I_1\left(\frac{r_1}{\ell_G}\right)-r_2 I_1\left(\frac{r_2}{\ell_G}\right)}{K_1\left(\frac{r_1}{\ell_G}\right)I_1\left(\frac{r_2}{\ell_G}\right)-I_1\left(\frac{r_1}{\ell_G}\right)K_1\left(\frac{r_2}{\ell_G}\right)}\right. \\
&\left.+I_1\left(\frac{r}{\ell_G}\right)\frac{r_2 K_1\left(\frac{r_2}{\ell_G}\right)-r_1 K_1\left(\frac{r_1}{\ell_G}\right)}{K_1\left(\frac{r_1}{\ell_G}\right)I_1\left(\frac{r_2}{\ell_G}\right)-I_1\left(\frac{r_1}{\ell_G}\right)K_1\left(\frac{r_2}{\ell_G}\right)}\right],
\end{aligned} \tag{16}$$

where η_L is the disorder-induced viscosity [96] and $\ell_L^2 = v_g^2\eta_L\tau_L/(2P_L)$ is the Gurzhi length in the leads.

In the limit $\ell_G \gg r_1, r_2$ (i.e., “clean system” with long mean free time $\tau_{\rm dis} \to \infty$) this simplifies to ($p = r_2/r_1$)

$$u_\vartheta \approx -\frac{BI\ell_L^2}{4\pi c r\eta_L}\left[2+\left(\frac{1}{\ell_G^2}-\frac{\eta_L}{\eta\ell_L^2}\right)\times\right. \tag{17}$$

$$\left.\times\,\frac{r^2\ln\left(\frac{r}{r_1}\right)+r^2p^2\ln\left(\frac{r_2}{r}\right)-r_2^2\ln p}{1-p^2}\right].$$

The corresponding voltage drop remains finite

$$U \approx \left(1-\frac{\eta\ell_L^2}{\eta_L\ell_G^2}\right)\frac{B^2Ir_2^2}{4\pi c^2\eta}\frac{(p^2-1)^2-4p^2\ln^2 p}{4p^2(p^2-1)} \tag{18a}$$

$$+\frac{I\ln p}{2\pi}\left(\frac{B^2}{c^2}\frac{v_g^2\tau_L}{3P_L}+\tilde{R}\right),$$

yielding the field-dependent bulk resistance ($R = U/I$)

$$R(B) \approx \frac{\ln p}{2\pi}R_0+\frac{B^2r_2^2}{4\pi c^2\eta}\frac{(p^2-1)^2-4p^2\ln^2 p}{4p^2(p^2-1)} \tag{18b}$$

$$+\frac{B^2v_g^4\ln p}{2c^2T^3}\left[\frac{\alpha_1^2\delta_I}{8\ln^3 2}\frac{1}{\tau_{\rm dis}^{-1}+\delta_I^{-1}\tau_{22}^{-1}}+\frac{T^3}{\mu^3}\tau_L\right],$$

assuming $\eta\ell_L^2/(\eta_L\ell_G^2) = 3P\tau_L/(2P_L\tau_{\rm dis}) \ll 1$ with $P_L = \mu_L^3/(3\pi v_g^2)$. The two field-dependent terms differ in their dependence on temperature, sample size, and coupling constant [35] opening a possibility to separate the two contributions from the experimental data and thus to measure the viscosity coefficient.

If the magnetic field is applied to the sample only (and not to the leads) u_ϑ vanishes in the leads and hence the terms with ℓ_L do not appear in the voltage drop (18). In that case, the field-dependent contribution to U does not contain $\tau_{\rm dis}$ in contrast to the known result in the strip geometry [23, 30].

A similar result can be obtained in the case of no-stress boundary conditions, where the tangential component of the velocity u_ϑ is still expressed in terms of the Bessel functions. In the clean limit ($\ell_G \gg r_1, r_2$) the voltage drop also remains finite

$$U \approx \frac{I}{2\pi}\left(\tilde{R}+\frac{B^2\ell_L^2}{c^2\eta}-\frac{B\eta_H}{ec\eta n_L}\right)\ln p \tag{19}$$

$$+\frac{r_2^2B^2I}{4\pi c^2\eta}\left[\frac{\left(p^2-1\right)\left(p^4+10p^2+1\right)}{12p^2\left(p^2+1\right)^2}-\frac{\ln p}{1+p^2}\right]$$

$$+\frac{I}{2\pi}\left[\frac{B^2}{c^2}\frac{\left(\ell_G^2-\ell_L^2\right)}{\eta}+\frac{B\eta_H}{ec\eta n_L}\right]\frac{p^2-1}{p^2+1},$$

where η_H is the Hall viscosity in the leads, which vanishes if the magnetic field is not allowed in the leads. In that case, the last term in the voltage drop (19) is proportional to $\tau_{\rm dis}$ and independent of viscosity. The second term in Eq. (19) remains similar to Eq. (18) and is inverse proportional to η. This term's dependence on the ratio p is distinct from both Eq. (18) and the third term in Eq. (19) and could be extracted by analyzing the data in a set of Corbino disks with different p.

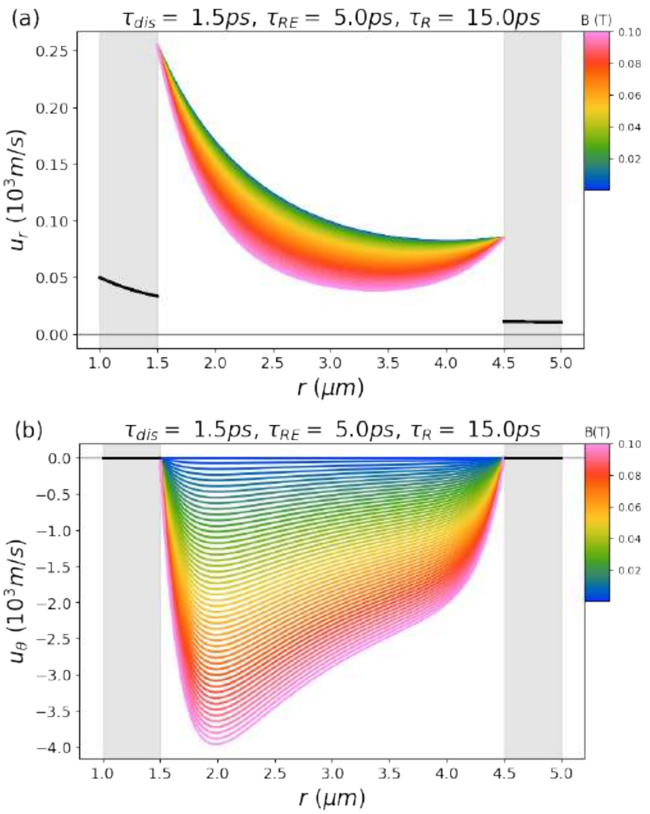


FIG. 3. Radial (top panel) and tangential (bottom panel) components of the hydrodynamic velocity $\boldsymbol{u}$ computed within the “supercollisions model” of energy relaxation. Black lines in the shaded regions show the drift velocity in the leads. Color curves correspond to different values of the external magnetic field according to the shown color coding. The top curve shows values at $\boldsymbol{B} = 0$ and is identical with the results of Ref. [65]. For the parameter values, see Fig 1.

In the opposite limit $\ell_G \ll r_1, r_2$, the leading contribution to the bulk voltage drop is independent of η. For no-slip boundary conditions and in the simplified case where the field is not allowed to penetrate the leads we find for the field-dependent part of U

$$R(B)-R(0) \approx \frac{B^2v_g^2\tau_{\rm dis}\ln p}{6\pi c^2P}+\frac{\ln p}{2\pi}\delta_I\alpha_1^2\tilde{R}_B \propto \tau_{\rm dis}B^2. \tag{20}$$

The voltage drop (20) is proportional to $\tau_{\rm dis}$ similarly to the result in the strip geometry (see Refs. [23, 30]). Of course, in the limit $\ell_G \ll r_1, r_2$ the mean free time $\tau_{\rm dis}$ cannot be arbitrarily large, hence the voltage drop (20) does not diverge. In the limit $\tau_{\rm dis} \to \infty$ the voltage drop crosses over to the above “clean” limit and is given by Eq. (18). However, the limiting expression (20) is independent of viscosity, and hence qualitatively similar to the usual result.

To summarize the results of this section, we have shown that bulk magnetoresistance in neutral graphene in the Corbino geometry exhibits a crossover between the "clean" limit of the large (compared to the disk radius) Gurzhi length to the limit of small Gurzhi length. In the former case, the field-dependent part of the bulk voltage drop is determined by viscosity, while in the latter limit it is proportional to the disorder mean free time similarly to the known result in the strip geometry.

B. Radial component of the velocity and the device resistance

The five equations (14) can be reduced to two coupled differential equations (for similar calculations in the strip geometry see Refs. [23, 48, 68, 82]). To simplify the arguments, we introduce the following notations

$$q = n_I u_r, \quad p = \delta j_{I,r}, \quad x = \frac{2n_I}{3P}\delta P, \quad y = \frac{12\ln 2}{\pi^2}\frac{n_I}{T}\mu_I. \tag{21}$$

In terms of the new variables, Eqs. (14a) and (14c) can be written as

$$\frac{1}{r}\frac{\partial(rq)}{\partial r} + \frac{1}{r}\frac{\partial(rp)}{\partial r} = -\frac{y}{\tau_R}, \tag{22a}$$

$$\frac{1}{r}\frac{\partial(rq)}{\partial r} = -\frac{x}{\tau_{RE}}. \tag{22b}$$

Equation (14e) can be rewritten as

$$\frac{\partial y}{\partial r} = -\frac{6}{\pi}\frac{\tilde{R}n_I}{R_0 T^2\tilde{\tau}}\,p - \frac{12\ln 2}{\pi}\frac{\alpha_1\omega_B^2}{e^2 v_g^2 R_0 T}\,q, \tag{23a}$$

where $\tilde{\tau} = \delta_I/(\tau_{\text{dis}}^{-1} + \delta_I^{-1}\tau_{22}^{-1})$. Finally, Eqs. (14b) and (14d) can be combined into

$$\frac{\partial x}{\partial r} = \frac{2\eta}{3P}\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial(rq)}{\partial r} - \frac{2}{v_g^2}\left[\tau_{\text{dis}}^{-1} + \frac{\omega_B^2}{e^2\tilde{R}}\frac{4T^2\ln^2 2}{3Pv_g^2}\right]q$$
$$+\alpha_1\tilde{\tau}\frac{\pi^2 T^2}{9Pv_g^2}\frac{\omega_B^2}{e^2\tilde{R}}\frac{\partial y}{\partial r}. \tag{23b}$$

Introducing the differential operator

$$\hat{\mathbb{D}}q = \frac{\partial}{\partial r}\frac{1}{r}\frac{\partial(rq)}{\partial r}, \tag{24}$$

we rewrite Eqs. (22) in the matrix form

$$\hat{\mathbb{D}}\begin{pmatrix} q \\ p \end{pmatrix} = \widehat{T}_S\begin{pmatrix} \partial x/\partial r \\ \partial y/\partial r \end{pmatrix}, \qquad \widehat{T}_S = \begin{pmatrix} \frac{1}{\tau_{\text{RE}}} & 0 \\ -\frac{1}{\tau_{\text{RE}}} & \frac{1}{\tau_R} \end{pmatrix}. \tag{25a}$$

Similarly, Eqs. (23) can be written in the matrix form

$$\begin{pmatrix} \partial x/\partial r \\ \partial y/\partial r \end{pmatrix} = -\widehat{M}\begin{pmatrix} q \\ p \end{pmatrix} + \widehat{V}\hat{\mathbb{D}}\begin{pmatrix} q \\ p \end{pmatrix}, \tag{25b}$$

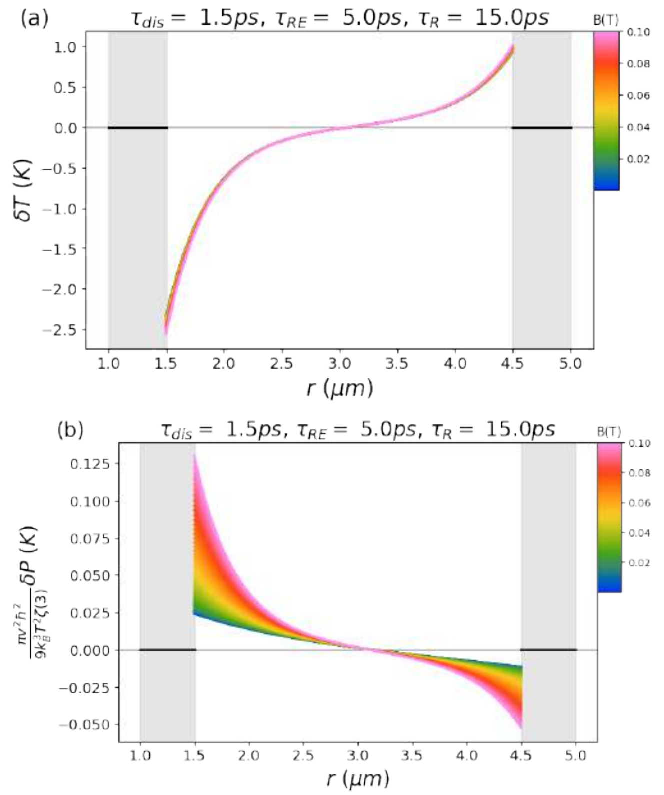


FIG. 4. Local variations of temperature (top panel) and pressure (bottom panel) in the Corbino device computed within the "supercollisions model" of energy relaxation. Black lines in the shaded regions indicate that the leads are at equilibrium. Color curves correspond to different values of the external magnetic field according to the shown color coding. Zero field values are identical with the results of Ref. [65]. For the parameter values, see Fig 1.

where

$$\widehat{V} = \begin{pmatrix} \frac{2\eta}{3P} & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\widehat{M} = \begin{pmatrix} \frac{16\ln^3 2}{3\pi}\frac{\delta_I\tilde{R}_B T^3}{v_g^4 P R_0\tilde{\tau}} + \frac{2}{v_g^2\tau_{\text{dis}}} & \frac{4\ln 2}{3}\frac{\alpha_1\delta_I n_I\tilde{R}_B T}{v_g^2 P R_0\tilde{\tau}} \\ \frac{24\ln^2 2}{\pi^2}\frac{\alpha_1\delta_I\tilde{R}_B}{v_g^2 R_0\tilde{\tau}} & \frac{6}{\pi}\frac{n_I\tilde{R}}{R_0 T^2\tilde{\tau}} \end{pmatrix}.$$

Finally, combining Eqs. (25) we find the equation for the variables p and q

$$\hat{\mathbb{D}}\begin{pmatrix} q \\ p \end{pmatrix} = \widehat{K}\begin{pmatrix} q \\ p \end{pmatrix}, \quad \widehat{K} = \left[1 - \widehat{T}_S\widehat{V}\right]^{-1}\widehat{T}_S\widehat{M}. \tag{26}$$

The obtained equation should be solved with the boundary conditions (9). The solution is straightforward albeit tedious. The results can be expressed in terms of linear combinations of the Bessel functions. Thus obtained solutions are not particularly instructive, hence we present the results of the calculation in graphical form.

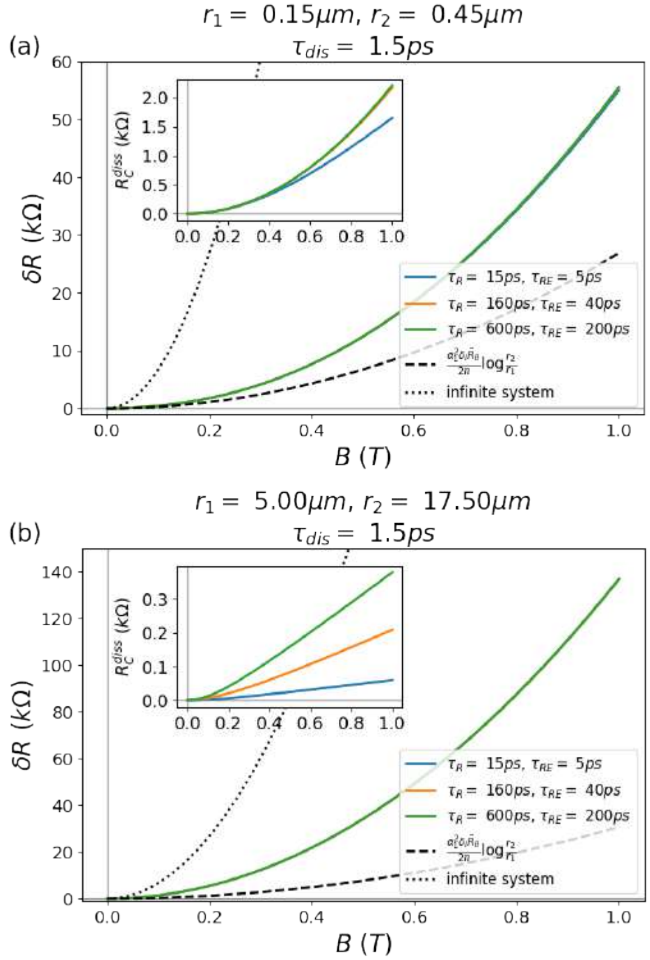


FIG. 5. Magnetoresistance of a small (top panel) and large (bottom panel) Corbino device computed within the "supercollisions model" of energy relaxation. The radii of the Corbino disks are shown above the plots. The black dotted line shows the quantity $\tilde{R}$, which is of the same order of magnitude as the magnetoresistance in the infinite system [23, 82]. Color curves correspond to three different sets of values of the relaxation times. For other parameter values (yielding $\ell_G = 0.2\,\mu$m), see Fig 1. The insets show the contact resistance due to viscous dissipation.

The radial component of the hydrodynamic velocity is shown in the top panel of Fig. 3. The drift velocity in the leads shows the standard Corbino profile, $u_r \propto 1/r$. At each interface, u_r exhibits a jump due to the mismatch of the entropy densities in the sample and leads. For high enough magnetic field, u_r has a sign change close to the interface. However, the corresponding change of direction is hardly seen in the overall flow diagram shown in Fig. 1, since the numerical value of the tangential component u_ϑ is much larger (see the bottom panel of Fig. 3).

The hydrodynamic velocity determines the energy current in the system. The nonuniform energy current results in local variations of the electronic temperature from its equilibrium value (see Fig. 4). The inhomogeneous temperature profile suggests that energy relaxation is less effective in strong magnetic fields. Fig. 1 shows the same data as Fig. 4 but in the form of the color map.

Finally we use the boundary conditions (12) to find the interface jumps of the electric potential which allows us to determine the device resistance. The procedure is the same as in the case of $\boldsymbol{B} = 0$ discussed in Ref. [65]. The results are shown in Fig. 5. For small enough samples (see the top panel in Fig. 5) the device resistance deviates only slightly from $\tilde{R}$ which is of the same order of magnitude as the magnetoresistance in the infinite system [23, 82]. In large samples the deviation is more pronounced and depends on the actual radius of the disk rather than on the ratio p (which is the same in both plots).

The quantitative results shown in this section were computed for a particular choice of the relaxation times. These values are largely phenomenological; however, the magnetoresistance shown in Fig. 5 hardly depends on them, while for larger samples (the bottom panel) the three curves are indistinguishable. However, the values of the relaxation times cannot be completely arbitrary. The point is that the matrix $\widehat{K}$ in Eq. (26) is not guaranteed to have real, positive eigenvalues (although its determinant is positive). In particular, the recombination time τ_R and energy relaxation time τ_{RE} cannot be very different. Within the physical model of supercollisions [59] these time-scales are of the same order of magnitude. Quasiparticle recombination involves supercollision scattering between the bands, while energy relaxation includes an additional contribution of intraband scattering. As a result, the energy relaxation time is shorter than τ_R, but not much shorter since the model does not involve any additional parameter. For such physical values of the relaxation times the eigenvalues of the matrix $\widehat{K}$ are real positive and the resulting magnetoresistance is well accounted for by the curves shown in Fig. 5 where, again, the particular values of τ_R and τ_{RE} do not have a strong quantitative impact on the overall resistance magnitude.

C. Energy relaxation due to electron-phonon interaction

Supercollisions are scattering events involving electron scattering off a phonon and an impurity. As such, this is a next-order process as compared to the direct electron-phonon scattering. The reason supercollisions might be important is that the speed of sound is much smaller than v_g. At high enough temperatures [59, 60] supercollisions indeed dominate, but at lower temperatures the direct electron-phonon scattering cannot be neglected.

Energy relaxation and quasiparticle recombination due to electron-phonon scattering was considered in Ref. [23] within the linear response theory. Since the macroscopic equations of the linear response theory coincide with the linearized hydrodynamic equations [22], we can directly incorporate the corresponding decay terms into our hydrodynamic theory. These decay terms appear in

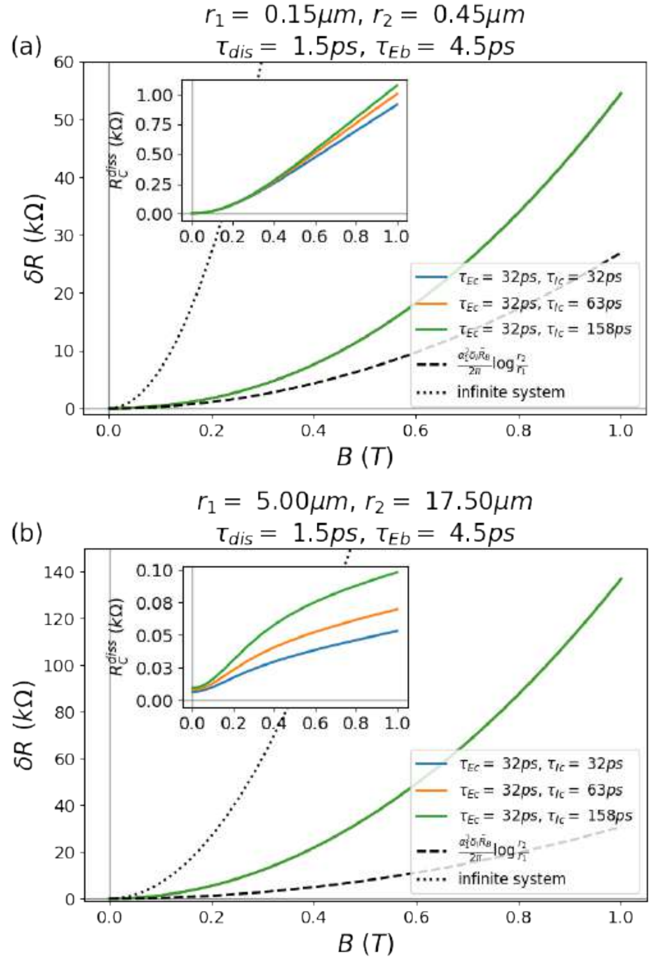


FIG. 6. Magnetoresistance in small (top) and large (bottom) Corbino devices computed within the "electron-phonon model" of energy relaxation (cf. Fig. 5).

Eq. (25a) through the matrix $\widehat{T}_S$. The model of electron-phonon interaction introduced in Ref. [23] corresponds to the following choice of this matrix

$$\widehat{T}_{ep} = -\frac{1}{|\Delta|}\begin{pmatrix} \frac{\gamma}{\tau_{Ec}} + \frac{1}{\tau_{Eb}} & -\frac{\gamma^2}{\mathcal{N}_2\tau_{Eb}} - \frac{\gamma}{\tau_{Ec}} \\ -\frac{\gamma^2\mathcal{N}_2}{\gamma\tau_{Ec}} - \frac{\mathcal{N}_2}{\tau_{Ic}} - \frac{1}{\tau_{Eb}} & \frac{2\gamma}{\tau_{Ec}} + \frac{\gamma^2}{\mathcal{N}_2\tau_{Eb}} + \frac{\mathcal{N}_2}{\tau_{Ic}} \end{pmatrix}, \tag{27}$$

where

$$\gamma = \frac{\pi^2}{12\ln^2 2}, \quad \mathcal{N}_2 = \frac{9\zeta(3)}{8\ln^3 2}, \quad \Delta = \gamma^2 - \mathcal{N}_2,$$

and $\tau_{Eb} \ll \tau_{Ec} \leq \tau_{Ic}$ describe the three independent components of the electron-phonon collision integral [23].

Repeating the above calculation with $\widehat{T}_{ep}$ instead of $\widehat{T}_S$, we arrive at the results that are largely similar to those obtained within the supercollision model, but with a few notable differences (see Figs. 6-11). Unless specified in the figure captions, the parameter values used for the quantitative computation are the same as in the case of supercollisions (see the caption to Fig. 1).

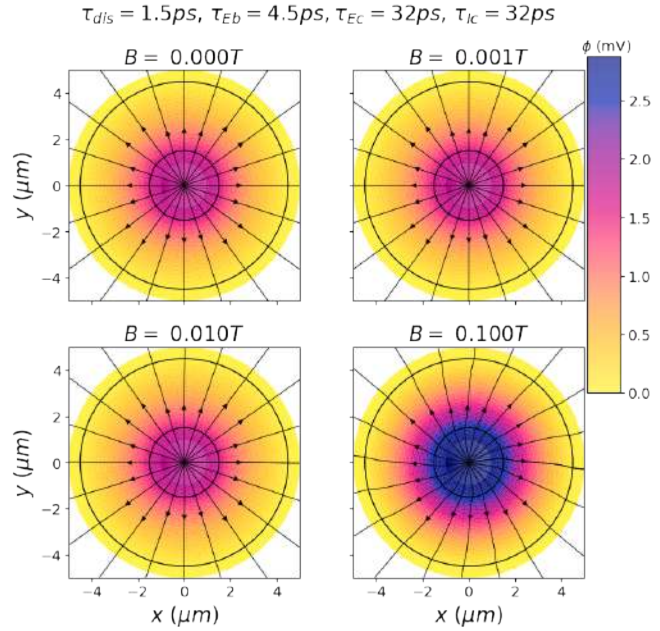


FIG. 7. Electric current density $\boldsymbol{j}$ and potential φ within the electron-phonon model of energy relaxation (cf. Fig. 2).

Magnetoresistance of the device is still positive and parabolic (see Fig. 6). In small devices, it is still largely determined by the quantity $\tilde{R}$ (shown by the black dotted line in Fig. 6 similarly to Fig. 5). In this case, variations of the electron-phonon relaxation rates still do not affect the result in any noticeable way. The results for large devices are also similar to the case of supercollisions: calculated magnetoresistance clearly exceeds $\tilde{R}$ and thus shows a strong dependence on the size of the device (but not on the ratio p).

The electric current density and potential in the device are seen largely the same as in the case of supercollisions, although the deviation of the current from the radial direction (i.e., its tangential component δj_ϑ) is somewhat smaller (see Fig. 7, cf. Fig. 2). This result seems to be consistent with the similarities in the magnetoresistance in the two cases.

The hydrodynamic velocity $\boldsymbol{u}$ is still dominated by its tangential component (see Figs. 8 and 9). The latter shows the behavior that is largely similar to that shown in the bottom panel of Fig. 3, although the magnitude of u_ϑ shows stronger growth with increasing magnetic field. In contrast, the temperature variation is "reversed": now the electronic temperature is increased around the inner contact and decreased close to the outer one (the opposite behavior to that seen in Figs. 1 and 4) (see Fig. 10).

The reversed behavior of the temperature variation corresponds to the change in the radial component of the hydrodynamic velocity u_r. While the jumps at the interfaces with the leads remain the same (insofar u_r on the sample side of the interface is larger than the drift velocity in the leads), the initial slope of u_r as a function

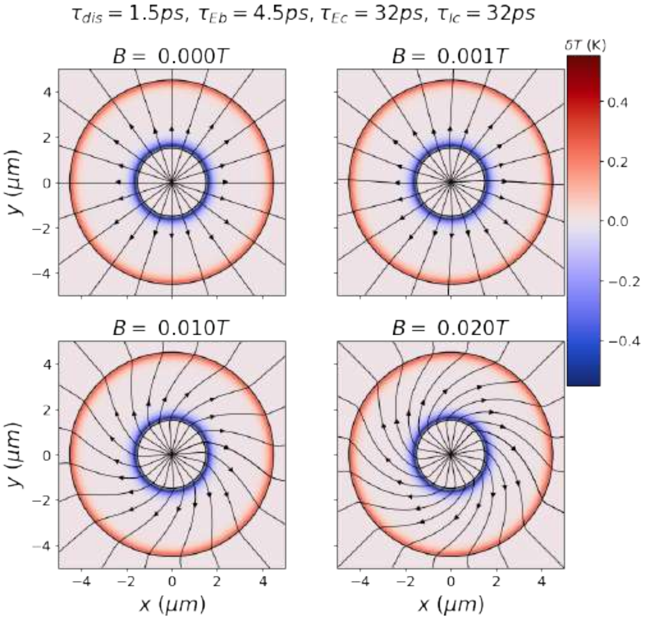


FIG. 8. Hydrodynamic velocity $\boldsymbol{u}$ and temperature δT distribution in the device obtained by solving the hydrodynamic equations at relatively low temperatures where energy relaxation is dominated by direct electron-phonon scattering (cf. Fig. 1).

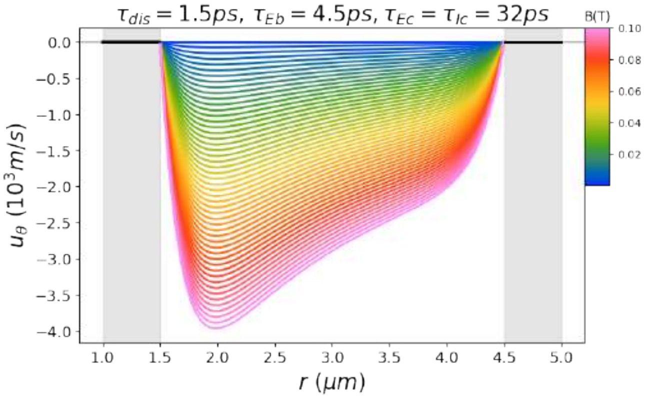


FIG. 9. Tangential component of the hydrodynamic velocity u_ϑ computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

of the radial coordinate has the opposite sign, which does not change with the increase in the magnetic field.

Overall, it is rather natural that the choice of the energy relaxation model mostly affects the energy flow in the device rather than the charge flow. This is a clear consequence of the decoupling of the energy and electric currents in neutral graphene. Although the two currents are being coupled by the magnetic field, the effect appears to be subleading. It is not surprising that the effect of this coupling is most pronounced in strong magnetic fields and large Corbino disks.

Contact resistance induced by viscous dissipation (see insets in Figs. 5 and 6) is also affected by the choice of the energy relaxation model. In the case of supercollisions its qualitative behavior exhibits a strong dependence on the size of the disk (see Fig. 5), while in the model of electron-phonon scattering this dependence is reduced to the magnitude only. The contact resistance is significantly stronger in small devices for both choices of the energy relaxation model as expected on general grounds.

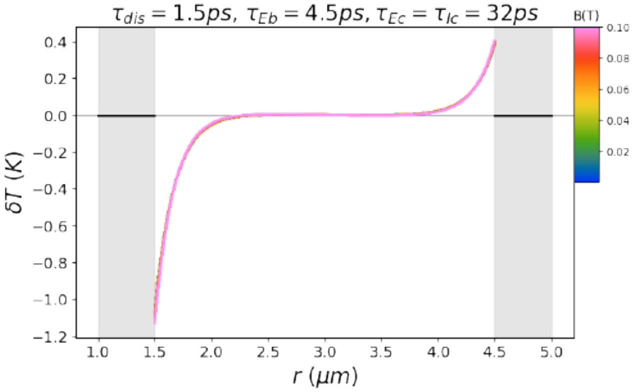


FIG. 10. Local temperature variation computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

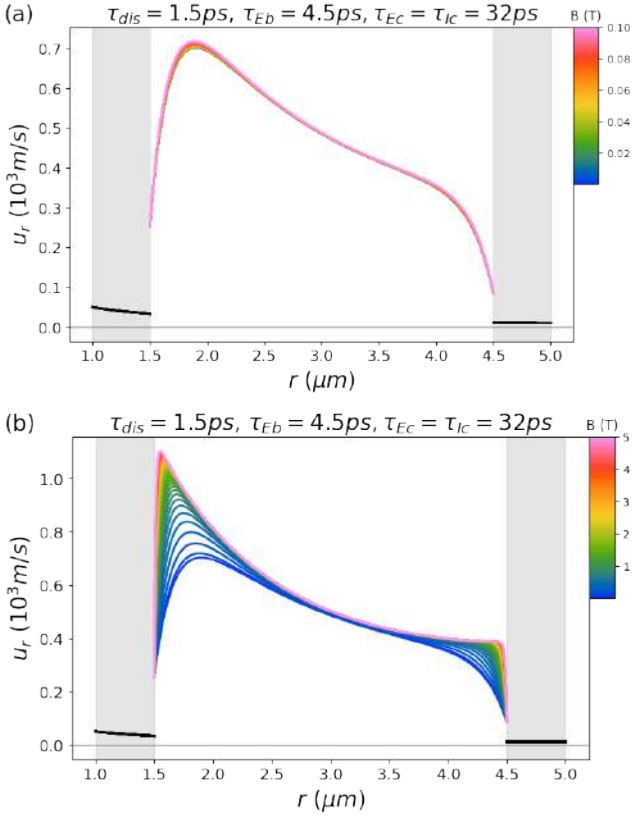


FIG. 11. Radial component of the hydrodynamic velocity u_r computed within the "electron-phonon model" of energy relaxation (cf. Fig. 3).

IV. SUMMARY

In this paper we considered hydrodynamic flows of charge and energy in neutral graphene Corbino disks. We have shown that the Corbino geometry offers a (in principle realizable) possibility to measure electronic viscosity in neutral graphene, a task that so far has appeared elusive. The viscosity coefficient could be extracted from the magnetoresistance data in the ultra-clean limit where the bulk contribution to the device resistance is independent of the electron-impurity scattering time. The bulk resistance dominates over the contact resistance for larger sized disks and hence can in principle be measured in laboratory experiments.

Corbino magnetoresistance in graphene is illustrated in Figs. 5 and 6, where the calculated magnetoresistance is shown for two models of energy relaxation. In both cases, the dependence $R(B)$ is parabolic, similarly to the known result in the strip geometry. The viscosity coefficient can be *in principle* determined experimentally by analyzing the data in a set of different Corbino disks (see Sec. III A). This is not a straightforward task since the magnetoresistance is given by a sum of viscosity-dependent and viscosity-independent terms. In the clean limit $\ell_G \ll r_1, r_2$ [see Eq. (18)], these terms exhibit distinct dependence on the sample size r_2, the ratio of the radii $p = r_2/r_1$, and temperature, making it possible to extract the viscosity coefficient from the experimental data. In the opposite limit [see Eq. (19)], the dominant contribution to magnetoresistance is independent of viscosity. Existing experiments appear to be in the crossover between these two limits. In this paper we have used parameter values yielding $\ell_G \approx 0.2\,\mu$m. The size of the Corbino disk used in a recent experiment [31] was $r_1 = 2\,\mu$m, $r_2 = 9\,\mu$m, which is closer to the "large Corbino disk" illustrated in panels (b) in Figs. 5 and 6 than to the clean limit. It is fair to say that at present extracting viscosity from Corbino magnetoresistance measurements would be extremely difficult. At the same time, we are not aware of any other way to measure the viscosity coefficient in neutral graphene. We believe that viscosity measurements and more generally experimental observation of purely viscous effects in neutral graphene will be more accessible in the near future with even cleaner samples (increasing $\tau_{\rm dis}$ by an order of magnitude).

The regime of linear magnetoresistance as seen in the strip geometry or infinitely sized models does not exist in the Corbino geometry. This can be easily understood by noting that the origin of linear magnetoresistance is in the accumulation of energy and quasiparticle density in the boundary region of a long strip where the sample edges provide a natural barrier for the lateral neutral flow of quasiparticles induced by the magnetic field. In a Corbino disk there is no such edge. The lateral currents (energy and imbalance) flow freely around the disk without accumulating quasiparticles at any point.

Unlike the case of a single-band conductor (e.g., doped graphene), at charge neutrality the electric field is not expelled from the bulk of the sample. Nevertheless bulk viscous dissipation does lead to a discontinuity of the electric potential at the sample-lead interfaces inducing an additional contact resistance. This resistance however is rather small as compared to the resistance of the whole device and should not have a strong effect on the viscosity measurements.

ACKNOWLEDGMENTS

The authors are grateful to P. Hakonen, V. Kachorovskii, A. Levchenko, A. Mirlin, J. Schmalian, A. Shnirman, and M. Titov for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 Program (Project No. 873028 HYDROTRONICS), by the German Research Foundation DFG project NA 1114/5-1 (B.N.N.), and by the German-Israeli Foundation for Scientific Research and Development (GIF) Grant No. I-1505-303.10/2019 (I.V.G.).

---

[1] M. Polini and A. K. Geim, Phys. Today **73**, 28 (2020).
[2] B. N. Narozhny, Riv. Nuovo Cim. **45**, 661–736 (2022).
[3] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).
[4] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).
[5] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).
[6] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).
[7] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, et al., Nature **583**, 537 (2020).
[8] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).
[9] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, et al., Nat. Phys. **13**, 1182 (2017).
[10] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, T. T. K. Watanabe, A. Lucas, A. F. Young,

and A. C. Bleszynski Jayich (2020), arXiv:2002.05065.

[11] J. Mayzel, V. Steinberg, and A. Varshney, Nat. Commun. **10**, 937 (2019).

[12] I. Matthaiakakis, D. Rodríguez Fernández, C. Tutschku, E. M. Hankiewicz, J. Erdmenger, and R. Meyer, Phys. Rev. B **101**, 045423 (2020).

[13] P. Rao and B. Bradlyn, Phys. Rev. X **10**, 021005 (2020).

[14] R. Cohen and M. Goldstein, Phys. Rev. B **98**, 235103 (2018).

[15] P. K. Kovtun, D. T. Son, and A. O. Starinets, Phys. Rev. Lett. **94**, 111601 (2005).

[16] T. Schäfer, Annu. Rev. Nucl. Part. Sci. **64**, 125 (2014).

[17] J. Zaanen, Nat. Phys. **9**, 609 (2013).

[18] K. Kaushansky, M. A. Lichtman, J. T. Prchal, M. Levi, L. J. Burns, and D. C. Linch, eds., *Williams Hematology* (McGraw Hill, New York, 2021).

[19] R. N. Gurzhi, Soviet Physics Uspekhi **11**, 255 (1968), [Usp. Fiz. Nauk **94**, 689 (1968)].

[20] L. S. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).

[21] U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).

[22] B. N. Narozhny, Ann. Phys. **411**, 167979 (2019).

[23] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

[24] A. B. Kashuba, Phys. Rev. B **78**, 085415 (2008).

[25] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[26] M. Müller, L. Fritz, and S. Sachdev, Phys. Rev. B **78**, 115406 (2008).

[27] M. Schütt, P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **83**, 155441 (2011).

[28] B. N. Narozhny, Phys. Rev. B **100**, 115434 (2019).

[29] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **103**, 115402 (2021).

[30] M. Müller and S. Sachdev, Phys. Rev. B **78**, 115419 (2008).

[31] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, et al., Nature **609**, 276 (2022).

[32] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[33] U. Vool, A. Hamo, G. Varnavides, Y. Wang, T. X. Zhou, N. Kumar, Y. Dovzhenko, Z. Qiu, C. A. C. Garcia, A. T. Pierce, et al., Nat. Phys. **17**, 1216 (2021).

[34] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[35] M. Kim, S. G. Xu, A. I. Berdyugin, A. Principi, S. Slizovskiy, N. Xin, P. Kumaravadivel, W. Kuang, M. Hamer, R. Krishna Kumar, et al., Nat. Commun. **11**, 2339 (2020).

[36] G. M. Gusev, A. D. Levin, E. V. Levinson, and A. K. Bakarov, Phys. Rev. B **98**, 161303(R) (2018).

[37] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Sci. Rep. **10**, 7860 (2020).

[38] G. M. Gusev, A. S. Jaroshevich, A. D. Levin, Z. D. Kvon, and A. K. Bakarov, Phys. Rev. B **103**, 075303 (2021).

[39] O. E. Raichev, G. M. Gusev, A. D. Levin, and A. K. Bakarov, Phys. Rev. B **101**, 235314 (2020).

[40] A. Gupta, J. J. Heremans, G. Kataria, M. Chandra, S. Fallahi, G. C. Gardner, and M. J. Manfra, Phys. Rev. Lett. **126**, 076803 (2021).

[41] G. Varnavides, A. S. Jermyn, P. Anikeeva, C. Felser, and P. Narang, Nat. Commun. **11**, 4710 (2020).

[42] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[43] A. Jaoui, B. Fauqué, and K. Behnia, Nat. Commun. **12**, 195 (2021).

[44] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys. Rev. B **92**, 165433 (2015).

[45] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).

[46] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore, Phys. Rev. Lett. **118**, 226601 (2017).

[47] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[48] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **97**, 085109 (2018).

[49] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, and M. Titov, Phys. Rev. B **98**, 125111 (2018).

[50] T. Holder, R. Queiroz, T. Scaffidi, N. Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, and A. Stern, Phys. Rev. B **100**, 245305 (2019).

[51] S. S. Pershoguba, A. F. Young, and L. I. Glazman, Phys. Rev. B **102**, 125404 (2020).

[52] S. Danz and B. N. Narozhny, 2D Materials **7**, 035001 (2020).

[53] J. L. M. Poiseuille, C. R. Acad. Sci. **11**, 961 (1840).

[54] J. L. M. Poiseuille, Annales de chimie et de physique (Series 3) **21**, 76 (1847).

[55] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, London, 1987).

[56] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Phys. Rev. Lett. **120**, 196801 (2018).

[57] H.-Y. Xie and M. S. Foster, Phys. Rev. B **93**, 195103 (2016).

[58] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[59] B. N. Narozhny and I. V. Gornyi, Frontiers in Physics **9**, 108 (2021).

[60] J. C. W. Song, M. Y. Reizer, and L. S. Levitov, Phys. Rev. Lett. **109**, 106602 (2012).

[61] M. W. Graham, S.-F. Shi, D. C. Ralph, J. Park, and P. L. McEuen, Nat. Phys. **9**, 103 (2013).

[62] A. C. Betz, S. H. Jhang, E. Pallecchi, R. Ferreira, G. Fève, J.-M. Berroir, and B. Plaçais, Nat. Phys. **9**, 109 (2013).

[63] K. S. Tikhonov, I. V. Gornyi, V. Y. Kachorovskii, and A. D. Mirlin, Phys. Rev. B **97**, 085415 (2018).

[64] J. F. Kong, L. Levitov, D. Halbertal, and E. Zeldov, Phys. Rev. B **97**, 245416 (2018).

[65] V. Gall, B. N. Narozhny, and I. V. Gornyi, Phys. Rev. B **107**, 045413 (2023).

[66] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[67] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[68] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. B **95**, 165410 (2017).

[69] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).
[70] O. M. Corbino, Nuovo Cimento **1**, 397 (1911).
[71] A. Tomadin, G. Vignale, and M. Polini, Phys. Rev. Lett. **113**, 235901 (2014).
[72] T. Holder, R. Queiroz, and A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).
[73] M. Shavit, A. V. Shytov, and G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).
[74] A. Hui, V. Oganesyan, and E.-A. Kim, Phys. Rev. B **103**, 235152 (2021).
[75] S. Li, A. Levchenko, and A. V. Andreev, Phys. Rev. B **105**, 125302 (2022).
[76] A. Stern, T. Scaffidi, O. Reuven, C. Kumar, J. Birkbeck, and S. Ilani (2021), arXiv:2110.15369.
[77] O. E. Raichev (2022), arXiv:2202.06623.
[78] A. Levchenko, S. Li, and A. V. Andreev, Phys. Rev. B **106**, L201306 (2022).
[79] B. N. Narozhny and M. Schütt, Phys. Rev. B **100**, 035125 (2019).
[80] B. N. Narozhny, Annals of Physics **454**, 169341 (2023).
[81] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).
[82] B. N. Narozhny, I. V. Gornyi, and M. Titov, Phys. Rev. B **104**, 075443 (2021).
[83] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).
[84] D. E. Sheehy and J. Schmalian, Phys. Rev. Lett. **99**, 226803 (2007).
[85] J. González, F. Guinea, and M. A. H. Vozmediano, Phys. Rev. B **59**, R2474 (1999).
[86] J. Gonzalez, F. Guinea, and M. A. H. Vozmediano, Nuclear Physics B **424**, 595 (1994).
[87] B. N. Narozhny, M. Titov, I. V. Gornyi, and P. M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).
[88] M. S. Steinberg, Phys. Rev. **109**, 1486 (1958).
[89] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).
[90] M. P. Marder, *Condensed Matter Physics* (Wiley, 2010).
[91] M. Kamada, V. Gall, J. Sarkar, M. Kumar, A. Laitinen, I. Gornyi, and P. Hakonen, Phys. Rev. B **104**, 115432 (2021).
[92] M. Kumar, A. Laitinen, and P. Hakonen, Nature Communications **9**, 2776 (2018).
[93] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).
[94] G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).
[95] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).
[96] I. S. Burmistrov, M. Goldstein, M. Kot, V. D. Kurilovich, and P. D. Kurilovich, Phys. Rev. Lett. **123**, 026804 (2019).

Observation of giant and tuneable thermal diffusivity of Dirac fluid at room temperature

Alexander Block[1,2], Alessandro Principi[3], Niels C.H. Hesp[1], Aron W. Cummings[2], Matz Liebel[1], Kenji Watanabe[4], Takashi Taniguchi[5], Stephan Roche[2,6], Frank H. L. Koppens[1,6], Niek F. van Hulst[1,6], Klaas-Jan Tielrooij[2,*]

[1]ICFO, The Institute of Photonic Sciences, The Barcelona Institute of Science and Technology, Castelldefels, Spain
[2]Catalan Institute of Nanoscience and Nanotechnology (ICN2), BIST & CSIC, Campus UAB, Bellaterra, Spain
[3]School of Physics and Astronomy, University of Manchester, Manchester, UK
[4]Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan
[5]International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan
[6]ICREA, Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain
*Correspondence to: klaas.tielrooij@icn2.cat

Conducting materials typically exhibit either diffusive or ballistic charge transport. However, when electron-electron interactions dominate, a hydrodynamic regime with viscous charge flow emerges[1–13]. More stringent conditions eventually yield a quantum-critical Dirac-fluid regime, where electronic heat can flow more efficiently than charge[14–22]. Here we observe heat transport in graphene in the diffusive and hydrodynamic regimes, and report a controllable transition to the Dirac-fluid regime at room temperature, using carrier temperature and carrier density as control knobs. We introduce the technique of spatiotemporal thermoelectric microscopy with femtosecond temporal and nanometre spatial resolution, which allows for tracking electronic heat spreading. In the diffusive regime, we find a thermal diffusivity of ~2,000 cm^2/s, consistent with charge transport. Remarkably, during the hydrodynamic time window before momentum relaxation, we observe heat spreading corresponding to a giant diffusivity up to 70,000 cm^2/Vs, indicative of a Dirac fluid. These results are promising for applications such as nanoscale thermal management.

In the diffusive transport regime, electrons and holes undergo random-walk motion with frequent momentum scattering events. When studied on ultrasmall length scales or ultrashort timescales, before momentum relaxation can occur, charges typically move ballistically – in a straight, uninterrupted line. Besides charge, electrons and holes carry electronic heat, with the thermal diffusivity linked to the charge mobility through the Wiedemann-Franz law. Under special conditions – when electron-electron interactions occur faster than momentum relaxation – the hydrodynamic regime emerges[1–22]. In this regime, the rapid collisions between particles can lead to viscous charge flow. The electron system then obeys macroscopic transport laws that are similar to the ones for classical fluid transport. During the last few years, signatures of viscous charge flow in this so-called Fermi-liquid regime were observed in 2D electron systems using electrical device measurements[3,4,10–12] and scanning probe microscopy[5,13,22] typically at cryogenic lattice temperatures and using ultraclean samples with high mobilities, $\mu > 100{,}000$ cm^2/Vs.

A second hydrodynamic regime, which has no analogue in classical fluids, can occur very close to the Dirac point. When the Fermi temperature ($T_F = E_F/k_B$, where E_F is the Fermi energy and k_B is the Boltzmann constant) becomes small compared to the electron temperature T_e, the system becomes a quantum-critical fluid[6,9,14,15,17]. In this Dirac-fluid regime, the non-relativistic description of the viscous fluid is replaced by its ultra-relativistic counterpart, which accounts for the presence of both particles and holes, as well as for their linear energy dispersion. In line with theoretical predictions in this regime[15], noise thermometry measurements indicated a deviation from the Wiedemann-Franz law[19], electrical measurements of the thermoelectric Seebeck coefficient demonstrated a deviation from the Mott relation[20], and a terahertz-probe study revealed the quantum-critical carrier scattering rate[21]. Very recently, scanning probe measurements showed that viscous charge flow occurs both in the Fermi-liquid and Dirac-fluid regimes at room temperature[22]. However, observing and controlling the flow of electronic heat in the hydrodynamic regime at room temperature has so far remained elusive.

Here, we follow electronic heat flow in the diffusive and hydrodynamic regimes at room temperature, and demonstrate a controlled Fermi-liquid to Dirac-fluid crossover, with a strongly enhanced thermal diffusivity close to the Dirac point. These observations are enabled by ultrafast spatiotemporal thermoelectric microscopy, a technique inspired by all-optical spatiotemporal diffusivity measurements[23–25], with the crucial difference that the observable is the thermoelectric current, which is directly sensitive to electronic heat[26]. We use a hexagonal boron nitride (hBN)-encapsulated graphene device that is both a Hall bar for electrical measurements and a split-gate thermoelectric detector (Fig. 1a). Since we use ultrashort laser pulses, with an instrument response time of Δt_{IRF} ~200 fs, to generate electronic heat, we are able to examine the system before momentum relaxation occurs, as we measure a momentum relaxation time τ_{mr} of ~350 fs (see Suppl. Fig. 1). In this temporal regime before momentum is relaxed, we enter the hydrodynamic window, because the electron-electron scattering time τ_{ee} is <100 fs [27], *i.e.* $\tau_{ee} < \Delta\tau_{IRF} < \tau_{mr}$. This is a different approach compared to most previous studies, where hydrodynamic effects were observed by using small system dimensions L in order to eliminate effects of momentum scattering, *i.e.* $v_F \cdot \tau_{ee} < L < v_F \cdot \tau_{mr}$ [3–5,10–13,19,22] ($v_F = 10^6$ m/s is the Fermi velocity). Our approach furthermore exploits elevated carrier temperatures, which greatly increases the accessibility of the Dirac-fluid regime, as for increasing carrier temperatures the crossover occurs increasingly far away from the Dirac point[14,17] (see Fig. 1b). As we will show, during the hydrodynamic window significantly more efficient heat spreading occurs in the Dirac-fluid regime than in the Fermi-liquid regime and in the diffusive regime (see Fig. 1c-d).

Our technique works by employing two ultrafast laser pulses that produce localized spots of electronic heat within tens of femtoseconds[27]. These spots are characterized by an increased carrier temperature $T_e > T_l$, with T_l the lattice temperature (300 K). The degree of spatial spreading of these electronic heat spots as a function of time is governed by the diffusivity D. We control the relative spatial and temporal displacement, Δx and Δt, of the two pulses with sub-100 nm spatial precision and ~200 fs temporal resolution. Each laser pulse is incident on opposite sides of a *pn*-junction at a distance $\Delta x/2$ from the junction. This *pn*-junction is created by applying opposite voltages $\pm\Delta U$ with respect to the Dirac point voltages to the two backgates that form a split-gate structure. The two photo-generated electronic heat spots spread out spatially and part of the heat can reach the *pn*-junction after a certain amount of time, generating a thermoelectric current at the junction through the Seebeck gradient[26]. The small region of the *pn*-junction thus serves as a local probe of electron temperature. While each of the heat spots can create thermoelectric current independently, we obtain spatiotemporal information by examining exclusively the signal that corresponds to heat generated by one of the pulses interacting with heat generated by the other pulse – the interacting heat signal ΔI_{TE}. Since the thermoelectric photocurrent scales sub-linearly with incident power, we can isolate this interacting heat signal ΔI_{TE} by modulating each laser beam at a different frequency, f_1 and f_2, and demodulating the thermoelectric current at the difference frequency $f_1 - f_2$. As illustrated in Fig. 1e-f, the higher the diffusivity D, the more interacting heat signal ΔI_{TE} remains for increasing Δx and Δt.

Figure 2a shows the measured interacting heat current ΔI_{TE} as a function of Δx and Δt. As expected, the largest ΔI_{TE} occurs for the largest spatiotemporal overlap at the *pn*-junction ($\Delta x = \Delta t = 0$). For increasing $|\Delta t|$, we find that the normalized signal extends further spatially, indicating the occurrence of heat spreading (see Fig. 2b). This spatial spread is quantified via the second moment $<\Delta x^2>$, which quantifies the width of the profile at different time delays (see Methods). Similar to recent all-optical spatiotemporal microscopy[24,25], we obtain spatial information beyond the diffraction limit by precise spatial sampling of diffraction-limited profiles. The experimentally obtained spatial spread as a function of Δt (Fig. 2c) is very similar to the calculated results (Fig. 2d), obtained by simulating the experiment with a given diffusivity D (see Methods and Suppl. Note 1). The white lines indicate the values of the spatial spread $<\Delta x^2>$ for different Δt. We also compare the simulated spatial spread $<\Delta x^2>$ *vs.* Δt (blue dashed line in Fig. 2e) with the theoretical expectation according to the heat diffusion equation, $<\Delta x^2> = <\Delta x^2>_{\mathrm{focus}} + 2D\Delta t$ (dash-dotted line in Fig. 2e). Here, D is the same diffusivity that was used as input for the simulation, and $<\Delta x^2>_{\mathrm{focus}}$ is the minimum second moment from the two overlapping pulses (see Suppl. Note 2 and Suppl. Figs. 2-5). The initial slope is the same for both the simulated heat spreading and the theoretical spreading following the heat diffusion equation.

We first discuss the experimental results in the diffusive regime, where $\Delta t > \tau_{\mathrm{mr}}$. For three different gate voltage combinations, corresponding to Fermi energies between 75 and 190 meV (T_{F} = 900 – 2200 K), we extract the spatial spread $<\Delta x^2>$ as a function of Δt from $\Delta I_{\mathrm{TE}}(\Delta x, \Delta t)$ (see symbols in Fig. 2e), and compare it with the results from simulations (solid lines). For these simulations, we have used the diffusivity values that we obtain directly from electrical measurements of charge mobility on the same device (see Suppl. Fig. 1), and the relation between mobility and diffusivity: $D = \mu E_{\mathrm{F}}/2e$ (see Methods). We find excellent agreement, if we account for short-lived ultrafast heat spreading around $\Delta t = 0$, which leads to a larger-than-expected initial spread at time zero $<\Delta x^2>_{\mathrm{min}}$, as we will explain below. Importantly, the agreement between the measured heat spread for $\Delta t > \tau_{\mathrm{mr}}$ and the one calculated using the measured charge mobilities shows that electronic heat and charge flow together, as expected in the diffusive regime. Furthermore, it confirms that our technique is a reliable method for obtaining thermal diffusivities in a quantitative manner.

We now turn to the non-diffusive regime, by exploring the behaviour in the hydrodynamic window, where $\Delta t < \tau_{\mathrm{mr}}$. Surprisingly, the experimentally obtained spatial spreads start at a minimum value $<\Delta x^2>_{\mathrm{min}}$ larger than 2 μm^2, rather than starting at an expected $<\Delta x^2>_{\mathrm{focus}}$ = 0.56 μm^2. A second device we measured reproduces this larger-than-expected spatial spread at time zero (see Suppl. Note 3 and Suppl. Fig. 6). We exclude the possibility of an experimental artefact such as an underestimation of the laser spot size, since we repeated the measurements while scanning through the laser focus, and measured the focus size (Suppl. Figs. 2-5). Furthermore, we observe that the offset depends on the Fermi energy, while keeping all other experimental parameters fixed. We therefore attribute the large experimentally observed minimum $<\Delta x^2>_{\mathrm{min}}$ to ultrafast initial heat spreading that occurs before momentum relaxation takes place, $\Delta t \lesssim 350$ fs (see schematic illustration of spatiotemporal heat spreading in Fig. 1d). The dynamics of this initial heat spreading are washed out by the finite time resolution Δt_{IRF}, and manifests as a large minimum $<\Delta x^2>_{\mathrm{min}}$ at time zero. The observed initial spatial spread suggests a thermal diffusivity of $D = (<\Delta x^2>_{\mathrm{min}} - <\Delta x^2>_{\mathrm{focus}})/2\Delta t_{\mathrm{IRF}} \approx$ 70,000 cm^2/s for the lowest measured E_F of 75 meV. Simulations of heat spreading with an input diffusivity of 100,000 cm^2/s are indeed consistent with the experimentally observed spread in the hydrodynamic window (see red line in Fig. 2e).

We attribute this observation of highly efficient initial heat spreading in the hydrodynamic time window to the presence of the quantum-critical electron-hole plasma. We can exclude that the observed initial spreading is the result of ballistic transport, as we calculate that the ballistic contribution to initial heat spreading would give only $<\Delta x^2>_{\mathrm{ball}}$ = 0.68 μm^2 (see Suppl. Note 2 and Suppl. Fig. 7). Besides, ballistic transport has a very weak dependence (<10 %) on carrier density in this range, as the Fermi velocity does not change significantly for the Fermi energies considered here[28].

In order to provide further evidence of hydrodynamic heat transport, we demonstrate the ability to control the crossover between the Fermi-liquid and quantum-critical Dirac-fluid regime via the ratio T_e/T_F, by independently varying T_e via the incident laser power and T_F via the applied gate voltages. A larger ratio results in less Coulomb screening and correspondingly stronger hydrodynamic effects due to electron-electron interactions. If T_e is significantly larger than T_F, electrons and hole coexist, and the Dirac-fluid regime becomes accessible (see Fig. 1b). We perform spatial scans in the hydrodynamic window at a temporal delay of $\Delta t = 0$, in a geometry with one laser pulse impinging on the junction, while scanning the other pulse across (*x*-axis) and along (*y*-axis) the junction region. Figure 3a-d shows four representative spatial ΔI_{TE} maps with varying T_e/T_F, yet similar signal magnitudes. Clearly, the signal is broader for larger T_e/T_F, indicating faster thermal transport. We repeat these measurements for a range of T_e and T_F values and quantify the initial heat spreading using Gaussian functions, with widths σ_x and σ_y, to describe ΔI_{TE} at $\Delta t = 0$ as a function of Δx or Δy (see Fig. 3e-f and Suppl. Fig. 8). As expected for a crossover from the diffusive Fermi-liquid regime to the hydrodynamic Dirac-fluid regime, both spatial spreads σ_x and σ_y increase substantially for increasing ratio T_e/T_F. These spreads correspond to a diffusivity up to 40,000 cm^2/s (see Methods), similar to the 70,000 cm^2/s we found earlier.

We compare our experimental results to Boltzmann transport calculations following Refs. [9,18], including carrier interactions and long-range impurity scattering. We model impurities as Thomas-Fermi screened Coulomb scatterers of density $0.24 \cdot 10^{12}/\mathrm{cm}^2$. Figure 3g shows the calculated thermal diffusivity D as a function of T_F and T_e, when considering only the hydrodynamic term due to electron-electron interactions, relevant in the hydrodynamic window where $\Delta t < \tau_{\mathrm{mr}}$. A higher electron temperature or lower Fermi temperature leads to strongly increased diffusivity, which is the same qualitative trend as for the experimental data taken at $\Delta t = 0$ in Fig. 3e-f, where a larger initial width originates from a larger diffusivity. The observed trend is clearly not consistent with calculations considering only the diffusive term due to scattering with impurities (Fig. 3h). The calculations thus support our interpretation of a crossover from a diffusive Fermi liquid to a hydrodynamic Dirac fluid for increasing ratio T_e/T_F.

A more quantitative comparison shows that the calculated D in the diffusive regime is around 2000 cm^2/s (see Fig. 3h), in quantitative agreement with the experiment in the diffusive regime. The obtained thermal diffusivity in the hydrodynamic window close to the Dirac point reaches values above 100,000 cm^2/s, even higher than our experimental estimates of 35,000 – 70,000 cm^2/s. Using the calculated diffusivities, we estimate the spatial spread at time zero σ_{calc} (see Methods), as shown in Fig. 3g. These are similar to the experimentally obtained ones, thus confirming our conclusion of highly efficient heat spreading in the Dirac-fluid regime at room temperature, with a diffusivity that is almost two orders of magnitude larger than in the diffusive regime. We note that the theoretical calculations predict that even higher diffusivities are attainable.

Finally, we discuss the (3D) thermal conductivity, in order to assess the ability of the Dirac fluid to transport useful amounts of heat. We find ~100 W/mK in the diffusive regime (see Methods), in agreement with *ab-initio* calculations[29]. In the Dirac-fluid regime, with an electron temperature of ~1000 K, we obtain a thermal conductivity of 18,000 – 40,000 W/mK. This is in agreement with Ref. [15], where values up to 100,000 W/mK were predicted theoretically for large T_e/T_F. The thermal conductivity we obtain is about three orders of magnitude larger than the one obtained in the Dirac-fluid regime at cryogenic temperatures[19]. Interestingly, our results show that in the Dirac-fluid window the electronic contribution to heat transport can be much larger than the phononic contribution with a conductivity of >2000 W/mK [30], which is already exceptionally high. Thus, the Dirac electron-hole plasma can contribute very significantly to thermal transport, extracting heat from hot spots much faster than predicted by classical limits.

In conclusion, our results show that the, until recently unreachable, physical phenomena associated with the Dirac fluid do not only offer an exciting playground for interesting physical phenomena, yet also hold great promise for applications, *e.g.* in thermal management of nanoscale devices. We note that the quantum-critical behaviour can be switched on and off using a modest gate voltage and in systems prepared by standard fabrication techniques. Finally, we believe that the optoelectronic technique we have introduced will be a valuable tool to reach a better understanding of the thermal behaviour of a broad range of quantum materials, with great promise for novel technological applications.

Methods

Fabrication of split-gate thermoelectric device

The split-gate device with Hall geometry consists of exfoliated, single layer graphene encapsulated by hBN, using standard exfoliation and dry transfer techniques. The hBN-graphene-hBN stack is placed on a pre-defined split-gate structure made of graphene, grown by chemical vapour deposition, where the gap between the two gates is ~100 nm, created via electron-beam lithography and reactive ion etching (RIE). The top hBN and graphene are etched into a Hall bar shape with laser lithography and RIE, keeping the split-gate intact, and not etching completely through the bottom hBN. Finally, the Ti/Au side contacts are created by a further step of lithography, RIE and metal evaporation. The fabrication steps are shown in Supplementary Figure 9.

Spatiotemporal thermoelectric current microscopy setup

Our setup enables us to follow electronic heat spreading in space and time, because we use the thermoelectric signal generated by electronic heat interacting at a fixed location (the *pn*-junction), while we vary the spatial displacement of our two laser pulses with respect to this junction, and vary the temporal delay between the two ultrashort pulses. This means that we are following in space and time the diffusion of light-induced electronic heat from the location of light incidence to the *pn*-junction. It is the thermoelectric effect at the *pn*-junction, governed by the Seebeck coefficient, that generates our observable signal, the thermoelectric current. We note that although the value of the Seebeck coefficient itself changes when changing E_F, and when entering the hydrodynamic regime[20], this only affects the magnitude of the thermoelectric current – not how electronic heat is diffusing outside of the *pn*-junction, which is what we are following with our spatiotemporal technique.

A sketch of the setup is shown in Supplementary Figure 10. A Ti:sapphire oscillator (886 nm centre wavelength, 76 MHz repetition rate), is split into two beam paths. Both beams are modulated with optical choppers, at frequencies f_1 = 741 Hz and f_2 = 529 Hz. The relative time delay between the two pulses is controlled by a mechanical delay line. The spatial offset of one beam with respect to the other is controlled with a mirror galvanometer, while the position of the sample with respect to the beams is controlled with a piezo scanning stage. The beams are focused onto the sample with a 40x/NA 0.6 objective lens. We collect the TE photocurrent between the source and drain contacts on either side of the junction via lock-in amplification. This signal is measured by demodulation of the amplified current across the source and drain contacts through the graphene sheet. By demodulating the current signal at the difference frequency of the two modulation frequencies, f_2 - f_1 = 211.7 Hz, we isolate the signal caused by the interaction of both heating sources, which we call the interacting heat current ΔI_{TE}. The temporal resolution of the setup of 200 fs is determined by the 20-80% rise time of transient absorption of graphene in the sample plane of the microscope (see Suppl. Fig. 11). The spatial accuracy is given by the signal-to-noise ratio, and is estimated to be below 100 nm.

We have used two distinct measurement geometries that each have their advantages and characteristics. For the data presented in Fig. 2, the two laser pulses are spatially offset symmetrically with respect to the gate junction region (by $\Delta x/2$ from the junction) by synchronized movement of the galvo mirrors (by Δx) and the piezo sample stage (by $\Delta x/2$). This measurement geometry is most suitable for extracting quantitatively the diffusivity, in particular in the diffusive regime, as shown by the simulations in Fig. 2. For the data presented in Fig. 3, where we focus on the hydrodynamic time window, we use a simpler "asymmetric" measurement geometry that gives a larger signal. Here, we keep one beam fixed on the junction while scanning the other beam by Δx, across the junction (Fig. 3e), and by Δy, along the junction (Fig. 3f), with fixed sample stage and moving the galvo mirrors only.

Estimating Fermi temperature controlled by gate voltage

During photocurrent measurements, the gate voltage U_x is applied to the left (x = "A") or right (x = "B") side of the split-gate. We always apply a symmetric voltage around the experimentally determined Dirac point voltage $U_{\mathrm{x}}^{\mathrm{DP}}$: $U_{\mathrm{A}} = U_{\mathrm{A}}^{\mathrm{DP}} + \Delta U$ and $U_{\mathrm{B}} = U_{\mathrm{B}}^{\mathrm{DP}} - \Delta U$. The gate electrode and the graphene form a capacitor with the dielectric hexagonal boron nitride (hBN), with a thickness of t_{hBN} = 70 nm, and a relative permittivity of $\epsilon_{\mathrm{hBN}} = 3.56$. The carrier density n is calculated via $n = \frac{\epsilon_0 \epsilon_{\mathrm{hBN}}}{e\, t_{\mathrm{hBN}}} \Delta U$, where ϵ_0 is the vacuum permittivity. We calculate the Fermi energy E_F and the Fermi temperature T_F via $E_{\mathrm{F}}^2 = \pi \hbar^2 v_{\mathrm{F}}^2 \cdot n$, and $T_{\mathrm{F}} = \frac{E_{\mathrm{F}}}{k_{\mathrm{B}}}$, where k_B is the Boltzmann constant.

Estimating carrier temperature controlled by laser power

The thermoelectric photovoltage is assumed to be proportional to the time-averaged increase of the electronic temperature T_e above the ambient temperature T_0, as in Ref. [31]. The sub-linear dependence of the thermoelectric current I_{TE} on optical power for the device under study here for illumination with a single pulsed laser (λ = 886 nm) is shown in Supplementary Figure 12. With a linear temperature scaling of the electronic heat capacity for graphene away from the Dirac point, $C_e(T) = \gamma T$, we integrate the heat energy per unit area $dQ = C_e dT$, i.e., $\int_{Q_0}^{Q_0+\Delta Q} \mathrm{d}Q = \int_{T_0}^{T_e} \gamma T \mathrm{d}T$. With the incident power P proportional to the absorbed heat energy per unit area ΔQ, we find that the peak T_e as a function of the laser power P scales as[31] $T_e = \sqrt[2]{T_0^2 + bP}$. Here, the parameter b is defined via $bP = 2\Delta Q/\gamma$, and is used to convert incident power to peak electron temperature (see Suppl. Fig. 12).

Simulation of the experiment

A detailed description of the simulation can be found in Supplementary Note 1 and Supplementary Figure 13. In brief, we calculate the spatiotemporal evolution of electronic heat generated by the two optical pulses in the graphene sheet via the heat equation with a finite difference method. We define Gaussian heating pulses and calculate their temperature rise via the experimentally measured nonlinear power scaling. We extract the differential TE current contribution as a function of Δx and Δt by the difference of the heating at the *pn*-junction region in the presence of both pulses with respect to simulations with only one pulse at a time, analogous to the experimental difference-frequency demodulation.

Quantifying the spatial spread

The following analysis is performed both on the experimental data and on the simulated data of $\Delta I_{TE}(\Delta x, \Delta t)$ for "symmetric experiments" with optical pulses incident at a distance Δx on each side of the *pn*-junction (c.f. Fig. 1-2). For each Δt of the datasets $\Delta I_{TE}(\Delta x, \Delta t)$ we calculate the width of the signal via the second moment, which for an ideal Gaussian profile is equal to the squared Gaussian width σ^2. The second moment is calculated from the pixels Δx_i (i = 1, ..., N) via

$$< \Delta x^2 > (\Delta t) = \frac{\sum_i \left| \Delta x_i - \overline{\Delta x} \right|^2 \Delta I_{\mathrm{TE}}(\Delta x_i, \Delta t)}{\sum_i \Delta I_{\mathrm{TE}}(\Delta x_i, \Delta t)} \text{, with the mean } \overline{\Delta x} = \frac{\sum_i \Delta x_i \, \Delta I_{\mathrm{TE}}(\Delta x_i, \Delta t)}{\sum_i \Delta I_{\mathrm{TE}}(x_i, t)} .$$

We note that the minimum second moment at the focus $<\Delta x^2>_{\mathrm{focus}}$ of 0.56 μm^2 comes from simulating the symmetric experiment, using as input the measured Gaussian beam width at the focus $\sigma_{\mathrm{focus}}^2$ = 0.14 μm^2 (see Suppl. Note 2). For the "asymmetric experiments" with one optical pulse always incident on the *pn*-junction (data of Fig. 3), we always consider the spatial profile only at time zero. Here we find that Gaussian fits with a background give the most reliable results. The entire set of data is shown in Supplementary Figure 8. For each dataset $\Delta I_{TE}(\Delta x)$ or $\Delta I_{TE}(\Delta y)$ taken at $\Delta t = 0$, we perform Gaussian fits using the function $\mathrm{f}(\Delta x) = a \exp\left(-\frac{\Delta x^2}{2\sigma^2}\right) + b$, where the Gaussian squared width σ^2 indicates the thermal spreading. Here, the minimum simulated Gaussian widths are $(\sigma_x^2)_{\mathrm{focus}} = 0.34$ µm^2 and $(\sigma_y^2)_{\mathrm{focus}} = 0.44$ µm^2 (see Suppl. Note 2). The experimentally obtained widths from this dataset as function of gate voltage and optical power are also shown in Supplementary Figure 8, showing an increase with power, *i.e.* larger T_{e}, and an increase towards the Dirac point, *i.e.* smaller T_{F}. We estimate the theoretical Gaussian widths in Fig. 3g using $\sigma^2_{\mathrm{calc}} = (\sigma_x^2)_{\mathrm{focus}} + 2D\,\Delta t_{\mathrm{IRF}}$, where D are the calculated diffusivities.

Electrical measurements

We characterize our device electrically with four-probe measurements (see Suppl. Fig. 1), finding a charge mobility μ of 30,000 – 50,000 cm^2/Vs, depending on carrier density. The measured mobilities correspond to a momentum scattering time τ_{mr} of 300 – 500 fs. Importantly, these scattering times are longer than the temporal resolution (instrument response function, IRF) of our measurement technique, $\Delta t_{\mathrm{IRF}} \approx 200$ fs, thus allowing us to probe our system before and after momentum scattering occurs, *i.e.* in the non-diffusive and diffusive regime. We use these measured charge mobilities to calculate the expected thermal diffusivity via the Einstein relation[32,33] $\mu_{\mathrm{e/h}} = \frac{e}{n_{\mathrm{e/h}}}\frac{\partial n_{\mathrm{e/h}}}{\partial E_{\mathrm{F}}} D_{\mathrm{e/h}}$, where e is the elementary charge, E_{F} is the Fermi energy, and $n_{\mathrm{e/h}}$ is the electron/hole carrier density. For highly doped graphene ($E_{\mathrm{F}} \gg k_{\mathrm{B}}T$) the simple carrier density expression $n_{\mathrm{e/h}} = \frac{E_{\mathrm{F}}^2}{\pi\hbar^2 v_{\mathrm{F}}^2}$, leads to the simple relation: $D_{\mathrm{e/h}} = \frac{E_{\mathrm{F}}}{2e}\mu_{\mathrm{e/h}}$. We note that we obtain the identical result by calculating D from the ratio of the 2D thermal conductivity $\kappa_{\mathrm{e,2D}}$ and the electronic heat capacity C_{e} and using the Wiedemann-Franz law: $\kappa_{\mathrm{e,2D}}/\sigma = \pi^2/3 \cdot (k_{\mathrm{B}}/e)^2 T_{\mathrm{e}}$, where k_{B} is the Boltzmann constant and e the elementary charge, together with the conductivity $\sigma = ne\mu$ and the following heat capacity for graphene (valid for $T_{\mathrm{e}} < T_{\mathrm{F}}$): $C_{\mathrm{e}} = \frac{2\pi\varepsilon_{\mathrm{F}} k_{\mathrm{B}}^2 T_{\mathrm{e}}}{3\hbar^2 v_{\mathrm{F}}^2}$. Given the measured mobilities, we expect thermal diffusivities around 2000 cm^2/s for our sample.

Thermal diffusivity and conductivity of the Dirac fluid

We estimate the enhanced thermal diffusivity of the Dirac fluid by comparing the measured width at time zero $\langle\Delta x^2\rangle_{\mathrm{min}}$ to the expected width $\langle\Delta x^2\rangle_{\mathrm{focus}}$ explained above, via $D = (\langle\Delta x^2\rangle_{\mathrm{min}} - \langle\Delta x^2\rangle_{\mathrm{focus}})/2\Delta t_{\mathrm{IRF}}$. We find values of 74,000 cm^2/s for the symmetric scan (Fig. 2), and 29,000 cm^2/s and 39,000 cm^2/s for the x- and y-directions of the asymmetric scan (Fig. 3), where $\langle\Delta x^2\rangle$ is replaced with (σ_x^2) and (σ_y^2), respectively. The same calculation for a second device (see Suppl. Note 3 and Suppl. Fig. 6) gives a diffusivity of 100,000 cm^2/s. The 3D thermal conductivity $\kappa_{3\mathrm{D}}$ of the Dirac fluid is calculated from the diffusivity D and the electronic heat capacity C_e, via $\kappa_{3\mathrm{D}} = DC_e/d$, where d is the thickness of graphene, 0.3 nm. For the Dirac fluid, we have $T_{\mathrm{e}} > T_{\mathrm{F}}$, and therefore use the “undoped” electronic heat capacity[34] $\frac{18\,\zeta(3)}{\pi(\hbar v_{\mathrm{F}})^2} k_{\mathrm{B}}^3 T_{\mathrm{e}}^2$, where $\zeta(3) \approx 1.202$. With the above estimate D = 35,000 - 70,000 cm^2/s and T_{e} = 1000 K, we obtain the 3D thermal conductivity $\kappa_{3\mathrm{D}}$ = 18,000 - 40,000 W/mK.

Dirac fluid crossover temperature

Following the treatment in Ref. [14], we find the crossover temperature from Fermi liquid to Dirac fluid, as a function of Fermi temperature as $T_{\mathrm{cross}}(T_{\mathrm{F}}) = T_{\mathrm{F}}\left(1 + \lambda \ln\left(\frac{T_0}{T_{\mathrm{F}}}\right)\right)$, where $\lambda = \mathrm{e}^2/16\epsilon_0\epsilon_{\mathrm{r}} v_{\mathrm{F}}\hbar \approx 0.55/\epsilon_{\mathrm{r}}$ for graphene with the dielectric environment $\epsilon_{\mathrm{r}} \approx 3.56$ for hBN. The temperature $T_0 = \frac{2\hbar v_{\mathrm{F}}\sqrt{\pi}}{3^{3/4} k_{\mathrm{B}} a_0} \approx 8.4 \cdot 10^4\ \mathrm{K}$, with the inter-atomic distance $a_0 = 1.42 \cdot 10^{-10}$ m. The resulting crossover temperature is shown in Fig. 1b and as a red dashed line in Fig. 3e-g. We note that the relatively high refractive index of the hBN encapsulant makes the Dirac fluid more easily accessible, as it lowers the crossover temperature compared to vacuum, by a factor of about 2 for the range of T_F studied here.

Acknowledgements

We thank Marco Polini and Pablo Piskunow for fruitful discussions. K.J.T., F.H.L.K., and S.R. acknowledge funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreements No. 804349 (ERC StG CUHL), No. 785219 and No. 881603 (Graphene Flagship). K.J.T. also acknowledges funding from a RyC fellowship No. RYC-2017-22330 and IAE project PID2019-111673GB-I00. ICN2 was supported by the Severo Ochoa program from Spanish MINECO (Grant No. SEV-2017-0706). ICFO acknowledges financial support from Spanish MINECO through the "Severo Ochoa" program (SEV-2015-0522), from Fundació Privada Cellex, from Fundació Mir-Puig, and from Generalitat de Catalunya through the CERCA program. A.B. acknowledges financial support from the International PhD fellowship program 'la Caixa' - Severo Ochoa. A.P. is supported by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS). M.L. acknowledges support by the Ministry of Science, Innovation, and Universities (MCIU/AEI: BES-2016-078727, RTI2018-099957-J-I00 and PGC2018-096875-B-I00). K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan, Grant Number JPMXP0112101001, JSPS KAKENHI Grant Number JP20H00354 and the CREST(JPMJCR15F3), JST. F.H.L.K. acknowledges financial support from the Government of Catalonia trough the SGR grant 1656. N.F.v.H. acknowledges financial support by the European Commission (ERC Advanced Grant 670949-LightNet), the Spanish MCIU (PGC2018-096875-B-I00) and the Catalan AGAUR (2017SGR1369).

References

1. Polini, M. & Geim, A. K. Viscous electron fluids. *Phys. Today* **73**, 28 (2020).
2. Müller, M., Schmalian, J. & Fritz, L. Graphene: A Nearly Perfect Fluid. *Phys. Rev. Lett.* **103**, 025301 (2009).
3. Gooth, J. *et al.* Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide. *Nat. Commun.* **9**, 4093 (2018).
4. Berdyugin, A. I. *et al.* Measuring hall viscosity of graphene's electron fluid. *Science* **364**, 162–165 (2019).
5. Sulpizio, J. A. *et al.* Visualizing Poiseuille flow of hydrodynamic electrons. *Nature* **576**, 75–79 (2019).
6. Foster, M. S. & Aleiner, I. L. Slow imbalance relaxation and thermoelectric transport in graphene. *Phys. Rev. B* **79**, 085415 (2009).
7. Narozhny, B. N., Gornyi, I. V., Titov, M., Schütt, M. & Mirlin, A. D. Hydrodynamics in graphene: Linear-response transport. *Phys. Rev. B* **91**, 035414 (2015).
8. Levitov, L. & Falkovich, G. Electron viscosity, current vortices and negative nonlocal resistance in graphene. *Nat. Phys.* **12**, 672–676 (2016).

9. Zarenia, M., Principi, A. & Vignale, G. Disorder-enabled hydrodynamics of charge and heat transport in monolayer graphene. *2D Mater.* **6**, 035024 (2019).

10. Moll, P. J. W., Kushwaha, P., Nandi, N., Schmidt, B. & Mackenzie, A. P. Evidence for hydrodynamic electron flow in PdCoO2. *Science* **351**, 1061–1064 (2016).

11. Bandurin, D. A. *et al.* Negative local resistance caused by viscous electron backflow in graphene. *Science* **351**, 1055–1058 (2016).

12. Krishna Kumar, R. *et al.* Superballistic flow of viscous electron fluid through graphene constrictions. *Nat. Phys.* **13**, 1182–1185 (2017).

13. Braem, B. A. *et al.* Scanning gate microscopy in a viscous electron fluid. *Phys. Rev. B* **98**, 241304 (2018).

14. Sheehy, D. E. & Schmalian, J. Quantum critical scaling in graphene. *Phys. Rev. Lett.* **99**, 226803 (2007).

15. Xie, H. Y. & Foster, M. S. Transport coefficients of graphene: Interplay of impurity scattering, Coulomb interaction, and optical phonons. *Phys. Rev. B* **93**, 195103 (2016).

16. Lucas, A., Crossno, J., Fong, K. C., Kim, P. & Sachdev, S. Transport in inhomogeneous quantum critical fluids and in the Dirac fluid in graphene. *Phys. Rev. B* **93**, 075426 (2016).

17. Lucas, A. & Fong, K. C. Hydrodynamics of electrons in graphene. *J. Phys. Condens. Matter* **30**, 053001 (2018).

18. Zarenia, M., Smith, T. B., Principi, A. & Vignale, G. Breakdown of the Wiedemann-Franz law in AB -stacked bilayer graphene. *Phys. Rev. B* **99**, 161407 (2019).

19. Crossno, J. *et al.* Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene. *Science* **351**, 1058–1061 (2016).

20. Ghahari, F. *et al.* Enhanced Thermoelectric Power in Graphene: Violation of the Mott Relation by Inelastic Scattering. *Phys. Rev. Lett.* **116**, 136802 (2016).

21. Gallagher, P. *et al.* Quantum-critical conductivity of the Dirac fluid in graphene. *Science* **364**, 158–162 (2019).

22. Ku, M. J. H. *et al.* Imaging viscous flow of the Dirac fluid in graphene. *Nature* **583**, 537–541 (2020).

23. Ruzicka, B. A. *et al.* Hot carrier diffusion in graphene. *Phys. Rev. B* **82**, 195414 (2010).

24. Zhu, T., Snaider, J. M., Yuan, L. & Huang, L. Ultrafast Dynamic Microscopy of Carrier and Exciton Transport. *Annu. Rev. Phys. Chem.* **70**, 219–244 (2019).

25. Block, A. *et al.* Tracking ultrafast hot-electron diffusion in space and time by ultrafast thermomodulation microscopy. *Sci. Adv.* **5**, eaav8965 (2019).

26. Gabor, N. M. *et al.* Hot carrier-assisted intrinsic photoresponse in graphene. *Science* **334**, 648–652 (2011).

27. Brida, D. *et al.* Ultrafast collinear scattering and carrier multiplication in graphene. *Nat.*

Commun. **4**, 1987 (2013).

28. Stauber, T. *et al.* Interacting Electrons in Graphene: Fermi Velocity Renormalization and Optical Response. *Phys. Rev. Lett.* **118**, 266801 (2017).
29. Kim, T. Y., Park, C. H. & Marzari, N. The Electronic Thermal Conductivity of Graphene. *Nano Lett.* **16**, 2439–2443 (2016).
30. Balandin, A. A. *et al.* Superior Thermal Conductivity of Single-Layer Graphene. *Nano Lett.* **8**, 902–907 (2008).
31. Tielrooij, K. J. *et al.* Out-of-plane heat transfer in van der Waals stacks through electron-hyperbolic phonon coupling. *Nat. Nanotechnol.* **13**, 41–46 (2018).
32. Zebrev, G. I. I. Graphene Field Effect Transistors: Diffusion-Drift Theory. in *Physics and Applications of Graphene - Theory* (IntechOpen, 2011).
33. Rengel, R. & Martín, M. J. Diffusion coefficient, correlation function, and power spectral density of velocity fluctuations in monolayer graphene. *J. Appl. Phys.* **114**, 143702 (2013).
34. Lui, C. H., Mak, K. F., Shan, J. & Heinz, T. F. Ultrafast photoluminescence from graphene. *Phys. Rev. Lett.* **105**, 127404 (2010).

Figure 1.

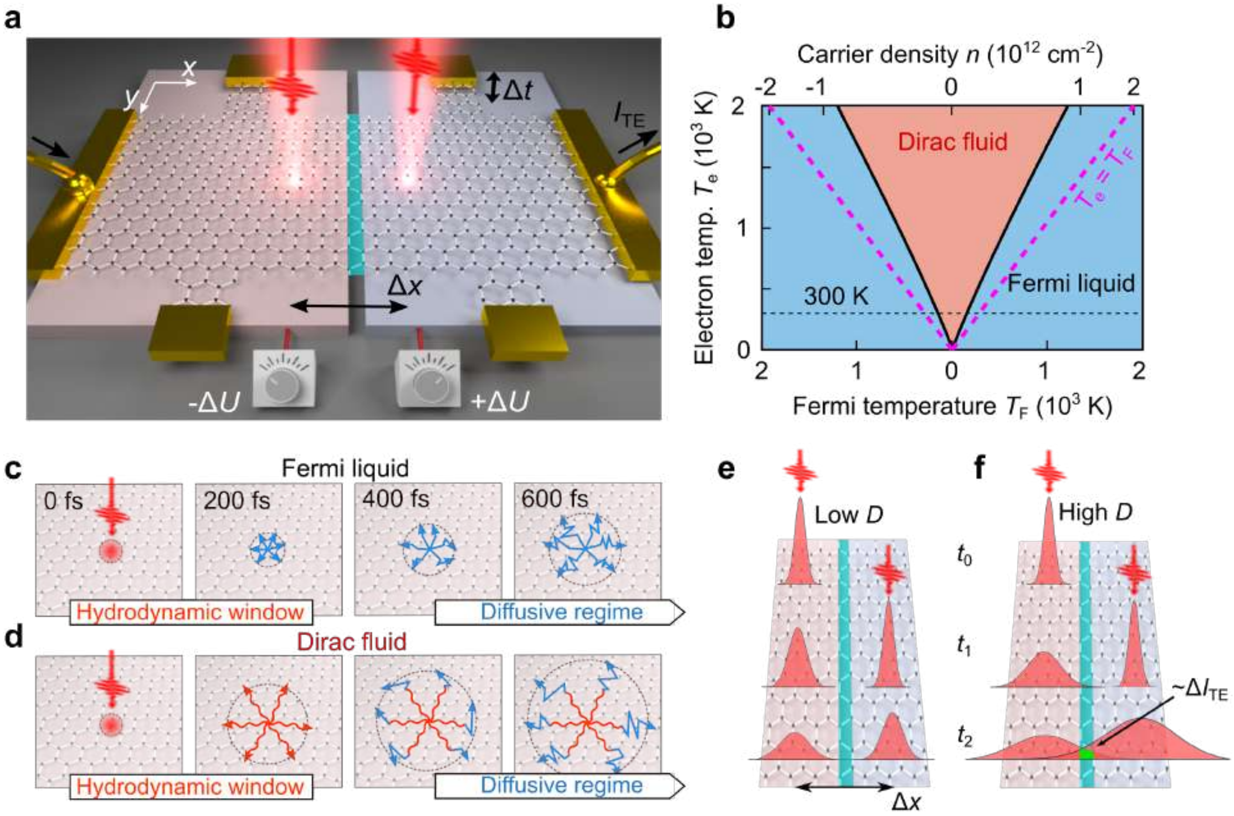


Fig. 1. Spatiotemporal thermoelectric microscopy and heat spreading regimes. (**a**) Concept of the experiment, where a graphene Hall-bar/thermoelectric device is illuminated by two femtosecond heat-generating pulses with a relative temporal offset Δt and a symmetric spatial offset Δx with respect to the *pn*-junction where electronic heat generates a thermoelectric current. The junction is created by applying $+\Delta U$ to one gate and $-\Delta U$ to the other. We isolate the differential thermoelectric current corresponding to light-induced electronic heat from both pulses that has travelled to the junction, where the heat adds up in a nonlinear fashion. (**b**) Phase diagram of the Dirac fluid regime, calculated following Ref. [14]. For increasing T_e the Dirac-fluid regime occurs increasingly far away from the Dirac point. (**c, d**) Illustration of light-triggered spreading of electronic heat in the Fermi liquid regime (**c**) and Dirac fluid regime (**d**). In both cases, for $\Delta t > \tau_{\mathrm{mr}}$, diffusive transport dominates (straight blue lines), while in the hydrodynamic window, with $\Delta t < \tau_{\mathrm{mr}}$, extremely efficient heat transport occurs in the Dirac-fluid regime (wavy red lines). (**e**, **f**) Sketch of the spatial broadening of the heat spots for low (**e**) and high (**f**) diffusivity, indicating a higher interacting heat at the junction region, hence higher ΔI_{TE} signal for higher D.

Figure 2.

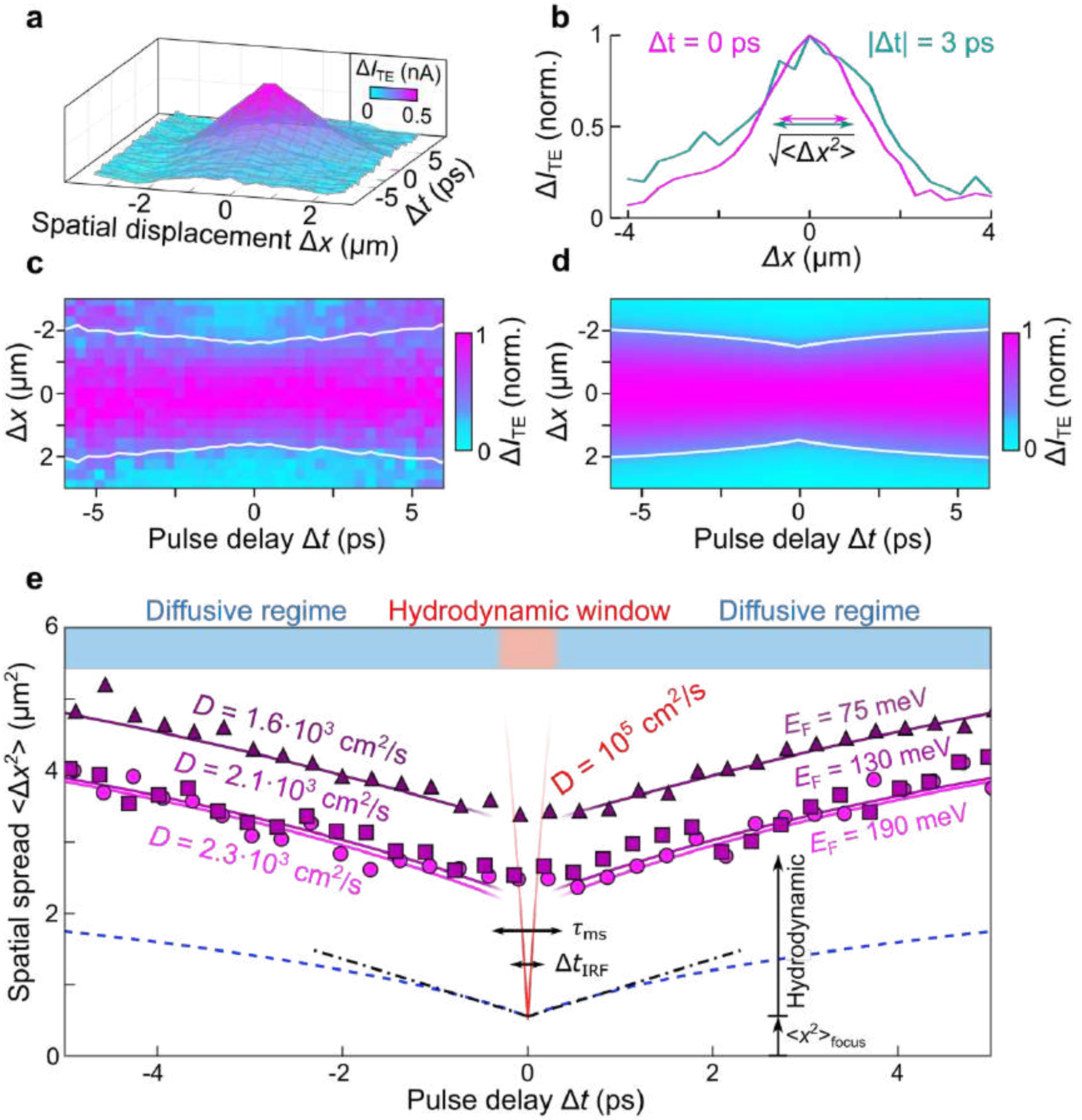


Fig. 2. Spatiotemporal tracking of heat spreading. (**a**) The experimental spatiotemporal differential thermoelectric current ΔI_{TE} as a function of Δx and Δt. (**b**) Normalized profiles, showing a larger spatial extent for larger |Δt|. (**c, d**) Experimental (**c**) and simulated (**d**) normalized ΔI_{TE} for each Δt, showing spatial broadening due to thermal transport as a function of Δt. The white line indicates the spatial spread $<\Delta x^2>$. (**e**) Spatial spread $<\Delta x^2>$ of ΔI_{TE}, as a function of Δt for three different Fermi energies (symbols), with simulation results using as input the diffusivities from electrical mobility measurements (purple solid lines), with offset due to ultrafast heat spreading around time zero. Simulation (blue dashed line) and theoretical heat equation (black dash-dotted line) results with the same input diffusivity and no ultrafast spreading around time zero. Heat spreading with ultrahigh diffusivity in the Dirac-fluid regime (red line), which lasts for a few hundred fs, explains the time zero offset.

Figure 3.

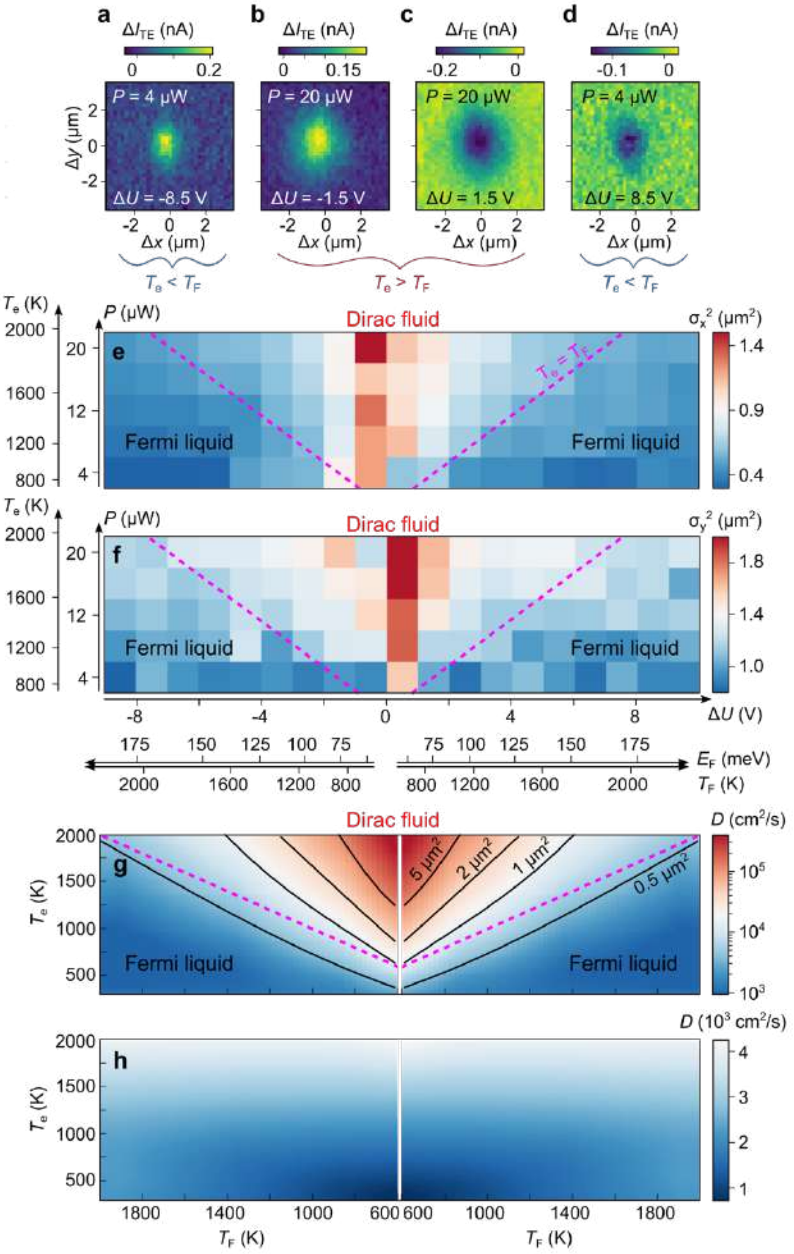


Fig. 3. Fermi-liquid to Dirac-fluid crossover. (**a**-**d**) Time-zero spatial maps of ΔI_{TE} for low optical power P and high gate voltage ΔU (**a**,**d**), and vice versa (**b**,**c**), for *np*- and *pn*- junction (**a**-**b** and **c**-**d**, respectively). For larger ratio T_e/T_F (*i.e.* larger $P/\Delta U$) the spatial extent is clearly larger. (**e**, **f**) Time-zero Gaussian widths for spatial scans with one pulse on the junction and the second one scanning across (**e**) and along (**f**) the graphene *pn*-junction, as a function of P and ΔU. The red dashed line shows the theoretical crossover temperature from Fermi liquid to Dirac fluid regime according to Ref. [14], thus showing our ability to controllably transition into the Dirac-fluid regime with strongly increase thermal diffusivity. (**g**, **h**) Calculation of the thermal diffusivity following Refs. [9,18] with only electron-electron interactions (**g**) and only long-range Coulomb scattering (**h**). The contours in (**g**) are the calculated time-zero spreads σ_{calc}^2 (see Methods).

Thermal transport in compensated semimetals: a mystery explained

Mohammad Zarenia[a,1], **Alessandro Principi**[b], **and Giovanni Vignale**[a]

[a]Department of Physics and Astronomy, University of Missouri, Columbia, Missouri 65211, USA.; [b]School of Physics, University of Manchester, Oxford Road, Manchester M13 9PL, UK.

arXiv:2008.01634v1 [cond-mat.mes-hall] 4 Aug 2020

It is well known that the electronic thermal conductivity of clean compensated semimetals can be greatly enhanced over the electric conductivity by the availability of an ambipolar mechanism of conduction, whereby electrons and holes flow in the same direction experiencing negligible Coulomb scattering as well as negligible impurity scattering. This enhancement – resulting in a breakdown of the Wiedemann-Franz law with an anomalously large Lorenz ratio – has been recently observed in two-dimensional monolayer and bilayer graphene near the charge neutrality point. In contrast to this, three-dimensional compensated semimetals such as WP_2 and Sb are typically found to show a reduced Lorenz ratio. This dramatic difference in behavior is generally attributed to different regimes of Fermi statistics in the two cases: degenerate electron-hole liquid in compensated semimetals versus non-degenerate electron-hole liquid in graphene. We show that this difference is not sufficient to explain the reduction of the Lorenz ratio in compensated semimetals. We argue that the solution of the puzzle lies in the ability of compensated semimetals to sustain sizeable regions of electron-hole accumulation near the contacts, which in turn is a consequence of the large separation of electron and hole pockets in momentum space. These accumulations suppress the ambipolar conduction mechanism and effectively split the system into two independent electron and hole conductors. We present a quantitative theory of the crossover from ambipolar to unipolar conduction as a function of the size of the electron-hole accumulation regions, and show that it naturally leads to a sample-size-dependent thermal conductivity.

Compensated semimetals | Hydrodynamic transport | Ambipolar | Unipolar | Thermoelectric |

1. Introduction

The thermal and electric conductivities of compensated semimetals such as single- and double-layer graphene near the charge neutrality point have been a topic of great interest in recent years - mostly because these systems can be made very clean and feature strong Coulomb interactions between non-degenerate electron and hole carriers near the point of contact of the conduction and valence bands. This clears the way for the observation of hydrodynamic transport, as opposed to conventional single-particle diffusive transport (1–19).

In this regime, the thermal resistivity ($\rho_{\rm th} = \kappa^{-1}$) – defined under the standard condition of zero electric current – is primarily controlled by momentum-non-conserving interactions (scattering from impurities and phonons), while the electric resistivity (σ^{-1}) is primarily controlled by momentum-conserving electron-hole collisions. The physical reason for this difference is well understood. The application of a thermal gradient causes electrons and holes to drift in the same direction (see Fig. 1a) This ambipolar mode of conduction is charge-neutral and therefore automatically satisfies the condition of zero electric current, which is essential to the measurement of the thermal conductivity. At the same time the thermal current is directly proportional to the total momentum of the electron-hole system, which cannot be changed by electron-hole collisions. Hence, except for momentum-non-conserving processes, such as electron-impurity collisions and umklapps, the thermal conductivity would be infinite. On the other hand, an electric field causes electrons and holes to drift in opposite directions (see Fig. 1a). Although the total momentum is now zero, electron-hole collisions transfer momentum between electrons and holes, giving rise to the Coulomb resistivity $\rho_{\rm el}$. Under these conditions the Lorenz ratio $L = \kappa/(\sigma T)$ is much higher than the conventional Lorenz ratio $L_0 = (\pi^2/3)\,(k_B/e)^2$ ($= 2.44\ \times 10^{-8}$ W Ω K^{-2}) of the Wiedemann-Franz law, and is given by (16)

$$L = L_0\left(1+\frac{1}{\Gamma^2}\right) \qquad [1]$$

where $\Gamma^2 = (3/\pi^2)(\rho_{\rm el,dis}/\rho_{\rm el}) \ll 1$ is the ratio of the ordinary Drude resistivity, $\rho_{\rm el,dis}$, to the Coulomb resistivity $\rho_{\rm el}$ – the smaller this is, the deeper we are into the hydrodynamic regime. (Notice that this formula is valid at or near the charge neutrality point, i.e., for chemical potential $\mu = 0$ or, at least, $\mu/(k_BT) \ll \Gamma$.) The resulting Lorenz ratio, $L > L_0$, is clearly seen in the experiments of Ref. (3), which we reproduce in Fig. 1b), and is well above what would be computed in a theory that does not take into account electron-hole scattering.

Notice that the presence of two kinds of carriers with opposite charges is essential to the enhancement of the Lorenz ratio. If we had only one kind of carriers, then the requirement of

Significance Statement

A fundamental obstacle to the realization of efficient thermoelectric devices is the tendency of electronic thermal and electric conductivities to track each other: their ratio (the so-called Lorenz ratio) is close to having a universal value in most materials. Semimetals, having two types of electric carriers of opposite polarity, offer different modalities of conduction for heat and charge and therefore are excellent candidates to achieve anomalously large or small Lorenz ratios. Recently, it has been experimentally realized that strong interaction between carriers in a clean semimetal can significantly increase or decrease the Lorenz ratio. Here, we explain why some semimetals exhibit greatly enhanced Lorenz ratios, while others exhibit reduced ones. This has important implications for the design of thermoelectric devices.

All authors contributed equally to the paper.

The authors have no competing interests.

---

[1] To whom correspondence should be addressed. E-mail: zareniam@missouri.edu

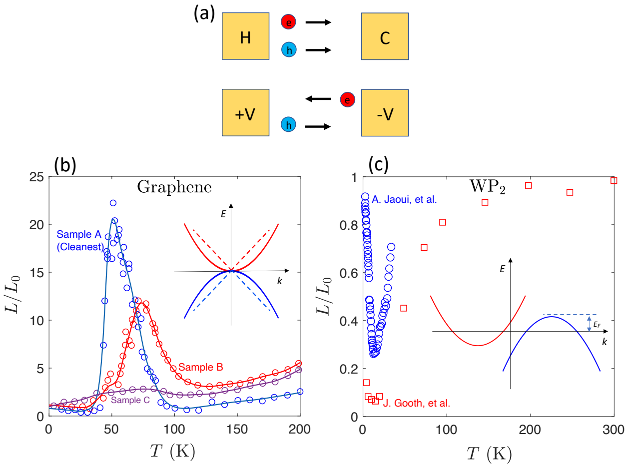


Fig. 1. (a) Schematic illustration of the difference between heat and charge current in a charge neutral system. A thermal current can be set up in a semimetal simply by letting electrons and holes drift with equal speeds in the same direction (upper row). Electric field causes electrons and holes to drift in opposite directions (lower row). (b) Experimental observation of the enhanced Lorenz ratio in monolayer graphene (the solid lines are guides to the eye) - data were reproduced from Ref. (3). (c) Experimental observation of the reduced Lorenz ratio in WP_2 - data were reproduced from Refs. (20, 21). The insets in (b) and (d) depict the low energy bands in graphene systems (dashed curves for monolayer and solid curves for bilayer) and in a compensated semimetal, respectively. In both cases, the chemical potential is taken as the zero of the energy.

zero-electric current in a thermal transport experiment would force these carriers to adopt a distribution in which their direction of drift changes sign depending on whether their energy is above or below the Fermi level. In this case, Coulomb interactions between the carriers would increase the thermal resistivity, producing a Lorenz ratio that is less than L_0 (22), exactly the opposite of what happens in the ambipolar case.

In view of the above discussion, it may come as a surprise that well-known compensated semimetals, such as WP_2 do not show, experimentally, any sign of ambipolar thermal transport. On the contrary, the Lorenz ratio of this and other compensated semimetals, is found to be lower than L_0 (23), which, as we have just seen, is a signature of interaction effects in *unipolar* transport. Earlier measurements on Bi (24, 25) also found a reduction of the Lorenz ratio rather than an enhancement. A cartoon of the band structure of a compensated semimetal with a negative indirect gap is shown in the inset of Fig. 1c. For simplicity, we assume parabolic bands of opposite curvature for electrons and holes. The electron and hole bands are well separated in momentum space, in contrast with those of graphene where electrons and holes coexist in the same region of momentum space. Experimental measurements of the thermal and electric conductivity, reproduced in Fig. 1c, clearly show the reduction of the Lorenz ratio in a range of temperatures $k_BT < \varepsilon_F$ in which electrons and holes can be safely regarded as degenerate Fermi liquids. Here, ε_F is the Fermi energy of electrons and holes measured from the bottoms of the respective bands, while the chemical potential is $\mu = 0$ as required for charge neutrality.

What is the reason for this difference?

The first explanation that comes to mind invokes the different regimes of Fermi statistics of electrons and holes in the two systems. Electrons and holes are degenerate Fermi liquids in WP_2, where the inverse quasiparticle lifetime scales as $(k_BT)^2/\varepsilon_F$; but, in graphene, they are non-degenerate, strongly interacting (Planckian) particles whose inverse lifetime scales as k_BT. The difference manifests in the behavior of the intrinsic electric resistivity (caused by electron-hole scattering): $\rho_{\rm el}$ is essentially independent of temperature in single- and double-layer graphene, but becomes proportional to $(k_BT/\varepsilon_F)^2 \ll 1$ in WP_2. The small value of $\rho_{\rm el}$ suggests that the "hydrodynamicity" parameter $1/\Gamma^2$ of Eq. (1) in WP_2 is much smaller than 1, consistent with the fact that electrons and holes are degenerate Fermi liquids. These considerations lead one to expect that L should be close to L_0, but not smaller than L_0. We note that recent theoretical calculations of the thermal conductivity of compensated semimetals (26) have yielded $L < L_0$ only because the ambipolar conduction channel was not allowed to be part of the solution. Those results for the thermal conductivity are qualitatively similar to what would be obtained by enforcing the zero electric current conditions separately for electrons and holes, without allowing for the possibility that the the electric currents of electrons and holes cancel against each other.

In view of the fact that the Lorenz ratio of WP_2 is indeed found to be less than L_0 we are left with the following problem: Why is the ambipolar channel of thermal conduction apparently disabled in WP_2, while it is clearly operative in graphene? In this paper we propose a resolution of this puzzle.

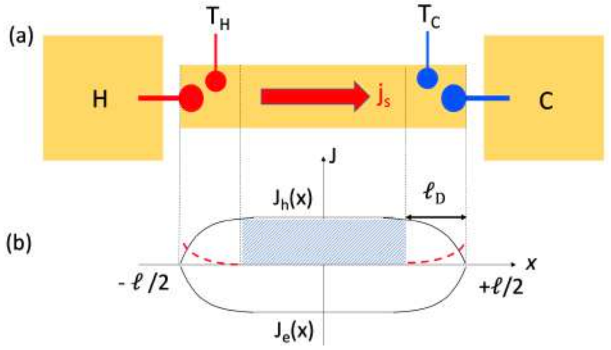


Fig. 2. (a) Schematics of the thermal conductivity measurement. (b) Spatial distribution of the electron and hole components of the electric current. ℓ_D is the diffusion length. The ambipolar transport region is shaded. The red dashed lines represent the enhanced density of electrons and holes near the contacts.

2. Ambipolar transport in the presence of contacts - qualitative description

In a typical thermal conductivity measurement (see Fig. 2a) no electric current is extracted from the system. This gives us the condition

$$j_e + j_h = 0 \qquad [2]$$

where j_e and j_h are the electric currents associated with electrons and holes respectively. At the contacts, however, both the electron and the hole currents are expected to vanish and therefore we have the boundary condition $j_e = j_h = 0$ at the contacts. In principle, this boundary condition could be homogeneously enforced all along the sample (we assume the sample is a channel of length ℓ extending from $-\ell/2$ to $+\ell/2$ along the x axis). Then the electrons and the holes would be effectively decoupled: there would be no reason for momentum or energy to flow preferentially from one group of carriers to the other. The thermal conductivity would be $\kappa = \kappa_e + \kappa_h$, κ_e and κ_h being the thermal conductivities of electrons and holes in isolation. The pattern of motion would reproduce that of a system with only one kind of carrier: the drift direction would change sign depending on whether their energy is above or below the Fermi level. Then electron-electron and hole-hole interactions would ensure that the Lorenz ratios $\kappa_e/T\sigma_e$ and $\kappa_h/T\sigma_h$, with σ_e and σ_h being the electric conductivities of electrons and holes in isolation, are lower than the non-interacting ratio L_0. The total electric conductivity σ is lower than $\sigma_e + \sigma_h$, due to the effect of electron-hole collisions. However, this effect can be neglected in the clean limit, because σ_e and σ_h are very large. Therefore, under this boundary condition we would expect the Lorenz ratio to be lower than L_0, as it is indeed observed to be in experiments.

But why should the boundary condition $j_e = j_h = 0$ be enforced homogeneously throughout the sample? Notice that j_e and j_h are not separately conserved, due to the possibility of electron-hole recombination. Therefore $j_e = 0$ at the contacts does not demand $j_e = 0$ everywhere. On the contrary, the principle of least entropy production (27) demands that the system take maximal advantage of the ambipolar channel of thermal conduction by keeping $j_e = -j_h \neq 0$ in the bulk. The way this is achieved is by creating regions of increased electron and hole density in the vicinity of the contacts. This is shown schematically by the red dashed lines in Fig. 2b. The excess densities of electrons and holes are identical, so that charge neutrality is preserved, but the local chemical potentials for electrons and holes shift in opposite direction. The gradients of electron and hole densities act as opposing forces, which gradually bring the electron and hole currents to zero. In the next section we will show that the size of the electron-hole accumulation regions is given by the diffusion length

$$\ell_D = \sqrt{D\tau_r} \qquad [3]$$

where D is the diffusion constant of electrons or holes, related to the electric conductivity by the usual Einstein relation ($D \sim v_F^2\tau$ in a degenerate Fermi liquid, $D \sim (k_BT/m)\tau$ in a non degenerate electron gas, where τ is the momentum relaxation time) and τ_r is the electron-hole recombination time. Notice that ℓ_D can be very large in a clean semimetal with a long electron-hole recombination time. For example, with a diffusion constant on the order of 10^4 cm^2/s and an electron-hole recombination time on the order of 10^{-6} s (e.g. see chapter 4 in (28)) we obtain $\ell \simeq 10^{-1}$ cm which is comparable to the size of experimental samples (20, 21). It is also worth noting that this mechanism of gradual suppression of the current is unique to ambipolar systems. In a unipolar system, carrier accumulation is inevitably associated with charge accumulation and the diffusion length is replaced by the much smaller screening length: the electric current is suppressed all over the sample by the uniform electric field generated by a surface charge layer.

The following qualitative picture emerges from our discussion. In a typical thermal conductivity measurement the system splits into three sections: (i) A central section of length $\ell - 2\ell_D$ (assuming $\ell > 2\ell_D$) in which thermal transport occurs via the ambipolar channel with $j_e = -j_h \neq 0$ and the thermal resistivity is given by $\rho_{\rm th,ambi}$ (ii) Two accumulation regions of length ℓ_D adjacent to the contacts, in which j_e and j_h are essentially zero and the thermal conductivity is given by $\rho_{\rm th,uni} = (\kappa_e + \kappa_h)^{-1}$. The thermal resistivities of the three sections add in series, leading us to the final result

$$\rho_{\rm th} = \alpha(\ell)\rho_{\rm th,uni} + [1 - \alpha(\ell)]\rho_{\rm th,ambi} \qquad [4]$$

where $\alpha(\ell) \simeq 2\ell_D/\ell$ for $\ell > 2\ell_D$, and $\alpha(\ell) = 1$ for $\ell < 2\ell_D$. This qualitative result will be substantiated in the next section by detailed calculations. In particular, we will derive a more accurate form of the weight function

$$\alpha(\ell) = \frac{2\ell_D}{\ell}\tanh\left(\frac{\ell}{2\ell_D}\right). \qquad [5]$$

If, as we expect in very clean samples, $\rho_{\rm th,ambi} \ll \rho_{\rm th,uni}$ Eq. (4) can be further simplified to

$$\rho_{\rm th} = \alpha(\ell)\rho_{\rm th,uni}. \qquad [6]$$

Here the dependence of the thermal resistivity on the sample size along the direction of flow is evident – as well as a distinct possibility to get $\rho_{\rm th} \simeq \rho_{\rm th,uni}$ when ℓ and $2\ell_D$ are comparable. No such complications arise in measurements of the electric conductivity, because the system remains homogeneous in those measurements.

According to this description, the difference between compensated semimetals like WP$_2$ and graphene arises from the difference between their electron-hole equilibration times. In WP$_2$ electrons and holes are well separated in momentum space, making the recombination process very slow. As a result, the diffusion length becomes comparable to the size of the sample and the thermal conductivity reduces to the sum of the thermal conductivities of electrons and holes in isolation, implying a Lorenz ratio lower than L_0, as discussed above. In single- and double-layer graphene, electrons and holes coexist in the same (small) region of momentum space. Transfer of non-equilibrium carriers between the conduction and valence bands is fast, preventing the establishment of different local chemical potential for electrons and holes. Therefore the diffusion length is negligible and the thermal resistivity plummets, leading to a Lorenz ratio higher than L_0.

Throughout this paper we have assumed that the current density remains uniform in the direction *perpendicular* to the flow. Thus, we have deliberately disregarded contributions to the resistances arising from the transverse electronic viscosity and boundary conditions which mandate the vanishing of the electronic current along the lateral boundaries of the channel. This corresponds to considering a wide conduction channel. A detailed analysis of narrow-channel effects is beyond the scope of this paper.

3. Ambipolar transport in the presence of contacts - quantitative theory

In this section we derive the 2×2 matrix of thermoelectric resistivities for a 1D channel ($-\ell/2 \leq x \leq \ell/2$). The latter relates electric and thermal currents to electric fields and thermal gradients. To simplify the following derivation, we now define $|j_{ns}\rangle = {}^t(j_n, j_s)$ and $|F_{ns}\rangle = {}^t(-eE, -k_{\rm B}\partial_x T)$ the vectors of thermoelectric currents and fields, respectively. Here $j_n = j_e + j_h$ and j_s are the electric and thermal currents, respectively, while E is the electric field and $\partial_x T$ the temperature gradient. Then, the resistivity matrix $\hat{\rho}$, such that $|F_{ns}\rangle = \hat{\rho}|j_{ns}\rangle$, has the form

$$\hat{\rho} = \begin{pmatrix} \rho_{\rm el} + Q^2\rho_{\rm th} & -Q\rho_{\rm th} \\ -Q\rho_{\rm th} & \rho_{\rm th} \end{pmatrix} , \qquad [7]$$

where $\rho_{\rm el}$ and $\rho_{\rm th}$ are the electric and thermal resistivities in reduced units. That is to say, they are the usual electric and thermal resistivities multiplied by e^2 and $k_{\rm B}^2 T$, respectively, while Q is the Seebeck coefficient in units of $k_{\rm B}/e$. Throughout this paper we work with these reduced units.

This well-establish two-mode description is however insufficient in describing thermoelectric transport in systems where conduction can occur via both electrons and holes, if one wishes to separately impose boundary conditions on the particle and hole currents j_e and j_h. It is in fact clear that, by its own construction, such description allows imposing boundary conditions only on the total electric current, j_n, which is the sum of electron and hole currents. To treat these currents separately, it is necessary to extend this theory by adding a third mode, the "imbalance" current $j_\delta = j_e - j_h$, as well as the corresponding imbalance field $F_\delta = -\partial_x(\mu_e - \mu_h)$. Here, μ_e and μ_h are the electron and hole chemical potentials, respectively. Indeed, by taking linear combinations of the imbalance and electric currents, it becomes possible to separately describe the propagation of electrons and holes.

We stress that the imbalance mode plays a rather special role in the present theory. From an experimental perspective, only two fields and currents, the electric and thermal ones, can be externally applied and measured. On the contrary, j_δ and F_δ are not directly accessible. They represent the internal rearrangement that the particle flow undergoes as a result of the application of external probes, while being subject to the boundary conditions. Their presence in the theory is vital to the correct implementation of boundary conditions and particle-hole recombination. However, in order to describe experiments, it is sufficient to down-fold such unfamiliar three-mode theory, resulting from the introduction of imbalance currents and fields, to the more conventional two-mode one of Eq. (7). Here we show that, by applying the boundary conditions on j_δ in the presence of particle-hole recombination, we are able to integrate out the imbalance current and reduce the three-mode thermoelectric resistivity to the more familiar 2×2 matrix of Eq. (7). From that we will then be able to read out the values of electric and thermal resistivities, as well as of the Wiedemann-Franz ratio and the Seebeck coefficient.

In the three-mode theory, the fields are related to the currents via a 3×3 resistivity matrix:

$$\left(\begin{array}{c} |F_{ns}\rangle \\ \hline F_\delta \end{array} \right) = \left(\begin{array}{c|c} \hat{\rho}_{ns} & |\rho_\delta\rangle \\ \hline \langle\rho_\delta| & \rho_{\delta\delta} \end{array} \right) \left(\begin{array}{c} |j_{ns}\rangle \\ \hline j_\delta \end{array} \right) . \qquad [8]$$

Here $\hat{\rho}_{ns}$ is a 2×2 block, whereas $|\rho_\delta\rangle$ is a two-component vector. Hereafter $\langle v|$ denotes the transposed of the vector $|v\rangle$. The vector of currents on the right-hand side of Eq. (8) specifies the state in which the system is prepared. Once such state is defined, this equation tells us which potential drops, thermal gradients and imbalance fields can be measured at the boundaries of the sample. We note that the specific forms of $\hat{\rho}_{ns}$, $|\rho_\delta\rangle$ and $\rho_{\delta\delta}$ are immaterial, as we proceed to show. The only property of the 3×3 matrix of Eq. (8) that we will use in what follows is that its determinant vanishes. We stress that such property is not generic to all thermoelectric matrices, but is a consequence of the existence of a conserved mode (the total momentum) in the present theory.

In fact, when electron-electron interactions are the dominant scattering mechanism, and barring Umklapp processes, the total momentum is a conserved quantity which must always be included in the theory, regardless of boundary conditions. This can be accomplished in two ways. One possibility is that the total momentum is already present explicitly in the 3×3 resistivity matrix of Eq. (8). This happens in very specific cases in which one of the three currents (j_n, j_s or j_δ) coincides with the momentum density. For example, in a parabolic band electron gas the electric current density coincides with the momentum density, whereas for massless Dirac fermions (e.g., in undoped graphene) the momentum density is directly proportional to the heat current density. Finally, in a gapless parabolic-band semimetal such as undoped bilayer graphene, the momentum density coincides with the imbalance current density j_δ.

In all these cases, the current that is proportional to the momentum *cannot* decay over time, since particle-particle collisions do not affect it. Once launched, it can only be relaxed by momentum-non-conserving scattering processes (e.g., electron-phonon collisions). By the very definition of hydrodynamic regime of transport, however, such processes seldom occur and are in fact neglected altogether in a first approximation. This fact has a striking consequence. If the system is prepared in a state in which only such conserved current exists, since it experiences neither resistance nor dissipation during its propagation, it *cannot* give rise to a drop in electric field or thermal gradient. Mathematically, if such *nontrivial* state is introduced on the right-hand side of Eq. (8), and is thus multiplied by the 3×3 resistivity matrix, it produces a null vector of fields. It is, therefore, a "zero mode" of the resistivity matrix. Since it is nontrivial, *i.e.* it is not the vector with all currents equal to zero, this in turn implies that the determinant of the 3×3 resistivity matrix of Eq. (8) must vanish.

In general, however, none of the three currents coincides with, or is directly proportional to the total momentum. Therefore, to include such mode one should in principle start from four-mode theory, the fourth component being the momentum. Then, via a down-folding procedure similar to the one we will describe momentarily, one can obtain the 3×3 resistivity matrix of Eq. (8). This procedure is shown in App. A: the end result is that the momentum mode is *implicitly* included in the 3×3 resistivity matrix and its determinant still vanishes (it is indeed possible to construct a current which is a combination of particle, thermal and imbalance ones that cannot decay over time). As we proceed to show, the vanishing of the determinant of the three-mode matrix, consequence of the presence of a conserved mode in the theory, plays a fundamental role in describing the transition between unipolar and ambipolar regimes, as well as the size-dependence of the thermal resistivity.

To take into account particle-hole recombination and boundary conditions in the thermoelectric transport, we now assume that the imbalance density, n_δ, satisfies the following continuity equation:

$$\partial_t n_\delta + \partial_x j_\delta = -\frac{\nu_0}{\tau_r}(\mu_e - \mu_h) , \qquad [9]$$

where τ_r is the electron-hole recombination time, while ν_0

is the density of states of electrons and holes (assumed to be equal) at the Fermi level. Eq. (9) can be derived from the Boltzmann equation. In general, the collision integral of electron-electron interactions does not conserve the imbalance density and therefore the latter decays over time with a typical time scale τ_r. Note that conservation of the imbalance density is obtained in the limit $\tau_r \to \infty$, hence Eq. (9) is completely general.

Before continuing, it is necessary to discuss which boundary conditions apply in different situations. In a typical measurement of the thermal resistivity, the channel is connected to two thermal reservoirs. There is no charge transfer to the reservoirs and only heat can be exchanged between them and the channel. Hence, the currents of electrons and holes have to vanish at the boundaries. This in particular implies that $j_\delta(\pm\ell/2) = 0$. This leads to the accumulation of electrons and holes at the boundaries. Such accumulation is required to stop the two currents from propagating in the channel. Hence, the imbalance field F_δ can be finite. On the contrary, when the electric resistivity is measured, a charge current is passed through the system and a voltage drop is detected. In this case, the imbalance current needs not to vanish at the boundaries and is in fact uniform throughout the channel. However, since there is no applied imbalance field, F_δ must vanish.

We will start by considering the measurement of the thermal resistivity. Taking the derivative of Eq. (9), in the steady state we get

$$\partial_x^2 j_\delta(x) = \frac{\nu_0}{\tau_r} F_\delta \ . \qquad [10]$$

From the last line of Eq. (8), we get that $F_\delta = \langle\rho_\delta|j_{ns}\rangle + \rho_{\delta\delta} j_\delta$. Using this into Eq. (10), and then solving by imposing the boundary conditions $j_\delta(\pm\ell/2) = 0$, we get

$$j_\delta(x) = -\frac{\langle\rho_\delta|j_{ns}\rangle}{\rho_{\delta\delta}}\big[1-\alpha(\ell,x)\big] \ . \qquad [11]$$

In this equation,

$$\alpha(\ell,x) = \frac{\cosh(x/\ell_D)}{\cosh\big[\ell/(2\ell_D)\big]} \ , \qquad [12]$$

where $\ell_D \equiv \sqrt{\tau_r/(\nu_0\rho_{\delta\delta})}$ is the recombination length. To get Eq. (11), we have assumed that the thermal and particle currents are constant throughout the channel, while the (electric, thermal and imbalance) fields depend on position. This implies that there is no loss of energy along the channel. This is compatible with the system being in the hydrodynamic regime: energy loss occurs via phonon emission, which is however assumed to occur at a much slower rate than electron-electron collisions.

We note that the hydrodynamic hypothesis also explains why ℓ_D can assume drastically different values in compensated semimetals and in, e.g., graphene systems. In a compensated semimetal, electron and hole Fermi surfaces are centered at distant points of the Brillouin zone. Electron-hole recombination occurs at the Fermi surface and requires a large transfer of momentum, much larger than the typical Fermi momenta of the involved particles. Therefore, it requires momentum-non-conserving scattering process to be effective in equilibrating particles and holes with each other. But this is exactly what is prevented in the hydrodynamic regime of transport, in which momentum-non-conserving collisions with impurities or phonons seldom occur. Hence, τ_r becomes very large, while electrons and holes are largely independent. When the recombination time $\tau_r \to \infty$, ℓ_D diverges and $\alpha(\ell,x) \to 1$. Under this condition, the imbalance current of Eq. (11) vanishes everywhere and the system behaves as two independent unipolar systems.

On the contrary in, e.g. graphene systems, electron-hole recombination (and therefore the equalization of their chemical potentials) occurs at a much faster rate, with typical time scales of few tens of femtoseconds. Hence, the typical relaxation times for imbalances in chemical potential are very short, *i.e.* $\tau_r \to 0$. In this case, ℓ_D vanishes and $\alpha(\ell,x) \to 0$. Since the imbalance current can be finite, the system displays ambipolar behavior. We stress that, in graphene, electron-hole recombination occurs in general much faster than cooling, which has typical time scales of few picoseconds (29).

Substituting the result of Eq. (11) into the first line of Eq. (8), we obtain an equation of the form $|F_{ns}\rangle = \tilde\rho(x)|j_{ns}\rangle$, where the 2×2 position-dependent thermoelectric matrix is

$$\tilde\rho(x) = \hat\rho_{ns} - \frac{|\rho_\delta\rangle\langle\rho_\delta|}{\rho_{\delta\delta}}\big[1-\alpha(\ell,x)\big] \ . \qquad [13]$$

According to the discussion above, from Eq. (13) we can define the unipolar and ambipolar resistivity matrices as

$$\begin{aligned}\hat\rho_{\rm uni} &\equiv \lim_{\alpha(\ell,x)\to1}\tilde\rho(x) = \hat\rho_{ns} \ ,\\ \hat\rho_{\rm ambi} &\equiv \lim_{\alpha(\ell,x)\to0}\tilde\rho(x) = \hat\rho_{ns} - \frac{|\rho_\delta\rangle\langle\rho_\delta|}{\rho_{\delta\delta}} \ . \end{aligned} \qquad [14]$$

Each infinitesimally thin slice of the channel at position x contributes a resistivity $\tilde\rho(x)$, which is in series to those of all other slices. The total resistivity of the channel is therefore obtained by summing the resistivities of the infinitesimally thin slices that compose it, and dividing the result by its total length ℓ. This is equivalent to averaging Eq. (13) over the length of the channel. We finally obtain the sought 2×2 thermoelectric matrix subject to thermal-measurement boundary conditions

$$\hat\rho = \hat\rho_{\rm uni}\alpha(\ell) + \hat\rho_{\rm ambi}\big[1-\alpha(\ell)\big] \ . \qquad [15]$$

Note that we can easily add a contribution due to momentum-non-conserving processes $\hat\rho_{\rm D}$ in series to $\hat\rho$ by replacing $\hat\rho_{\rm uni} \to \hat\rho_{\rm uni} + \hat\rho_{\rm D}$ and $\hat\rho_{\rm ambi} \to \hat\rho_{\rm ambi} + \hat\rho_{\rm D}$. Comparing Eq. (15) with the definition (7), we immediately identify the thermal resistivity

$$\rho_{\rm th} = \alpha(\ell)\rho_{\rm th,uni} + \big[1-\alpha(\ell)\big]\rho_{\rm th,ambi} \ . \qquad [16]$$

Using Eqs. (8) and (15) we can also derive the Seebeck coefficient as

$$Q = \frac{\alpha(\ell)\rho_{\rm th,uni}Q_{\rm uni} + \big[1-\alpha(\ell)\big]\rho_{\rm th,ambi}Q_{\rm ambi}}{\alpha(\ell)\rho_{\rm th,uni} + \big[1-\alpha(\ell)\big]\rho_{\rm th,ambi}} \ , \qquad [17]$$

where $Q_{\rm uni}$ and $Q_{\rm ambi}$ are the Seebeck coefficients of the system in the unipolar and ambipolar regimes, respectively.

To derive the electric resistivity, we have to start again from Eq. (8) and apply the boundary condition $F_\delta = 0$. Since we impose no condition on the imbalance current, the latter has a uniform value which is determined by the applied fields. In this case, the electric resistivity is simply that of the ambipolar channel, *i.e.* $\rho_{\rm el} = \rho_{\rm el,ambi}$. In fact, given that the boundary conditions do no treat particles and holes separately, there is no such thing as a "unipolar" electric resistivity. This can be seen also mathematically, by setting $F_\delta = 0$ in Eq. (8) and

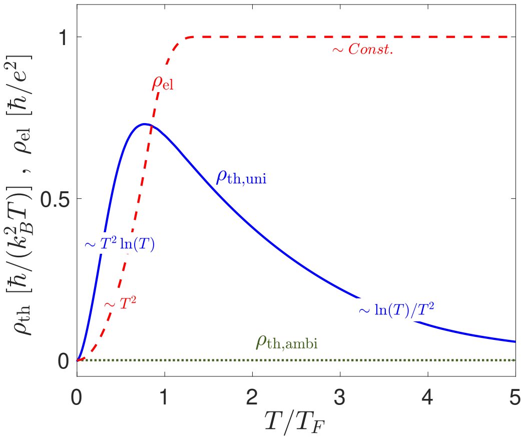


Fig. 3. Intrinsic unipolar and ambipolar thermal resistivities (as defined in the text) as well as the intrinsic electric resistivity as functions of temperature (scaled with T_F). In the presence of disorder the total electric and thermal resistivities, respectively, become $\rho_{el} + \rho_{\rm el,dis}$ and $\rho_{\rm th,uni/ambi} + \rho_{\rm th,dis}$ (not shown in the figure), where $\rho_{\rm el,dis}$ is constant at low T (impurity-dominated regime) and linearly scales with T when phonons are relevant. Making use of the Wiedemann-Franz law for a non-interacting disordered system, we have assumed $\rho_{\rm el,dis}/\rho_{\rm th,dis} = \pi^2/3$ (in reduced units).

solving its last line. The result is $j_\delta = -\langle \rho_\delta | j_{ns}\rangle / \rho_{\delta\delta}$, which is identical to Eq. (11) in the limit $\alpha(\ell, x) \to 0$. Therefore, as expected, the system behaves as purely ambipolar and $\hat{\rho} = \hat{\rho}_{\rm ambi}$. Therefore, the Lorenz ratio reads

$$L = \frac{\rho_{\rm el}}{\alpha(\ell)\rho_{\rm th,uni} + \big[1-\alpha(\ell)\big]\rho_{\rm th,ambi}}\,. \qquad [18]$$

In the absence of interactions and in the unipolar regime, L tends to L_0, the value prescribed by the Wiedemann-Franz law.

4. Numerical Results and Discussion

In this section we present numerical results for Eqs. (16), (17), and (18) for the thermal resistivity, the Seebeck coefficient, and the Lorenz ratio, respectively. We assume the semimetal to be perfectly compensated, *i.e.* the number of electrons equals the number of holes. For the sake of clarity, we start by discussing the electrical and thermal resistivities in the ideal intrinsic limit, in which momentum-non-conserving interactions are completely neglected. These will be re-introduced later, in order to regularize the results pertaining to the Lorenz ratio.

Our results are summarized in Fig. 3. We plot the dimensionless electrical and thermal resistivities, $\rho_{\rm el}$ and $\rho_{\rm th}$, as functions of temperature. The thermal resistivities $\rho_{\rm th,uni}$ and $\rho_{\rm th,ambi}$, subject to homogeneous unipolar and ambipolar boundary conditions, are plotted as solid and dotted lines respectively. As discussed above, the unipolar boundary condition sets electron and hole currents, j_e and j_h, individually to zero. This is equivalent to requiring that both the electric and the imbalance currents vanish. On the other hand, the ambipolar boundary condition sets only the total current $j_e + j_h$ to zero. Note that, as explained in the previous section, the distinction between unipolar and ambipolar boundary conditions does not apply to the electric resistivity. In Fig. 3 we see that, while the ambipolar thermal resistivity vanishes, the unipolar thermal resistivity remains finite. In the latter case the system is essentially equivalent to two independent electron and hole fluids with no electron-hole interactions.

Since our model enjoys particle-hole symmetry, the thermal conductivity of electrons and holes are equal. Thus, $\rho_{\rm th,uni} = 1/(2\kappa_{\rm e})$. We use the well-established kinetic-equation methods of Refs. (16) and (17) to calculate the intrinsic thermal conductivity $\kappa_{\rm e}$ of a single parabolic band at a fixed carrier density, defined by the Fermi energy ε_F. These techniques can be viewed as a simplified version of the calculations performed in Ref. (26). We rely on a simpler Ansatz for the non-equilibrium distribution function (16) and find $\rho_{\rm th,uni} \simeq I_{\rm th}/(2D_{\rm th}^2)$, where $D_{\rm th} = \sum_{\boldsymbol{k}} (\partial f_{\boldsymbol{k}}/\partial\epsilon_{\boldsymbol{k}}) \boldsymbol{v}_{\boldsymbol{k}} \cdot \boldsymbol{v}_{\boldsymbol{k}} [(\epsilon_{\boldsymbol{k}} - \varepsilon_F)/k_B T]^2 \simeq 9\zeta(3)/(4\pi\hbar^2\beta)$ is the thermal Drude weight, while $I_{\rm th}$ is the Coulomb collision integral projected onto the thermal channel. In the degenerate Fermi liquid regime ($T \ll T_F$), $D_{\rm th} \sim T$ and $I_{\rm th} \sim T^4$ we find*

$$\rho_{\rm th,uni}(T \ll T_F) \sim T^2 \ln(T). \qquad [21]$$

This behavior of the thermal resistivity is consistent with previous results obtained for degenerate Fermi liquid graphene (22). We find that $\rho_{\rm th,uni}$ peaks around $T \simeq T_F$ and decreases as $\sim \ln(T)/T^2$ for $T \gg T_F$ (*i.e.* in the non-degenerate regime), see blue solid curve in Fig. 3.

Next, we study the the temperature dependence of the intrinsic electric resistivity $\rho_{\rm el}$ of the compensated semimetal. We distinguish the *Planckian* regime, in which there is only one energy scale $k_B T$ ($T \gg T_F$), from the Fermi-liquid one, in which there are two energy scales, $k_B T$ and ε_F. The temperature dependence of $\rho_{\rm el}$ (in reduced units) is given by,

$$\rho_{\rm el} \sim \frac{1}{D_{\rm el}(T)\tau_{\rm eh}(T)}, \qquad [22]$$

where $\tau_{\rm eh}(T)$ is the electron-hole scattering rate and $D_{\rm el}(T)$ the Drude weight in the electric channel, which is defined as $D_{\rm el} = \sum_{\boldsymbol{k}} \boldsymbol{v}_{\boldsymbol{k}}^2 (\partial f_{\boldsymbol{k}}/\partial\epsilon_{\boldsymbol{k}})$. In the Planckian regime the maximum scattering rate allowed by the energy-time uncertainty principle (31) is $1/\tau_{\rm eh}(T) \sim k_B T/\hbar$ while in the Fermi liquid regime $1/\tau_{\rm eh}(T) \sim (k_B T)^2/\hbar\varepsilon_F$. Similarly, in the Planckian regime the Drude weight $D_{\rm el} \propto k_B T$, whereas in the Fermi liquid regime it is independent of temperature. Hence, the intrinsic resistivity $\rho_{\rm el}$ is independent of temperature in the Planckian regime, while it scales as $\sim T^2$ in the Fermi liquid regime. The red dashed curve in Fig. 3 shows precisely these limiting behaviors.

We now introduce momentum-non-conserving interactions and discuss the calculation of the Lorenz ratio of Eq. (18).

---

[*] For a system with a single parabolic band, the velocity $v_{\boldsymbol{k}} \sim \boldsymbol{k}$ coincides with the momentum zero mode, and therefore only the thermal moment of the collision integral, $I_{\rm th}$, associated with the relaxation of energy (thermal) currents, survives. This implies that, while the electric resistivity is exactly zero, the thermal resistivity remains finite. We make use of standard approximations for the Coulomb collision integral (screened interaction plus Fermi golden rule), previously used for graphene systems (16, 17), and find $I_{\rm th}$ to be

$$I_{\rm th} = -\frac{1}{4\pi(k_B T)^3} \sum_{\boldsymbol{q}} \int_{-\infty}^{\infty} d\omega \frac{|V(\boldsymbol{q},\omega)|^2}{\sinh^2(\hbar\omega/2k_B T)} [(\Im\Pi_1)^2 - \Im\Pi_0 \Im\Pi_2], \qquad [19]$$

where $V(\mathbf{q},\omega) = v_q/|1 - v_q\Pi_0(\mathbf{q},\omega)|$ is the screened electron-electron Coulomb interaction and $v_q = 2\pi e^2/(\epsilon q)$. Here, ϵ is the dielectric constant that accounts for the surrounding medium as well as screening from remote bands. We set $\epsilon = 1$ in our calculation. The response functions $\Pi_n(\mathbf{q},\omega)$ are defined as

$$\Pi_n(\mathbf{q},\omega) = 2\sum_{\boldsymbol{k}} \frac{(\tilde{\epsilon}_{\boldsymbol{k}} v_{\boldsymbol{k}} - \tilde{\epsilon}_{\boldsymbol{k}+\boldsymbol{q}} v_{\boldsymbol{k}+\boldsymbol{q}})^n (f^0_{\boldsymbol{k}} - f^0_{\boldsymbol{k}+\boldsymbol{q}})}{\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{k}+\boldsymbol{q}} + \hbar\omega + i0^+}, \qquad [20]$$

where $\tilde{\epsilon}_{\boldsymbol{k}} = \epsilon_{\boldsymbol{k}} - \varepsilon_F$ is the band energy measured from the Fermi energy. At $T \ll T_F$ (degenerate regime), Π_0 is the well-known zero-temperature 2D Lindhard function (30) and we find that for $\hbar\omega \ll \varepsilon_F$, $[\Pi_1^2 - \Pi_0\Pi_2] \sim \omega^6$. This results in $I_{\rm th} \sim T^4$.

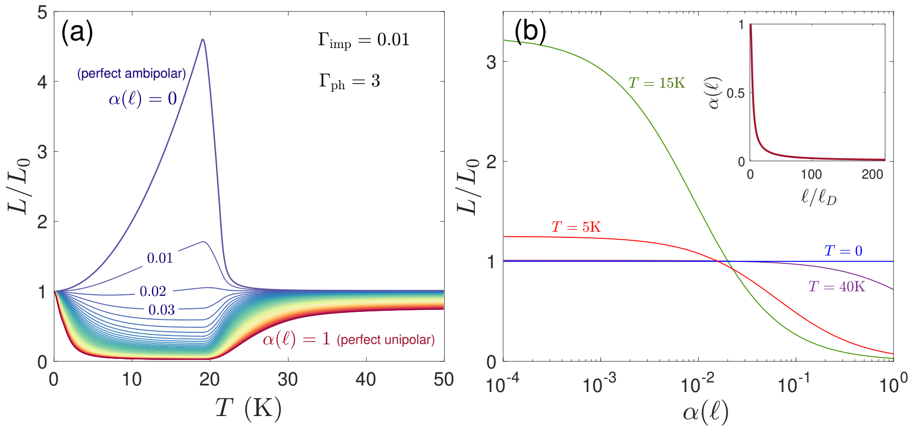


Fig. 4. Lorenz ratio as a function of (a) temperature for different values of $\alpha(\ell)$ from 0 (perfect ambipolar regime) to 1 (perfect unipolar regime) with a step of 0.01 and (b) as a function of $\alpha(\ell)$ (in logarithmic scale) for different temperatures as labeled. The strength of the charge impurity- and phonon-limited resistivities were defined through the hydrodynamicity parameters $\Gamma^2_{\rm imp} = \rho_{\rm el,imp}/\rho_{\rm el}(T \gg T_F)$ and $\Gamma^2_{\rm ph} = \rho_{\rm el,ph}/\rho_{\rm el}(T \gg T_F)$, respectively. The inset in (b) shows $\alpha(\ell)$, Eq. (5), as a function of ℓ/ℓ_D. The onset temperature for phonons is taken as $T_{\rm ph} = 25$K. In the hydrodynamic regime, i.e. $0 \lesssim T \lesssim 25$K in this figure, the Lorenz ratio deviates from its standard (impurity/phonon)-limited value L_0. Due to the ambipolar constraint $j_e + j_h = 0$, the violation of the Lorenz ratio is a large enhancement while with the unipolar situation ($j_e = j_h = 0$) the Lorenz number is drastically reduced below L_0.

Momentum-non-conserving scattering is necessary to regularize results in the ambipolar limit: as it is clear from Fig. 3 and the definition 18, the ratio between the electrical and thermal resistivity ($\rho_{\rm th,ambi}$) would diverge if the contribution of disorder were neglected. To include disorder we consider the following simple but realistic model. At low temperatures impurities are the dominant disorder mechanism while, as temperature increases, electron-phonon scatterings become more important.

The Drude resistivity due to scattering against impurities is here defined as $\rho_{\rm el,imp} = m^*/(n\tau_{\rm imp})$, where m^* and n are the electron effective mass and density, respectively. The impurity scattering rate $1/\tau_{\rm imp}$ is assumed to be independent of temperature for both short- and long-range impurities. Since the particle density n and the effective masses are fixed in compensated semimetals, once the electron and hole Fermi energies are set, the electric resistivity due to scattering against impurities, $\rho_{\rm el,imp}$, is independent of T.

As the temperature increases, the resistivity due to collisions with phonons $\rho_{\rm el,ph}$ becomes the dominant contribution to the total electric resistivity. Above the Bloch-Grüneisen temperature, $\rho_{\rm ph}$ increases linearly with T (32). We therefore posit the following model for the momentum-non-conserving scattering: $\rho_{\rm el,dis}(T \lesssim T_{\rm ph}) \simeq \rho_{\rm el,imp} \sim$ Const. and $\rho_{\rm el,dis}(T \gtrsim T_{\rm ph}) \simeq \rho_{\rm el,ph} \sim T$, where $T_{\rm ph}$ is defined as the onset temperature at which phonons start to become the dominant scattering mechanism. Assuming that the Wiedemann-Franz law is satisfied when only momentum-non-conserving (electron-impurity or electron-phonon) processes are taken into account, we obtain in particular that the thermal resistivity of impurities in reduced units is $\rho_{\rm th,imp} = (3/\pi^2)\rho_{\rm el,imp}$.

Figures 4a and 4b show the results for the Lorenz ratio of Eq. (18) as a function of (a) temperature for different values of $\alpha(\ell)$ (corresponding to different sample lengths) and (b) as a function of $\alpha(\ell)$ for different temperatures. The onset temperature for phonon-dominated scattering is taken to be $T_{\rm ph} \simeq 25$ K. We determine the strength the charge impurity as well as phonon resistivities, respectively through the hydrodynamicity parameters $\Gamma^2_{\rm imp} = \rho_{\rm imp}/\rho_{\rm el}(T \gg T_F)$, and $\Gamma^2_{\rm ph} = \rho_{\rm ph}/\rho_{\rm el}(T \gg T_F)$ (i.e. the ratio of the charge impurity/phonon resistivities to the intrinsic Coulomb resistivity in the non-degenerate regime, $T \gg T_F$). For the results in Figs. 4a and 4b we have taken $\Gamma_{\rm imp} = 0.01$ and $\Gamma_{\rm ph} = 3$.

We observe that the Wiedemann-Franz law is violated in two radically different ways, depending on whether we are in the ambipolar ($\ell \gg \ell_D$) or unipolar ($\ell \ll \ell_D$) limit. While in the former we observe a large enhancement of the Lorenz ratio, in the latter we observe a moderate reduction. The ambipolar limit is the situation realized in graphene systems (see Fig. 1b), while the unipolar one occurs in compensated semimetals as WP$_2$ (see Fig. 1c). We stress that the large enhancement of the Lorenz ratio cannot be explained without taking into account strong electron-hole scattering in the electric conduction channel.

Fig. 5 displays a 2D plot of L/L_0 as functions of temperature and scaled sample length (ℓ/ℓ_D). Based on the behavior of the Lorenz ratio we identify a phase diagram of possible transport regimes in a charge-neutral system. When $T \to 0$ as well as for temperatures $T \gtrsim T_{\rm ph}$, $L/L_0 \to 1$ resulting from the disorder-limited transport in these regimes, i.e. impurity-dominated at $T \to 0$ and phonon-dominated at $T \gtrsim T_{\rm ph}$). In the hydrodynamic regime ($0 \lesssim T \lesssim T_{\rm ph}$), one can tune the WF ratio from an enhancement when $\ell \gg \ell_D$ (bipolar condition) to a reduction when $\ell \ll \ell_D$ (unipolar condition).

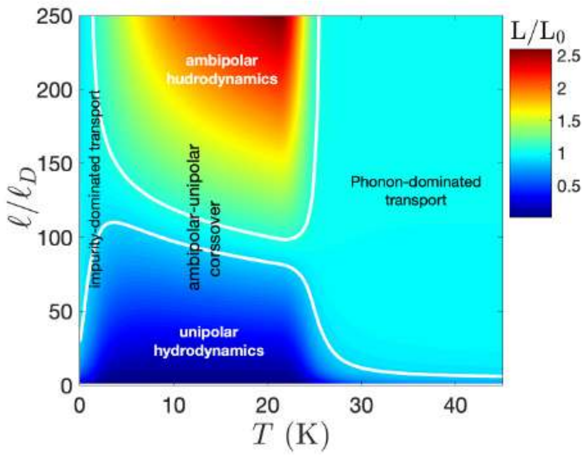


Fig. 5. 2D plot of the Lorenz ratio as a function of ℓ/ℓ_D (sample length scaled with the diffusion length) and T. Different transport regimes are indicated on the figure. The onset temperature of phonons $T_{\rm ph} = 25$K and the hydrodynamicity parameters $\Gamma_{\rm imp} = 0.01$ and $\Gamma_{\rm ph} = 3$ were chosen to be the same as in Fig. 4. The enhancement of the Lorenz ratio (in the ambipolar hydrodynamic regime) is relevant for graphene systems and is a consequence of electron-hole scattering, which selectively enhances the electric resistivity. The reduced Lorenz number (in the unipolar hydrodynamic regime) is relevant for compensated semimetals, where the electron and hole bands are well separated in momentum space.

Finally, we calculate the Seebeck coefficient according to Eq. (17). For a symmetric electron-hole system ($Q_e = -Q_h$), the Seebeck coefficient always vanishes at the charge neutrality point, both in the absence and in the presence of disorder. Note that in the intrinsic regime at perfect compensation, in which momentum-non-conserving processes are absent, $\rho_{\rm th,ambi} = 0$. This in turn implies that $\rho_{\rm th} = \alpha(\ell)\rho_{\rm th,uni}$ and $Q = Q_{\rm uni}$. In this case, while the thermal resistivity depends on the system size, the Seebeck coefficient is independent of it.

5. Outlook

The breakdown of the Wiedemann-Franz law, which results in an anomalously large Lorenz ratio near the charge neutrality has been recognized to occur in clean graphene samples and has inspired a considerable amount of theoretical work. On the contrary, experiments in three-dimensional semimetals such as WP_2 and antimony show a radically different result. Although a phenomenology similar to that of graphene would naively be expected, a puzzling *reduced* Lorenz ratio is observed at low temperatures, when both electrons and holes form degenerate Fermi liquids.

In this study, we have shown that this apparent contradiction is explained by the completely different transport situations realized in the two systems: truly ambipolar transport in graphene versus two independent channels of unipolar transport in compensated semimetals. In contrast to graphene and its bilayer, electron and hole pockets in compensated semimetals are well distanced in momentum space, resulting in a long recombination time. Since both electron and hole currents must separately vanish at the contacts, this results in a suppression of the bulk ambipolar conduction mechanism. Effectively, electrons and holes behave as two independent and decoupled Fermi liquid throughout the channel (a situation analogous to unipolar transport).

The violation of the Wiedemann-Franz law in both the ambipolar and unipolar transport regimes occurs in a temperature window in which the so-called *hydrodynamic* regime of transport is realized, *i.e.* when momentum-conserving collisions among particles constitute the dominant scattering mechanism (we note that in a Fermi liquid the WF law is satisfied when disorder scattering dominates). We have presented a simple theory for a general unipolar/ambipolar system and demonstrated a crossover from ambipolar to unipolar conduction as a function of a weight function (related to the electron-hole recombination time) which naturally leads to a sample-size-dependent thermal conductivity as observed in experiments. Although our theory has been presented for a two-dimensional system, the results would qualitatively remain valid for three-dimensional ones.

6. Acknowledgments

We thank Kamran Behnia and Joseph Heremans for useful discussions. M.Z. and G.V. were supported by the U.S. Department of Energy (Office of Science) under grant No. DE-FG02-05ER46203. A.P. was supported by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS).

7. References

1. D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, L. A. Ponomarenko, A. K. Geim, and M. Polini. Negative local resistance caused by viscous electron backflow in graphene. *Science*, 351(6277):1055–1058, 2016. ISSN 0036-8075. .
2. Fereshte Ghahari, Hong-Yi Xie, Takashi Taniguchi, Kenji Watanabe, Matthew S. Foster, and Philip Kim. Enhanced thermoelectric power in graphene: Violation of the mott relation by inelastic scattering. *Phys. Rev. Lett.*, 116:136802, Mar 2016. . URL https://link.aps.org/doi/10.1103/PhysRevLett.116.136802.
3. Jesse Crossno, Jing K. Shi, Ke Wang, Xiaomeng Liu, Achim Harzheim, Andrew Lucas, Subir Sachdev, Philip Kim, Takashi Taniguchi, Kenji Watanabe, Thomas A. Ohki, and Kin Chung Fong. Observation of the dirac fluid and the breakdown of the wiedemann-franz law in graphene. *Science*, 2016. ISSN 0036-8075. .
4. Alessandro Principi, Giovanni Vignale, Matteo Carrega, and Marco Polini. Bulk and shear viscosities of the two-dimensional electron liquid in a doped graphene sheet. *Phys. Rev. B*, 93:125410, Mar 2016. . URL https://link.aps.org/doi/10.1103/PhysRevB.93.125410.
5. B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin. Hydrodynamics in graphene: Linear-response transport. *Phys. Rev. B*, 91:035414, Jan 2015. . URL https://link.aps.org/doi/10.1103/PhysRevB.91.035414.
6. U. Briskot, M. Schütt, I. V. Gornyi, M. Titov, B. N. Narozhny, and A. D. Mirlin. Collision-dominated nonlinear hydrodynamics in graphene. *Phys. Rev. B*, 92:115426, Sep 2015. . URL https://link.aps.org/doi/10.1103/PhysRevB.92.115426.
7. Lars Fritz, Jörg Schmalian, Markus Müller, and Subir Sachdev. Quantum critical transport in clean graphene. *Phys. Rev. B*, 78:085416, Aug 2008. . URL https://link.aps.org/doi/10.1103/PhysRevB.78.085416.
8. Markus Müller, Lars Fritz, and Subir Sachdev. Quantum-critical relativistic magnetotransport in graphene. *Phys. Rev. B*, 78:115406, Sep 2008. . URL https://link.aps.org/doi/10.1103/PhysRevB.78.115406.
9. Hong-Yi Xie and Matthew S. Foster. Transport coefficients of graphene: Interplay of impurity scattering, coulomb interaction, and optical phonons. *Phys. Rev. B*, 93:195103, May 2016. . URL https://link.aps.org/doi/10.1103/PhysRevB.93.195103.
10. D. Svintsov, V. Vyurkov, V. Ryzhii, and T. Otsuji. Hydrodynamic electron transport and nonlinear waves in graphene. *Phys. Rev. B*, 88:245444, Dec 2013. . URL https://link.aps.org/doi/10.1103/PhysRevB.88.245444.
11. D. Svintsov. Hydrodynamic-to-ballistic crossover in dirac materials. *Phys. Rev. B*, 97:121405, Mar 2018. . URL https://link.aps.org/doi/10.1103/PhysRevB.97.121405.
12. Derek Y. H. Ho, Indra Yudhistira, Nilotpal Chakraborty, and Shaffique Adam. Theoretical determination of hydrodynamic window in monolayer and bilayer graphene from scattering rates. *Phys. Rev. B*, 97:121404, Mar 2018. . URL https://link.aps.org/doi/10.1103/PhysRevB.97.121404.
13. Alessandro Principi and Giovanni Vignale. Intrinsic charge and spin conductivities of doped graphene in the fermi-liquid regime. *Phys. Rev. B*, 91:205423, May 2015. . URL https://link.aps.org/doi/10.1103/PhysRevB.91.205423.
14. Andrew Lucas and Sankar Das Sarma. Electronic hydrodynamics and the breakdown of the wiedemann-franz and mott laws in interacting metals. *Phys. Rev. B*, 97:245128, Jun 2018. . URL https://link.aps.org/doi/10.1103/PhysRevB.97.245128.
15. Andrew Lucas and Sean A. Hartnoll. Kinetic theory of transport for inhomogeneous electron fluids. *Phys. Rev. B*, 97:045105, Jan 2018. . URL https://link.aps.org/doi/10.1103/PhysRevB.97.045105.
16. Mohammad Zarenia, Alessandro Principi, and Giovanni Vignale. Disorder-enabled hydrodynamics of charge and heat transport in monolayer graphene. *2D Materials*, 6(3):035024, may 2019. . URL https://doi.org/10.1088%2F2053-1583%2Fab1ad9.
17. Mohammad Zarenia, Thomas Benjamin Smith, Alessandro Principi, and Giovanni Vignale. Breakdown of the wiedemann-franz law in ab-stacked bilayer graphene. *Phys. Rev. B*, 99: 161407, Apr 2019. . URL https://link.aps.org/doi/10.1103/PhysRevB.99.161407.
18. Mohammad Zarenia, Indra Yudhistira, Shaffique Adam, and Giovanni Vignale. Enhanced hydrodynamic transport in near magic angle twisted bilayer graphene. *Phys. Rev. B*, 101: 045421, Jan 2020. . URL https://link.aps.org/doi/10.1103/PhysRevB.101.045421.
19. Mohammad Zarenia, Shaffique Adam, and Giovanni Vignale. Temperature collapse of the electric conductivity in bilayer graphene. *Phys. Rev. Research*, 2:023391, Jun 2020. . URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.023391.
20. Alexandre Jaoui, B. Fauqué, Carl Willem Rischau, Alaska Subedi, Chenguang Fu, Johannes Gooth, Nitesh Kumar, Vicky Süβ, Dmitrii L. Maslov, Claudia Felser, and Kamran Behnia. Departure from the wiedemann-franz law in wp2 driven by mismatch in t-square resistivity prefactors. *npj Quantum Materials*, 3(1):64, Dec 2018. ISSN 2397-4648. . URL https://doi.org/10.1038/s41535-018-0136-x.
21. J. Gooth, F. Menges, N. Kumar, V. Süβ, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann. Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide. *Nature Communications*, 9(1):4093, Oct 2018. ISSN 2041-1723. . URL https://doi.org/10.1038/s41467-018-06688-y.
22. Alessandro Principi and Giovanni Vignale. Violation of the wiedemann-franz law in hydrodynamic electron liquids. *Phys. Rev. Lett.*, 115:056603, Jul 2015. . URL https://link.aps.org/doi/10.1103/PhysRevLett.115.056603.
23. Alexandre Jaoui, Benoît Fauqué, and Kamran Behnia. Thermal resistivity due to electron viscosity in bulk antimony, 2020.
24. C. F. Gallo, B. S. Chandrasekhar, and P. H. Sutter. Transport properties of bismuth single crystals. *Journal of Applied Physics*, 34(1):144–152, 1963. . URL https://doi.org/10.1063/1.1729056.
25. C Uher and H. J. Goldsmid. Separation of the electronic and lattice thermal conductivities in bismuth crystals. *Phys. Stat. Sol.*, 65:765, 1974.
26. Songci Li and Dmitrii L. Maslov. Lorentz ratio of a compensated metal. *Phys. Rev. B*, 98: 245134, Dec 2018. . URL https://link.aps.org/doi/10.1103/PhysRevB.98.245134.
27. D. Pines and P. Nozières. *The Theory of Quantum Liquids: Normal Fermi liquids.* W.A. Benjamin, 1966. URL https://books.google.com.sg/books?id=GP1QAAAAMAAJ.
28. Ben G. Streetman. *Solid State Electronic Devices.* Prentice-Hall, Inc., USA, 1990. ISBN 0138229414.
29. Yuzhong Chen, Yujie Li, Yida Zhao, Hongzhi Zhou, and Haiming Zhu. Highly efficient hot

electron harvesting from graphene before electron-hole thermalization. *Science Advances*, 5 (11), 2019. . URL https://advances.sciencemag.org/content/5/11/eaax9958.

30. Gabriele Giuliani and Giovanni Vignale. *Quantum Theory of the Electron Liquid*. Cambridge University Press, 2005. .
31. Sean A. Hartnoll. Theory of universal incoherent metallic transport. *Nature Physics*, 11(1): 54–61, Jan 2015. ISSN 1745-2481. . URL https://doi.org/10.1038/nphys3174.
32. Safa Kasap, Cyril Koughia, and Harry E. Ruda. *Electrical Conduction in Metals and Semiconductors*, pages 1–1. Springer International Publishing, Cham, 2017. ISBN 978-3-319-48933-9. . URL https://doi.org/10.1007/978-3-319-48933-9_2.

Appendix

A. Explicitly accounting for momentum: the four-mode theory

In this appendix we consider the general case in which the total momentum, the conserved mode of hydrodynamic conduction, does not coincide with either the electric, thermal or imbalance currents. Given its importance in determining the transport properties of the system, it is necessary to include it explicitly. In the resulting four-mode description, the total momentum is added to the list of currents flowing in the system. To stress the fact that it is a zero mode of the resulting 4×4 resistivity matrix, *i.e.* a nontrivial vector with eigenvalue zero, we will call the total momentum j_0. A force F_0 that couples explicitly to it will also be included. The goal of this appendix is therefore to show how the 4×4 resistivity matrix that connects the four currents to the four fields can be down-folded to obtain the 3×3 matrix of Eq. (8). We will guide the reader through this process and show that the determinant of the resulting resistivity matrix vanishes, as stated in Sect. 3. This in turn implies that the zero mode, although not explicit, is still included in the three-mode theory. Therefore, no information about the physical implications of the conservation of momentum is lost in the down-folding process.

The derivation here parallels that given in Sect. 3. Fields and currents are now related by the 4×4 resistivity matrix $\hat{\rho}$. Explicitly,

$$\left(\begin{array}{c} F_0 \\ \hline |F_{ns\delta}\rangle \end{array}\right) = \left(\begin{array}{c|c} \rho_{00} & \langle\rho_0| \\ \hline |\rho_0\rangle & \hat{\rho}_{ns\delta} \end{array}\right) \left(\begin{array}{c} j_0 \\ \hline |j_{ns\delta}\rangle \end{array}\right) . \qquad [23]$$

Here $|j_{ns\delta}\rangle = {}^t(j_n, j_s, j_\delta)$ and $|F_{ns\delta}\rangle = {}^t(-eE, -k_{\rm B}\partial_x T, F_\delta)$, $\hat{\rho}_{ns\delta}$ is a 3×3 block, and $|\rho_0\rangle$ is a three-component vector. The determinant of the 4×4 matrix in Eq. (23) is

$$\det(\hat{\rho}) = \det(\hat{\rho}_{ns\delta})\left[\rho_{00} - \langle\rho_0|\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle\right] , \qquad [24]$$

and it is equal to zero since there is at least one nontrivial vector with eigenvalue zero (*i.e.* the total momentum). Indeed, if the system is prepared in a state such that only $j_0 \neq 0$, the vector on the left-hand side of Eq. (23) is a null vector. By assumption, the only zero mode is the momentum, so the determinant of the 3×3 block $\hat{\rho}_{ns\delta}$ is finite and such matrix is therefore invertible. Thus, it must be that

$$\rho_{00} = \langle\rho_0|\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle . \qquad [25]$$

The property (25) plays a pivotal role in the following proof.

To down-fold the four-mode theory of Eq. (23) into the three-mode one of Eq. (8) we need to apply the boundary conditions on the momentum. Since no external field that couples specifically to the momentum is applied, we will set $F_0 = 0$, while the momentum j_0 is allowed to assume an arbitrary value. The latter is determined by the values of the electric and thermal currents and fields, as well as by the boundary conditions imposed on the imbalance current. The first line of Eq. (23) implies that

$$j_0 = -\frac{\langle\rho_0|j_{ns\delta}\rangle}{\rho_{00}} . \qquad [26]$$

When this relation is substituted into the last line of Eq. (23) we get

$$|F_{ns\delta}\rangle = \left[\hat{\rho}_{ns\delta} - \frac{|\rho_0\rangle\langle\rho_0|}{\rho_{00}}\right]|j_{ns\delta}\rangle . \qquad [27]$$

The matrix in square brackets on the right-hand side of Eq. (27) is the 3×3 resistivity matrix of Eq. (8). We now prove that its determinant is zero. To do so, it is sufficient to show that there exists a nontrivial vector $|j_{ns\delta}\rangle$ such that, when the matrix acts on it, the result is exactly zero. It is easy to see that such vector is $\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle$. Indeed, using the property (25), we have

$$\begin{aligned} |F_{ns\delta}\rangle &= \left[\hat{\rho}_{ns\delta} - \frac{|\rho_0\rangle\langle\rho_0|}{\rho_{00}}\right]\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle \\ &= |\rho_0\rangle - \frac{|\rho_0\rangle\langle\rho_0|\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle}{\langle\rho_0|\hat{\rho}^{-1}_{ns\delta}|\rho_0\rangle} = 0 , \end{aligned} \qquad [28]$$

which proves the assertion. Hence, the 3×3 resistivity matrix of Eq. (8) can be assumed to have determinant equal to zero and to implicitly retain the information about the conservation of momentum by electron-electron interactions.

Interplay between two mechanisms of resistivity

A Kapustin and G Falkovich

(Dated: July 24, 2024)

Abstract

Mechanisms of resistivity can be divided into two basic classes: one is dissipative (like scattering on phonons) and another is quasi-elastic (like scattering on static impurities). They are often treated by the empirical Matthiessen rule, which says that total resistivity is just the sum of these two contributions, which are computed separately. This is quite misleading for two reasons. First, the two mechanisms are generally correlated. Second, computing the elastic resistivity alone masks the fundamental fact that the linear-response approximation has a vanishing validity interval at vanishing dissipation. Limits of zero electric field and zero dissipation do not commute for the simple reason that one needs to absorb the Joule heat quadratic in the applied field. Here, we present a simple model that illustrates these two points. The model also illuminates the role of variational principles for non-equilibrium steady states.

I. INTRODUCTION

This is a methodological note intended to explain the basic interplay between the two mechanisms of resistivity using a simple model analyzed in more detail in [1].

Consider a classical particle that moves under the influence of a uniform force E in a medium with the temperature T and randomly placed elastic scatterers. We denote the momentum relaxation rate due to elastic scattering (averaged over all momenta) by ρ_e. Finite temperature provides for additional momentum relaxation with the rate ν and for a random force, which leads to diffusion in the momentum space with the diffusivity νT.

We define the resistivity ρ as a linear-response factor relating the mean momentum to the force: $\bar{p} = E/\rho$. It is, thus, the mean relaxation rate of the momentum direction. An empirical Matthiessen rule (M-rule) suggests that (see, e.g. [3])

$$\rho \approx \nu + \rho_e \, . \tag{1}$$

Just how off is the estimate (1)? Below, we show that the rule is exact when the elastic scattering is momentum-independent in two dimensions. We also compute ρ in two limits where one or another mechanism dominates and show that the addition of another mechanism generally enhances resistivity much more than the Matthiessen rule suggests.

It is also clear that computing the elastic contribution in the limit $\nu \to 0$ does not make much sense since energy conservation requires

$$\nu \langle p^2 \rangle = E \cdot \bar{p} = E^2/\rho \, . \tag{2}$$

That means that the E^2-corrections to the linear-response theory diverge as $1/\nu$; that is, nonzero friction is necessary for the elastic resistivity to make sense. The behavior of the nonequilibrium distribution function in this limit iss discussed in detail in [1]. On the contrary, we show below that the validity interval of the linear-response theory expands as $1/\rho_e$ when $\rho_e \to 0$.

II. THE BASIC EQUATION

Let us consider the simplest kinetic (Fokker-Planck) equation on the momentum distribution $f(\mathbf{p}, t)$, which satisfies the equation:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial p^i}(\nu p_i - E_i) f + T\nu \Delta f + \hat{I} f \ . \tag{3}$$

The first term on the right is due to a constant force and linear friction. Random kicks from the thermostat provide diffusion in the momentum space, described by the Laplacian. Let us stress that the ν-terms is the simplest model of a thermostat; our classical consideration only qualitatively corresponds to the scattering of quantum electrons on phonons. On the other hand, it describes a variety of classical stochastic systems. The last term is a linear operator $\hat{I}$ describing elastic scattering.

Statistical isotropy of scattering means that the angular harmonics are eigenfunctions of the scattering operator: $\hat{I}f_l = \gamma_l f_l$. Mean momentum and resistivity are determined by the first-harmonic rate, which we denote $\gamma_1 = \gamma$. It generally depends on p. One universal limit is that of scattering by a small angle proportional to the time of interaction and inversely proportional to p. For a finite-range scattering potential, the time is also $\propto 1/p$ so that the deflection angle is $1/p^2$. The rate of meeting scatterers is proportional to the momentum p. As a result, small-angle scattering leads to angular diffusion: $\hat{I} = \frac{W}{p^3}\Delta_\Omega$ so that $\gamma = (d-1)Wp^{-3}$ [2]. That typically occurs when the average momentum exceeds the potential strength measured by W (proportional to the 2-point correlation function of the potential [1]). For lower momenta, the scattering is by angles of order unity; the rate of momentum loss in many such cases is proportional to the momentum itself or independent of momentum. It is also instructive to consider power-law functions $\gamma(p) \propto p^a$ with different a.

For every momentum, two mechanisms have their scattering rates added according to (3). The first-harmonic correction to $f(\mathbf{p})$ is inversely proportional to the sum of the rates for every momentum p. The conductivity is proportional to the correction integrated over momenta (see e.g. [3]). The total resistivity (inverse conductivity) is then bounded by $\rho \geq \nu + \rho_e$, where $\rho_e = \gamma$ is the average of $\gamma(p)$ over p. Only when both relaxation rates are independent of momenta, we have the M-rule equality: $\rho = \nu + \rho_e$. Here we shall see that in our case, it is enough that elastic scattering is momentum-independent for the Matthiessen rule to hold.

The model is characterized by two dimensionless parameters, $B = \gamma/\nu$ and $F = E/\nu T^{1/2}$. The first one determines the relative role of the two mechanisms of resistivity. The second one characterizes the strength of the field.

Without the external force, $E = 0$, the equation has an equilibrium Maxwell isotropic solution $f_0(p) \propto \exp(-p^2/2T)$, which is independent of ν. It realizes the maximum of entropy $S = -\int f \log f \, d\mathbf{p}$ for a given mean energy $\int (p^2/2) f \, d\mathbf{p}$.

Without scattering, $\hat{I} = 0$, the solution has a Gaussian form for arbitrary E since (3) has a symmetry which shifts $\mathbf{p}$ and $\mathbf{E}$ simultaneously:

$$f_0(p,\theta) = (2\pi)^{-d/2} \exp[-|\mathbf{p} - \mathbf{E}/\nu|^2/2T] \,. \tag{4}$$

The distribution (4) gives a linear current-field relation, $\langle p \rangle = E/\nu$, for any E. It also realizes the entropy maximum under the condition of the energy production-dissipation balance, $\nu\langle p^2/2\rangle = \mathbf{E}\langle \mathbf{p}\rangle$. Indeed, it realizes the extremum of the functional $\int f[-\log f + \lambda(\nu p^2/2 - \mathbf{pE})] \, d\mathbf{p}$. Even though (4) looks like a shifted equilibrium whose entropy does not depend on E, it is a non-equilibrium state with energy dissipation and entropy production.

III. LINEAR RESPONSE AT WEAK ELASTIC SCATTERING

Let us describe the effect of scattering on the distribution and the linear resistivity in the limit $B = \gamma/\nu \to 0$. Even though (4) is valid at arbitrary E, we could compute the $\hat{I}$-corrections to it only in the limit $F = E/\nu T^{1/2} \to 0$. We assume $f = f_0(p) + f_1(\mathbf{p}, E) + f_2(\mathbf{p}, E, \gamma)$, where $f_0 \propto e^{-p^2/2T}$ and $f_1 = f_0(pE/\nu T)\cos\theta$ are given by (4). We assume $f_1 \gg f_2 \propto BF \propto \gamma E$. Substituting it into (3) gives

$$\frac{\partial}{\partial p^i} E_i f_2 - \frac{\partial}{\partial p^i} \nu p_i f_2 - T\nu \frac{1}{p}\frac{\partial}{\partial p} p \frac{\partial f_2}{\partial p} - \frac{T\nu}{p^2}\Delta_\Omega f_2 \approx \hat{I} f_1 = \cos\theta \gamma(p) f_0 \frac{pE}{\nu T} \,. \tag{5}$$

Since we are interested in the contribution of scattering to resistivity, we consider only the first angular harmonic, $f_2 = f_0 \chi(p)\cos\theta$, which satisfies the equation

$$p\frac{\partial \chi}{\partial p} - \frac{T}{p}\frac{\partial \chi}{\partial p} - T\frac{\partial^2 \chi}{\partial p^2} + \frac{T(d-1)\chi}{p^2} = -\gamma(p)\frac{pE}{\nu^2 T} \,. \tag{6}$$

For a general $\gamma(p) \propto p^a$, the solution has different asymptotics for large and small p: $\chi \propto -p^{a+1}E/(a+1)\nu^2 T$ for $p \gg T$ and $\chi \propto -Wp^{a+3}E/[d-1-(a+3)^2]\nu^2 T$ for $p \ll T$. In this case, one needs numerics to compute the solution and the correction to resistivity from weak elastic scattering. Fortunately, for two specific (and physical!) values of a, the solution has a simple power form, $\chi = bp^c$, where b, c are constants to be determined. Only solutions with $b < 0$ make physical sense since the scattering must diminish the current. Action by the first term in (6) gives $-b\nu c p^c \cos\theta$. Action by the third term gives $b\nu c p^{c-2}[2p^2 - cT]\cos\theta$. Action by the fourth term gives $b(d-1)\nu T p^{c-2}\cos\theta$. Then the equation (6) gives

$$b\nu c p^c + b\nu p^{c-2} T \left[d - 1 - c^2\right] = -\gamma(p)\frac{pE}{\nu T} \,. \tag{7}$$

Since the terms on the left have different powers of p, the power-law solution exists only for $c = 0$ and $c = \sqrt{d-1}$. The former case corresponds to the small-angle scattering when $\hat{I} = Wp^{-3}\Delta_\Omega$ and $\gamma(p) = (d-1)Wp^{-3}$. In that case we get

$$f_2 = -\frac{WE}{(\nu T)^2} f_0 \cos\theta \,. \tag{8}$$

The correction (8) diminishes the current and gives corrections to the conductivity and resistivity:

$$\sigma = \frac{1}{\nu}\left(1 - \frac{W}{\nu T^{3/2}}\right) \,, \quad \rho \approx \nu + \frac{W}{T^{3/2}} \,. \tag{9}$$

In the next section, we compute the elastic resistivity for small-angle scattering: $\rho_e = (d-1)W\sqrt{2\pi}/32T^{3/2}$. Comparison with that value shows that the Matthiessen's rule, $\rho = \nu + \rho_e$, is quite off in the limit of weak elastic scattering. For $d = 3$ the W-addition to the resistivity (9) is more than six times larger (and for $d = 2$ twelve times larger) than the rule predicts since there is a strong positive correlation between the two mechanisms of the momentum relaxation. Indeed, an angular scattering enhances the frictional relaxation of the x-momentum by bringing more particles from other directions.

The case $c = \sqrt{d-1}$ for $d = 2$ corresponds to γ independent of p:

$$f_2 = -\frac{\gamma p E}{\nu^2 T} f_0(p) \cos\theta = -f_1 \frac{\gamma}{\nu} \,, \tag{10}$$

which gives the conductivity correction $-\gamma/\nu^2$ and the resistivity as follows:

$$\rho = \nu + \gamma \,. \tag{11}$$

We see that only for the momentum-independent scattering rate in two dimensions is the Matthiessen's rule valid. In this case, (11) is valid for arbitrary relation between γ and ν. Indeed, the linear-response first angular harmonic is exactly equal to

$$f_1 = \frac{pE}{(\nu + \gamma)T} f_0 \cos\theta \,. \tag{12}$$

The terms we neglected are quadratic in E and contribute to the zeroth and second harmonics. They can be accounted for in the next orders. The terms cubic in E contribute to the third angular harmonic.

IV. LOW-FRICTION LIMIT

One may assume that in the limit $\nu \to 0$, $T \to 0$ one can neglect the thermostat-related terms and write:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial p^i} E_i f = \hat{I} f \ . \tag{13}$$

Yet this equation does not have a steady state for the simple reason that the second term pumps energy while the last one does not change it. Despite that, one can find the resistivity within the linear response theory, assuming that $f(\mathbf{p}) = f_0(p) + f_1(\mathbf{p}) + O(E^2)$, where $f_0(p)$ is isotropic and $f_1 \propto E$. Then for small-angle scattering we can write

$$\frac{\partial}{\partial p^i} E_i f_0 = \frac{W}{p^3} \Delta_\Omega f_1 \ . \tag{14}$$

$$f_1 = -E \cos\theta \frac{p^3 f_0'}{(d-1)W} \tag{15}$$

Taking $f_0 = \delta(p - p_0)$, we obtain $\bar{p} = \frac{2p_0^3 E}{(d-1)W}$, which gives the resistivity for a given energy $E_0 = p_0^2/2$:

$$\rho_e(E_0) = \frac{(d-1)W}{2(2E_0)^{3/2}} \ . \tag{16}$$

Resistivity at a fixed temperature is obtained by using $f_0 \propto \exp(-p^2/2T)$:

$$\rho_e(T) = \frac{\sqrt{2\pi}(d-1)W}{32T^{3/2}} \ . \tag{17}$$

Let us now account for small friction. The Matthiessen rule would predict just adding ν to resistivity: $\rho = \sqrt{2\pi}(d-1)W/32T^{3/2} + \nu$. Let us show that this is not the case. We look for the correction in the form $f = f_0(p) + f_1(\mathbf{p}, W) + f_2(\mathbf{p}, W, \nu)$, where $f_0 \propto e^{-p^2/2T}$ and $f_1 = E f_0 \cos\theta p^4/T(d-1)W$ due to (15). Substituting into (3) and assuming $f_1 \gg f_2$, we obtain (for $d = 3$):

$$f_2 = f_1 \frac{\nu(9pT - 2p^3)}{W} \ . \tag{18}$$

Non-surprisingly, friction decreases the number of fast particles and increases the number of slow ones. The resulting resistivity is as follows:

$$\rho = \frac{\sqrt{2\pi}W}{16T^{3/2}} + \nu \frac{630\pi}{512} \ . \tag{19}$$

For a general case of scattering by order-unity angles, we can simply put $\hat{I} f_1 = -\gamma f_1$, which gives $f_1 = f_0 E \cos\theta/\gamma$. The elastic resistivity is simply γ and $\rho = \nu + \gamma$.

We thank A. Finkelstein and B. Spivak for useful discussions. The work of A. K. was supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0011632 and by the Simons Investigator Award. The work of G.F. is supported by the Excellence Center at WIS, by the Simons grants 662962 and 617006, the NSF-BSF grant 2020765, and by the EU Horizon grants No 873028 and 823937.

---

[1] A. Kapustin, arXiv:2402.09672 [cond-mat.stat-mech]

[2] L. van Hove, Quantum-mechanical perturbations giving rise to a statistical transport equation, *Physica* **21**, 517-540 (1954).

[3] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Ch. 16 (Brooks/Cole 1976).

Giant nonlocality in nearly compensated 2D semimetals

S. Danz,[1] M. Titov,[2, 3] and B.N. Narozhny[1, 4]

[1]*Institut für Theorie der kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[2]*Radboud University Nijmegen, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*
[3]*ITMO University, 197101 St. Petersburg, Russia*
[4]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409 Moscow, Russia*

(Dated: February 12, 2020)

In compensated two-component systems in confined, two-dimensional geometries, nonlocal response may appear due to external magnetic field. Within a phenomenological two-fluid framework, we demonstrate the evolution of charge flow profiles and the emergence of a giant nonlocal pattern dominating charge transport in magnetic field. Applying our approach to the specific case of intrinsic graphene, we suggest a simple physical explanation for the experimental observation of giant nonlocality. Our results provide an intuitive way to predict the outcome of future experiments exploring the rich physics of many-body electron systems in confined geometries as well as to design possible applications.

The trend towards miniaturization of electronic devices requires a deeper understanding of the electron flow in confined geometries. In contrast to the electric current in household wiring, charge flow in small chips with multiple leads may exhibit complex spatial distribution patterns depending on the external bias, electrostatic environment, chip geometry, and magnetic field. Traditionally, such patterns were detected using nonlocal transport measurements [1–7], i.e. by measuring voltage drops between various leads other than the source and drain. Devised to study ballistic propagation of charge carriers in mesoscopic systems, these techniques were recently applied to investigate possible hydrodynamic behavior in ultra-pure conductors [8–12], where the unusual behavior of the nonlocal resistance is often associated with viscosity of the electronic system [13–17].

Nonlocal resistance measurements have also been used to study edge states accompanying the quantum Hall effect [18–23]. While the exact nature of the edge states has been a subject of an intense debate, the nonlocal resistance, R_{NL}, appears to be an intuitively clear consequence of the fact that the electric current flows along the sample edges and not through the bulk. Such a current would not be subject to exponential decay [24] exhibited by the bulk charge propagation leading to a much stronger nonlocal resistance.

In recent years the focus of the experimental work on electronic transport has been gradually shifting towards measurements at nearly room temperatures [6, 8–10, 21]. A particularly detailed analysis of the nonlocal resistance in a wide range of temperatures, carrier densities, and magnetic fields was performed on graphene samples [21]. Remarkably, the nonlocal resistance measured at charge neutrality remained strong well beyond the quantum Hall regime, with the peak value $R_{NL} \approx 1.5\,\mathrm{k}\Omega$ at $B = 12\,\mathrm{T}$ and $T = 300\,\mathrm{K}$, three times higher than that at $T = 10\,\mathrm{K}$.

In this Letter, we argue that the giant nonlocality observed in intrinsic graphene at high temperatures can be attributed to the presence of two types of charge carriers (electrons and holes): at the neutrality point, the two bands (the conductance and valence bands) touch creating a two-component electronic system. Physics of such systems is much richer than in their single-component counterparts. Observed phenomena that are directly related to the two-band structure of the neutrality point include giant magnetodrag in graphene [25, 26] and linear magnetoresistance [27, 28]. Both effects have been explained within a phenomenological framework [26, 27] allowing for a two-component (electron-hole) system coupled by the external magnetic field. We generalize this approach to investigate evolution of the spatial distribution of the electron current density in the experimentally relevant Hall bar geometry. In sufficiently strong magnetic fields, the current density forms a giant nonlocal pattern where the current is flowing not only in the bulk, but also along the boundaries leading to strong nonlocal resistance, see Fig. 1. Such patterns can be directly observed in laboratory experiments using the modern imaging techniques [29–31]. Tuning the model parameters to the specific values available for graphene, we arrive at a quantitative estimate of the nonlocal resistance [21].

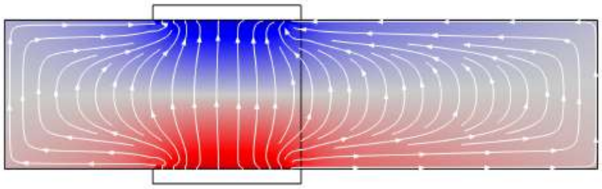


FIG. 1: Giant nonlocality in a compensated semimetal in magnetic field. The arrows indicate the current flow and the color map shows the electrochemical potential (see the main text and Figs. 2 and 3 for specific parameters).

To highlight the difference between the one- and two-component systems, we briefly recall the macroscopic description of electronic transport in the standard (former) case. Allowing for nonuniform charge density, the linear

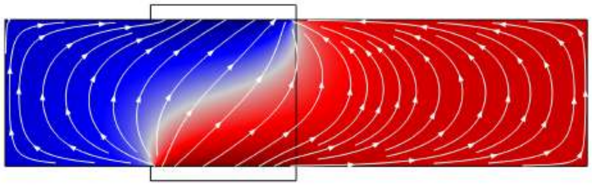


FIG. 2: Classical Hall effect in a one-component electronic system. The current density (shown by the arrows) and the electrochemical potential (shown by the color map) were obtained from Eqs. (1) for a sample of the width $W = 1\,\mu$m and length $L = 4\,\mu$m with the carrier density $n = 10^{12}\,\mathrm{cm}^{-2}$ at the temperature $T = 240\,$K and in magnetic field $B = 0.2\,$T.

relation between the electric current $\boldsymbol{J}$ and the external fields $\boldsymbol{E}$, $\boldsymbol{B}$ could be formulated as [17, 32, 33]

$$r_0\boldsymbol{J} = \boldsymbol{E} + r_H\boldsymbol{e_B}\times\boldsymbol{J} + \frac{1}{e\nu_0}\boldsymbol{\nabla}n, \tag{1a}$$

where $e > 0$ is the unit charge, ν_0 is the density of states (DoS), n is the carrier density, $\boldsymbol{e_B}$ is the unit vector in the direction of the magnetic field, and r_0 and r_H are the longitudinal and Hall resistivities. Within the Drude-like description, $r_H = \omega_c\tau r_0$ (ω_c is the cyclotron frequency and τ is the mean free path). The relation Eq. (1a) is applicable to a wide range of electronic systems from simple metals [34, 35] to doped graphene [11, 36]. The transport coefficients r_0 and r_H could be treated as phenomenological or could be derived from the underlying kinetic theory [11, 32, 37].

In addition to Eq.(1a), the electric current satisfies the continuity equation, which for stationary currents reads

$$\boldsymbol{\nabla}\cdot\boldsymbol{J} = 0. \tag{1b}$$

Charge density inhomogeneity induces electric field, so that Eq. (1a) should be combined with the corresponding electrostatic problem. Most recent experiments were performed in gated structures, where the relation between the electric field and charge density simplifies [27, 38]. In two-dimensional (2D) samples

$$\boldsymbol{E} = \boldsymbol{E}_0 - \frac{e}{C}\boldsymbol{\nabla}n, \tag{1c}$$

where $C = \epsilon/(4\pi d)$ is the gate-to-sample capacitance per unit area, d is the distance to the gate, ϵ is the dielectric constant, and $\boldsymbol{E}_0$ is the external field.

In a two-terminal (slab) geometry, solution of Eqs. (1) is a textbook problem. In the absence of magnetic field, the resulting electrochemical potential is governed by the relation of the mean free path to the system size, exhibiting either a flat (in short, ballistic samples) or linear (in long, diffusive samples) spatial profile. Most recently, these solutions were used as benchmarks in the imaging experiment [29] and the numerical solution of the hydrodynamic equations in doped graphene [17]. In external magnetic field, the system exhibits the classical Hall effect, which in short samples is accompanied by nontrivial current flow patterns [39].

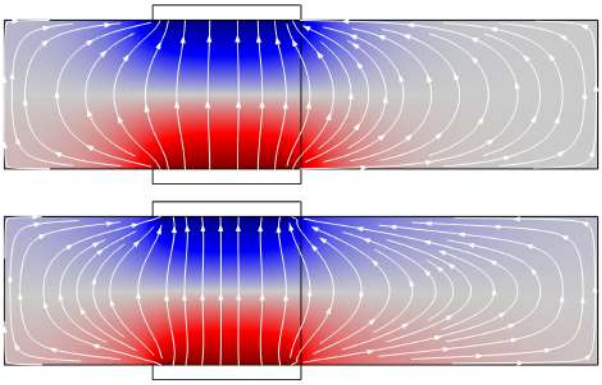


FIG. 3: Charge flow in compensated semimetals. Top: Ohmic flow in the absence of magnetic field. Bottom: emergent nonlocality in weak magnetic field $B = 0.2\,$T. The associated potential on the sample boundaries grows with the increasing field, see Fig. 1 for the pattern at $B = 2\,$T. Stronger fields expel the current from the bulk such that it flows along the boundary.

In a four-terminal Hall bar geometry, the electric current still fills the whole sample, but decays exponentially [24] away from the direct path between source and drain. The resulting flow pattern was calculated (in the context of doped graphene) in Refs. [14, 15, 17]. In magnetic field, the pattern gets skewed due to the classical Hall effect, but exhibits no qualitatively new features, see Fig. 2.

Let us now extend the transport equations (1) to a two-component system. Keeping in mind applications to graphene, we re-write Eq. (1a) for the quasiparticles in the conduction band ("electrons") in the form

$$-\boldsymbol{j}_e = eD\nu_e\boldsymbol{E} + \omega_c\tau\boldsymbol{j}_e\times\boldsymbol{e_B} + D\boldsymbol{\nabla}n_e, \tag{2a}$$

where $\boldsymbol{j}_e$ is the electron flow density (carrying the electric current $\boldsymbol{J}_e = -e\boldsymbol{j}_e$) and ν_e is DoS. The "holes" (i.e., the quasiparticles in the valence band) are described by

$$-\boldsymbol{j}_h = -eD\nu_h\boldsymbol{E} - \omega_c\tau\boldsymbol{j}_h\times\boldsymbol{e_B} + D\boldsymbol{\nabla}n_h. \tag{2b}$$

Here the electric current carried by the holes is $\boldsymbol{J}_h = e\boldsymbol{j}_h$ and DoS may differ from that of electrons, $\nu_h \neq \nu_e$. For simplicity, we assume that the the cyclotron frequency, mean free time, and diffusion constant for the two bands coincide (a generalization is straightforward, but doesn't lead to qualitatively new physics).

The total electric current in the two component system is given by $\boldsymbol{J} = -e\boldsymbol{j}$, where $\boldsymbol{j} = \boldsymbol{j}_e - \boldsymbol{j}_h$. Introducing also the total quasiparticle flow $\boldsymbol{j}_I = \boldsymbol{j}_e + \boldsymbol{j}_h$, we find (cf. Ref. [37])

$$\boldsymbol{j} + eD(\nu_e + \nu_h)\boldsymbol{E} + \omega_c\tau\boldsymbol{j}_I\times\boldsymbol{e_B} + D\boldsymbol{\nabla}n = 0, \tag{3a}$$

$$\boldsymbol{j}_I + eD(\nu_e - \nu_h)\boldsymbol{E} + \omega_c\tau\boldsymbol{j}\times\boldsymbol{e_B} + D\boldsymbol{\nabla}\rho = 0, \tag{3b}$$

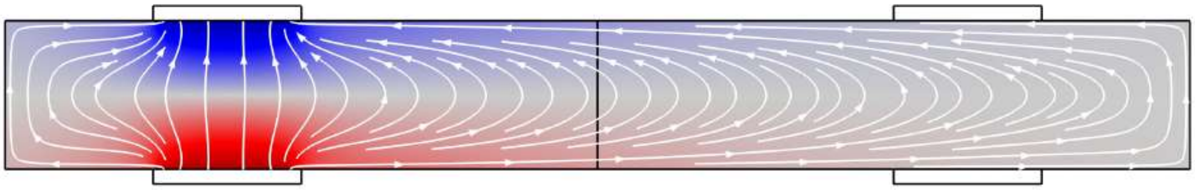


FIG. 4: Giant nonlocality in the Hall bar geometry. The sample has a width $W = 1\,\mu$m and length $8\,\mu$m, with the distance between contacts $L = 5\,\mu$m. The driving current is $I = 0.1\,\mu$A. The flow pattern was computed for $B = 0.8\,$T, cf. Fig. 1.

where $n = n_e - n_h$ is the carrier density per unit charge (the charge density being $-en$) and $\rho = n_e + n_h$ is the total quasiparticle density. The transport equations have to be supplemented by continuity equations reflecting the particle number conservation. The electric current satisfies Eq. (1b), but the total number of quasiparticles [40] can be affected by electron-hole recombination processes leading to a weak decay term in the continuity equation

$$\boldsymbol{\nabla}\cdot\boldsymbol{j}_I = -\delta\rho/\tau_R, \tag{3c}$$

where $\delta\rho$ is the deviation of the quasiparticle density from its equilibrium value and τ_R is the recombination time.

Under the assumption of electron-hole symmetry (e.g., at the charge neutrality point in graphene), $\nu_e = \nu_h$, we recover the phenomenological model of Ref. [27]. In the two-terminal geometry this model yields unsaturating linear magnetoresistance in classically strong fields [28].

Now we analyze the behavior of the phenomenological model (3) in the four-terminal Hall bar geometry. In the absence of the magnetic field, the system exhibits a typical Ohmic flow [14, 15, 17], see the top panel in Fig. 3. Applying the field we find a qualitative change in the flow pattern – the emergence of a boundary flow and the associated electrochemical potential at the sample edges. Increasing the field leads to the nonlocal pattern growing until it fills the whole sample, see Figs. 1 and 4. Stronger fields essentially expel the current from the bulk with the charge flow being concentrated along the sample boundaries, which leads to strong nonlocal resistance.

The nonlocal flow pattern emerging in magnetic field, see Figs. 1, 3 and 4, has to be contrasted with the vortices appearing in the viscous hydrodynamic flow (e.g., in doped graphene [14, 15, 17, 41]). In the latter case, vorticity appears due to the constrained geometry of the flow and the particular boundary conditions [15, 17, 42]: neglecting Ohmic effects, the solution of the hydrodynamic equations can be obtained by introducing the stream function, which obeys a biharmonic equation independent of viscosity, which however affects the distribution of the electrochemical potential. In contrast, within the model (3) the "Ohmic" scattering represents the only source of dissipation and hence cannot be omitted. One can still introduce the stream function, but now it is determined not only by the sample geometry, but also by the Ohmic scattering and magnetic field. As a result, the flow pattern does not exhibit vortices, unlike those suggested recently for the hydrodynamic flow in intrinsic graphene [41] (in the absence of magnetic field).

Nonlocal resistance in graphene subjected to external magnetic field was studied experimentally in Ref. [21]. At high enough temperatures where signatures of the quantum Hall effect are washed out, strong (or "giant") nonlocality was observed at the neutrality point. The effect vanishes in zero field as well as with doping away from neutrality. Both features are consistent with the model (3): in zero field the model exhibits usual Ohmic flow patterns, see Fig. 3, while at sufficiently high doping levels the effects of the second band are suppressed – the two equations (3a) and (3b) become identical showing the response typical of one-component systems, see Fig. 2.

Having discussed the qualitative features of the charge flow in two-component systems, we now turn to a quantitative calculation of nonlocal resistance in graphene. Although the model (3) is applicable to any semimetal, graphene is a by far better studied material with readily available experimental values for model parameters. Here we use the data measured in Refs. [8, 9, 21, 26, 43] and theoretical calculations of Refs. [11, 12, 26, 37, 41].

DoS of the quasiparticles in graphene has been evaluated in, e.g., Refs. [11, 12, 36, 37], and has the form

$$\nu_e + \nu_h = 2\mathcal{T}/(\pi v_g^2), \qquad \nu_e - \nu_h = 2\mu/(\pi v_g^2), \tag{4}$$

where μ is the chemical potential, v_g is the quasiparticle velocity in graphene, and $\mathcal{T} = 2T\ln[2\cosh(\mu/2T)]$. The generalized cyclotron frequency is $\omega_c = eBv_g^2/(c\mathcal{T})$ and the diffusion coefficient has the usual form $D = v_g^2\tau/2$. At charge neutrality, $\mu = 0$ and $\mathcal{T} = 2T\ln 2$, while in the degenerate regime $\mathcal{T}(\mu \gg T) = \mu$. The latter confirms that all coefficients in Eqs. (3a) and (3b) become identical with doping. Similarly, the continuity equations (1b) and (3c) should coincide in the degenerate regime. In graphene this happens by means of the fast decay of the recombination rate [26]. Close to neutrality we assume

$$\tau_R^{-1} = g^2 T/\cosh(\mu/T), \tag{5}$$

where g is determined by the corresponding matrix element. The above expression [26] reflects the exponential

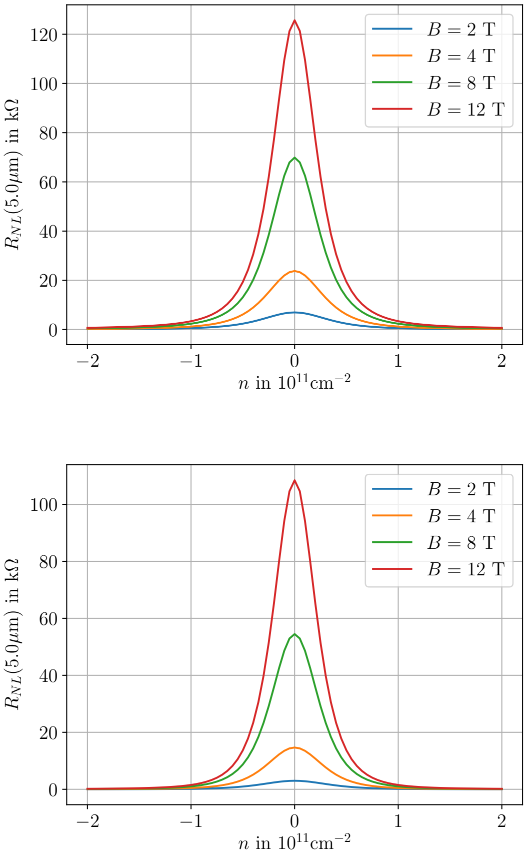


FIG. 5: Nonlocal resistance measured in the Hall bar geometry, see Fig. 4, as a function of carrier density. Top: Coulomb scatterers; bottom: short-ranged impurities. The impurity model parameters are chosen to represent the mobility at $n = 10^{11}\,\mathrm{cm}^{-2}$ reported in Ref. [21]. The range of magnetic fields and carrier densities as well as the distance to the gate ($d = 50\,\mathrm{nm}$) is taken from Ref. [21], see Fig.2.

decay of the two-band physics away from charge neutrality, which is responsible for the fast decay of R_{NL} as a function of carrier density [21], see Fig. 5. Finally, the mean-free time, τ, in graphene is a non-trivial function of temperature and carrier density [11, 12, 36, 43, 44], which strongly depends on the model of the impurity potential [45–50]. However, these dependencies are typically not exponential and hence do not affect the exponential decay of the nonlocal resistance.

In Fig. 5 we demonstrate the decay of R_{NL} for two impurity models – the Coulomb scatterers and short-ranged impurities – showing nearly identical behavior. Such robustness of the model (3) with respect of the functional dependence of the mean free time justifies the inaccuracy of our description of electronic transport in graphene, where close to charge neutrality the resistivity is strongly affected by electron-electron interaction. The data shown in Fig. 5 were obtained by solving Eqs. (3) in the Hall bar geometry of Fig. 4 using the estimate [41] for the recombination length scale, $\ell_R = v_g\tau_R \approx 10\,\mu$m (a previous calculation of Ref. [26] put it at a smaller value $1.2\,\mu$m), which leads to similar results for the nonlocal resistance, but with a smaller peak value at charge neutrality.

The results for R_{NL} shown in Fig. 5 are extremely similar to those reported in Ref. [21] with the exception of the values at neutrality, which are grossly exaggerated. There are several reasons for this behavior. Firstly, by ignoring the effects of electron-electron interaction, we strongly underestimate the usual resistivity of intrinsic graphene. Secondly, we ignore viscous effects. Furthermore, DoS in real graphene never really vanishes “at neutrality” due to electrostatic potential fluctuations [51]. As a result, the minimal carrier concentration is often as high as $10^{10}\mathrm{cm}^{-2}$, essentially cutting off the lower density range around the peak in Fig. 5. Finally, Eq. (5) is a rather crude estimate that needs to be improved.

To conclude, we have argued that the observed giant nonlocality in neutral graphene in non-quantizing magnetic fields at relatively high temperatures observed in Ref. [21] is a direct consequence of the two-band nature of the quasiparticle spectrum in graphene. As such, this effect is not specific to graphene and should be observable in any compensated two-component system. Our theory does not involve spin-related phenomena including the effect of Zeeman splitting invoked in Ref. [21]. The latter should be independent of the field direction, however, the effect was not observed in the nearly parallel field studied in Ref. [51]. Assuming the g-factor to be equal to 2, we estimate the Zeeman splitting as $E_z \approx 0.35\,\mathrm{meV} \approx 4\,\mathrm{K}$ at $B = 10\,\mathrm{T}$. The corresponding residual quasiparticle density (at $T = 0$) is given by $\rho_Q = E_z^2/(4\pi v_g^2) \approx 2.2 \times 10^6\,\mathrm{cm}^{-2}$. As a result, we expect the effects of Zeeman splitting to be observable at temperatures and carrier densities much lower than those typical to nonlocal measurements discussed here.

With material-specific parameters, our phenomenological model is capable of a quantitative description of the effect. For graphene, a more precise calculation involving solution of the full system of hydrodynamic equations near charge neutrality is required to reach perfect agreement with the data, however the present approach shows that the effect is more general and does not require additional assumptions of electronic hydrodynamics.

The authors are grateful to I.V. Gornyi, A.D. Mirlin, J. Schmalian, J.A. Sulpizio, M. Schütt, A. Shnirman, and Y. Tserkovnyak for fruitful discussions. This work was supported by the German Research Foundation DFG within FLAG-ERA Joint Transnational Call (Project GRANSPORT), by the European Commission under the EU Horizon 2020 MSCA-RISE-2019 program (Project 873028 HYDROTRONICS), and by the Russian Science Foundation Project No. 17-12-0 (MT). BNN acknowledges the support by the MEPhI Academic Excellence Project, Contract No. 02.a03.21.0005.

---

[1] W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, Phys. Rev. Lett. **58**, 2347 (1987).

[2] H. van Houten, C. W. J. Beenakker, J. G. Williamson, M. E. I. Broekaart, P. H. M. van Loosdrecht, B. J. van Wees, J. E. Mooij, C. T. Foxon, and J. J. Harris, Phys. Rev. B **39**, 8556 (1989).

[3] A. K. Geim, P. C. Main, P. H. Beton, P. Streda, L. Eaves, C. D. W. Wilkinson, and S. P. Beaumont, Phys. Rev. Lett. **67**, 3014 (1991).

[4] K. L. Shepard, M. L. Roukes, and B. P. Van der Gaag, Phys. Rev. Lett. **68**, 2660 (1992).

[5] Y. Hirayama, A. D. Wieck, T. Bever, K. von Klitzing, and K. Ploog, Phys. Rev. B **46**, 4035 (1992).

[6] G. Mihajlović, J. E. Pearson, M. A. Garcia, S. D. Bader, and A. Hoffmann, Phys. Rev. Lett. **103**, 166601 (2009).

[7] R. V. Gorbachev, J. C. W. Song, G. L. Yu, A. V. Kretinin, F. Withers, Y. Cao, A. Mishchenko, I. V. Grigorieva, K. S. Novoselov, L. S. Levitov, et al., Science **346**, 448 (2014).

[8] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, et al., Science **351**, 1055 (2016).

[9] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Nat. Commun. **9**, 4533 (2018).

[10] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. K. Kumar, A. Principi, I. Torre, M. B. Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, et al., Science **364**, 162 (2019).

[11] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, and J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[12] A. Lucas and K. C. Fong, J. Phys: Condens. Matter **30**, 053001 (2018).

[13] I. Torre, A. Tomadin, A. K. Geim, and M. Polini, Phys. Rev. B **92**, 165433 (2015).

[14] L. Levitov and G. Falkovich, Nat. Phys. **12**, 672 (2016).

[15] G. Falkovich and L. Levitov, Phys. Rev. Lett. **119**, 066601 (2017).

[16] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).

[17] S. Danz and B. N. Narozhny (2019), arXiv:1910.14473.

[18] P. L. McEuen, A. Szafer, C. A. Richter, B. W. Alphenaar, J. K. Jain, A. D. Stone, R. G. Wheeler, and R. N. Sacks, Phys. Rev. Lett. **64**, 2062 (1990).

[19] J. K. Wang and V. J. Goldman, Phys. Rev. B **45**, 13479 (1992).

[20] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[21] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, A. S. Mayorov, M. I. Katsnelson, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levitov, et al., Science **332**, 328 (2011).

[22] X.-P. Zhang, C. Huang, and M. A. Cazalilla, 2D Materials **4**, 024007 (2017).

[23] K. Komatsu, Y. Morita, E. Watanabe, D. Tsuya, K. Watanabe, T. Taniguchi, and S. Moriyama, Science Advances **4**, eaaq0194 (2018).

[24] L. J. van der Pauw, Philips Tech. Rev. **20**, 223 (1958).

[25] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, et al., Nat. Phys. **8**, 896 (2012).

[26] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, et al., Phys. Rev. Lett. **111**, 166601 (2013).

[27] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[28] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, Y. B. Vasilyev, P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, M. Titov, B. N. Narozhny, et al., Phys. Rev. B **93**, 195430 (2016).

[29] L. Ella, A. Rozen, J. Birkbeck, M. Ben-Shalom, D. Perello, J. Zultak, T. Taniguchi, K. Watanabe, A. K. Geim, S. Ilani, et al., Nat. Nanotechnol. **14**, 480 (2019).

[30] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, et al. (2019), arXiv:1905.10791.

[31] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, et al., Nature **576**, 75 (2019).

[32] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon Press, London, 1981).

[33] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).

[34] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1965).

[35] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[36] M. I. Katsnelson, *Graphene* (Cambridge University Press, 2012).

[37] B. N. Narozhny, Annals of Physics **411**, 167979 (2019).

[38] I. L. Aleiner and B. I. Shklovskii, Phys. Rev. B **49**, 13721 (1994).

[39] A. Shik, J. Phys. Condens. Matter **5**, 8963 (1993).

[40] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[41] H.-Y. Xie and A. Levchenko, Phys. Rev. B **99**, 045434 (2019).

[42] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).

[43] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[44] B. N. Narozhny, M. Titov, I. V. Gornyi, and P. M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012).

[45] T. Ando, J. Phys. Soc. Jpn. **75**, 074716 (2006).

[46] N. Shon and T. Ando, J. Phys. Soc. Jpn. **67**, 2421 (1998).

[47] K. Nomura and A. H. MacDonald, Phys. Rev. Lett. **96**, 256602 (2006).

[48] V. V. Cheianov and V. I. Fal'ko, Phys. Rev. Lett. **97**, 226801 (2006).

[49] I. L. Aleiner and K. B. Efetov, Phys. Rev. Lett. **97**, 236801 (2006).

[50] P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. B **74**, 235443 (2006).

[51] F. Chiappini, S. Wiedmann, M. Titov, A. K. Geim, R. V. Gorbachev, E. Khestanova, A. Mishchenko, K. S. Novoselov, J. C. Maan, and U. Zeitler, Phys. Rev. B **94**, 085302 (2016).

Transport properties of strongly coupled electron-phonon liquids

Alex Levchenko[a], Jörg Schmalian[b,c]

[a] *Department of Physics, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA*
[b] *Institut für Theorie der Kondensierten Materie, Karlsruher Institut für Technologie, D-76131 Karlsruhe, Germany*
[c] *Institut für Quantenmaterialien und Technologien, Karlsruher Institut für Technologie, D-76021 Karlsruhe, Germany*

Abstract

In this work we consider the hydrodynamic behavior of a coupled electron-phonon fluid, focusing on electronic transport under the conditions of strong phonon drag. This regime occurs when the rate of phonon equilibration due to e.g. umklapp scattering is much slower than the rate of normal electron-phonon collisions. Then phonons and electrons form a coupled out-of-equilibrium state where the total quasi-momentum of the electron-phonon fluid is conserved. A joint flow-velocity emerges as a collective hydrodynamic variable. We derive the equation of motion for this fluid from the underlying microscopic kinetic theory and elucidate its effective viscosity and thermal conductivity. In particular, we derive decay times of arbitrary harmonics of the distribution function and reveal its corresponding super-diffusive relaxation on the Fermi surface. We further consider several applications of this theory to magneto-transport properties in the Hall-bar and Corbino-disk geometries, relevant to experiments. In our analysis we allow for general boundary conditions that cover the crossover from no-slip to no-stress flows. Our approach also covers a crossover from the Stokes to the Ohmic regime under the conditions of the Gurzhi effect. In addition, we consider the frequency dependence of the surface impedance and non-equilibrium noise. For the latter, we notice that in the diffusive regime, a Fokker-Planck approximation, applied to the electron-phonon collision integral in the Eliashberg form, reduces it to a differential operator with Burgers type nonlinearity. As a result, the non-equilibrium distribution function has a shock-wave structure in the energy domain. The consequence of this behavior for the Fano factor of the noise is investigated. In conclusion we discuss connections and limitations of our results in the context of recent electron-phonon drag measurements in Dirac and Weyl semimetals, and layout directions for further extensions and developments.

Keywords: Electron-phonon scattering, drag viscosity, superdiffusion, magnetotransport, noise

---

Contents

1. Introduction and motivation

Hydrodynamic effects of electronic transport in quantum materials are of significant current interest in condensed matter physics; see reviews [1, 2] and references therein. Various transport measurements in electrostatically defined wires in the two-dimensional electron gas in Ga(Al)As hetero-structures [3–7], monolayer and bilayer graphene [8–15], quasi-two-dimensional delafossite metals $PdCoO_2$ and $PtCoO_2$ [16–18], Dirac semimetal $PtSn_4$ [20, 21], type-II Weyl semimetal tungsten phosphide WP_2 [22, 23], and antimony Sb [24], provided substantial evidence for viscosity-dominated electronic response. Recently, direct imaging techniques, employing scanning gate microscopy [25, 26], a nanotube single-electron transistor [27], and quantum spin magnetometry realized with nitrogen vacancy centers in diamond [28, 29], revealed signatures of the Poiseuille profile of electron flow in narrow graphene channels and mesoscopic Ga(Al)As. While experimental findings of electronic hydrodynamics in solid-state systems are mounting, the conceptual questions remain in particular in regard to different microscopic scattering mechanisms that govern transition to the hydrodynamic regime in different materials, and how they manifest in transport coefficients.

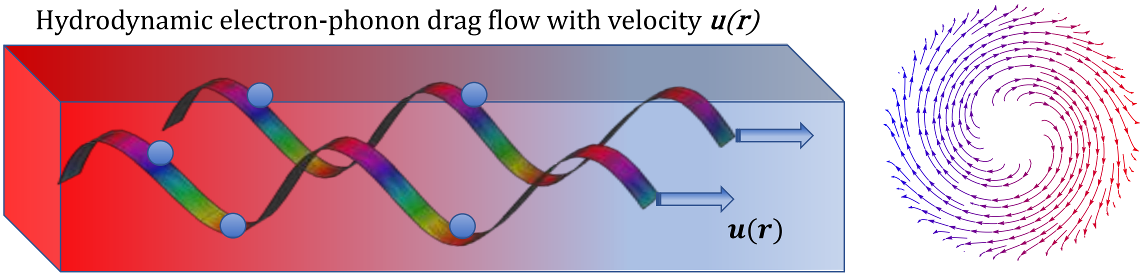


Figure 1: [Left panel]: When the total momentum of an interacting electron-phonon system decays slowly, a coupled out-of-equilibrium state – the electron-phonon fluid – emerges. As a result, a joint flow-velocity $\boldsymbol{u}(\boldsymbol{r})$ emerges as a collective hydrodynamic variable. Such a state displays viscous flow, super-diffusion in phase space and gives rise to shock-wave phenomena in the energy domain. In the figure we indicate the flow of the coupled electron wave and the moving ions that indicate the spread of acoustic waves. [Right panel]: Swirling magneto-flow profile of $\boldsymbol{u}(\boldsymbol{r})$ for a Corbino disc in an applied magnetic field, discussed in Section 3.3.

Typically the description of electronic conduction processes in solids requires a kinetic theory that is based on the formalism of the Boltzmann equation [30–32]. In this framework, microscopic scattering processes of momentum and energy relaxation are captured by collision terms between electrons and impurities, phonons or other relevant excitations. The electrical and thermal conductivities are then related to these microscopic length and time scales for momentum and energy relaxation. In contrast, a hydrodynamic description relies on the existence of locally conserved quantities. In this regime momentum and energy conserving electron-electron (ee) collisions are frequent and occur on shortest length and time scales. In this picture, the resistance, for example, can be related to the electronic viscosity and the thermal conductivity [33–35]. More generally, for conductors in which the underlying electron liquid lacks Galilean invariance the resistivity is determined by the entire thermoelectric matrix of the intrinsic kinetic coefficients [36–40].

Since electronic scattering lengths are strongly temperature dependent and highly sensitive to the type of scattering, one often argues that the hydrodynamic regime sets in at intermediate temperatures. Indeed, at lowest temperatures when the electron-electron and electron-phonon scattering mean free paths diverge, the electronic momentum is relaxed by scattering with impurities and boundary inhomogeneities. At higher temperatures, when phonon excitation branches are activated, electron-phonon scattering is the main mechanism that relaxes both momentum and energy of the electronic system. In between these two limits, and provided samples of sufficient purity, there is a range of temperatures where the electron fluid attains local equilibrium on the length scale of electron-electron collisions, which is short compared to the scales at which the conservation laws break down. Then the dynamics of the electron fluid can be treated hydrodynamically. This is certainly the scenario that occurs in graphene [1, 2, 41, 42] and perhaps very high mobility semiconductor hetero-structures of moderately-strongly correlated electrons at low densities [6, 43].

The above assertion that electron-phonon scattering is destructive for establishing electronic hydrodynamic regime by relaxing electronic momentum relies on a crucial assumption that the phonons are in thermal equilibrium. This transport situation has been considered in multiple works and much is known about momentum and energy relaxation rates from the solution of the Boltzmann equation, see for example Refs. [44–47]. Recently detailed *ab initio* calculations provided firm results for the relevant electron-phonon scattering processes in semimetals accounting for complexities of their Fermi surfaces and microscopic details of electron-phonon coupling [48]. However, it was pointed out by Peierls [49] early on that in fact the non-equilibrium nature of the current-carrying electronic distribution should lead, through electron-phonon scattering, to a phonon distribution that is also out of equilibrium. As a consequence, the total quasi-momentum of a combined electron-phonon systems would be conserved in the absence of umklapp processes. The electrons and phonons would then drift along together, maintaining their nonzero crystal momentum and also a nonzero current, see Figure 1 for the schematic illustration. The drift velocity can be treated as an emergent hydrodynamic soft mode whose relaxation occurs at the longer time scale of umklapp scattering due to phonon nonlinearities or assisted by scattering with electrons.

The transport theory of this phonon drag effect was developed by Gurevich in the context of thermoelectric phenomena [50]. Later Gurzhi [51], Nielsen and Shklovskii [52], and Gurevich and Shklovskii [53], and Gurevich and Laikhtman [54] put forward hydrodynamic description of phonons in dielectrics and coupled electron-phonon liquids in metals and semiconductors (see also a detailed review [55]). Of particular relevance to our work, Steinberg [56] and Gurzhi and Kopeliovich [57] considered the problem of electric conductivity of pure metals with an account of phonon drag. The electron viscosity was determined in Ref. [56], while Ref. [57] analyzed the case of a metal with open Fermi surface consisting of large electron (or hole) groups interconnected arbitrarily by a narrow necks. In this situation the dominant cause of the low-temperature resistance is due to umklapp events occurring in collisions between electrons and phonons which remains effective down to lowest temperatures. Because of the kinematic constraints of momentum and energy conservations in scattering, the change of electron momentum in each act of collision is small and scattering occurs preferentially at small angles. As a result, electrons effectively diffuse in momentum space. This enables one to reduce the full kinetic equation to a form of a Fokker-Planck type and account for umklapp processes by imposing periodic boundary conditions on the non-equilibrium distribution function.

In this work, we in large parts develop alternative derivations of the classic works [56, 57], which allow us to make extensions or draw additional conclusions for the behavior of electron-phonon fluids. For example, we consider a complimentary scenario of a Peierls mechanism of umklapp scattering mediated by phonon-phonon collisions. The rate of these processes is exponential in temperature whereas the rate of normal electron-phonon collisions is a power-law. The interplay between the two leads to a pronounced peak in the temperature dependent thermopower that can be observed as one lowers the temperature. This feature is considered as one of the hallmarks of strong electron-phonon interactions as recently seen in semimetals [20, 21]. Even though we face similar technical aspects of the problem as was already considered in Refs. [56, 57], we perform a somewhat different route to analyze the problem. We do not expand the collision integrals in the limit of small momenta but rather choose to work directly with the fully coupled collision integrals. While it will not be possible to solve these equations analytically, it is nevertheless possible to extract the main qualitative predictions from them in particular with regard to the temperature dependence of the drag viscosity and thermal conductivity of the coupled electron-phonon fluid. Furthermore, we believe that our approach may be more suitable if one wants to quantitatively describe realistic materials with a more complex shape of the Fermi surface. The rich physics that emerges if one includes such effects and anisotropies of the underlying crystal was recently elucidated in Refs. [58, 59].

Candidate materials for electron-phonon fluid behavior are clearly the delafossite metals $PdCoO_2$ and $PtCoO_2$ [16–18]. The temperature dependence of the bulk resistivity observed in Ref. [19] is fully consistent with phonon-drag behavior, i.e. inelastic scattering at low T has an exponential temperature dependence, rather the Bloch-Grüneisen behavior $\rho \propto T^5$ that occurs without phonon drag. In addition, hydrodynamic flow always requires strong momentum-conserving collisions. If collisions in the delafossite metals would be due to electron-electron scattering, their large Fermi surface would immediately give rise to equally strong umklapp processes. Hence, it seems that these systems have weak electron-electron scattering but are governed by electron-phonon scattering with phonon drag. In addition, evidence for electron-phonon fluid behavior was reported for the semimetal $PtSn_4$ [21], another material that displays very low resistivity at low temperatures and shows a pronounced phonon drag peak in the low temperature thermopower [20].

The remainder of the paper is organized as follows. In Sec. (2) we formulate the generic kinetic theory of coupled integro-differential equations for non-equilibrium distributions of electrons and phonons. We linearize these equations and study parity properties of the collision kernel. We also estimate rates of electron-phonon and phonon-electron collisions. Even though they originate from the same matrix elements, the respective mean free paths are parametrically different due to distinct phase space restrictions for fermions and bosons. As a methodological exercise, we illustrate how Bloch's law for the electron-phonon resistivity follows from the solution of the integral Boltzmann equation when phonons are taken at equilibrium, and discuss how this solution is violated when complete dragging of phonons is imposed. Analyzing the conservation laws of the problem, we demonstrate how a joint drift velocity of the coupled electron-phonon system emerges as hydrodynamic variable, even though both constituents of the fluid have vastly different quasiparticle velocities. We finally consider a partially equilibrated case of phonon drag with rare momentum relaxing collisions and derive the hydrodynamic equation of motion for the flow of the coupled electron-phonon liquid. This analysis reveals the intrinsic viscosity and thermal conductivity in the drag regime. In Sec. (3) we apply this hydrodynamic description to several practical examples of viscous resistive effects and the Gurzhi effect in particular [60]. We consider flows in different geometries of a Hall bar, a quantum wire, a Corbino disk, and allow for boundary conditions with arbitrary slip length that enables us to cover the crossover from no-slip to no-stress regimes. We also consider effects of a magnetic field, and the Hall viscosity, in particular for the viscous magnetoresistance and study finite-frequency responses in the context of the skin effect. Lastly we briefly touch upon the non-equilibrium thermometry of electron-phonon collisions via shot noise in the diffusive regime. We summarize our findings in Sec. (4) and discuss open questions and directions for future research. Various technical calculations are delegated to several supplementary appendices that expand on properties and methods of analysis of the electron-phonon collision operator.

2. From kinetic to hydrodynamic theory

2.1. Electron-phonon interaction

In many practical situations and for a broad range of temperatures, the electron-phonon interaction is dominated by processes with single-phonon emissions or absorptions [30]. Two-phonon processes could become important when one-phonon processes are forbidden or suppressed by the conservation laws or by symmetry restrictions for the transition matrix elements [32]. We restrict our attention to single-phonon processes exclusively. Furthermore, we will treat only the situation of scattering by long-wavelength acoustic phonons with a single electronic band.

For spatially inhomogeneous and time-dependent conditions the coupled kinetic equations for non-equilibrium electron $n(\boldsymbol{p},\boldsymbol{r},t)$ and phonon $N(\boldsymbol{q},\boldsymbol{r},t)$ distribution functions read (hereafter $\hbar = k_B = 1$):

$$\frac{\partial n}{\partial t} + \boldsymbol{v}\frac{\partial n}{\partial \boldsymbol{r}} + e\boldsymbol{v}\boldsymbol{E}\frac{\partial f}{\partial \varepsilon} = \mathrm{St}_{\mathrm{ep}}\{n,N\} + \mathrm{St}_{\mathrm{ei}}\{n\}, \tag{2.1}$$

$$\frac{\partial N}{\partial t} + \boldsymbol{s}\frac{\partial N}{\partial \boldsymbol{r}} = \mathrm{St}_{\mathrm{pe}}\{n,N\} + \mathrm{St}^{N}_{\mathrm{pp}}\{N\} + \mathrm{St}^{U}_{\mathrm{pp}}\{N\}. \tag{2.2}$$

Here $\boldsymbol{v} = \partial_{\boldsymbol{p}}\varepsilon$ and $\boldsymbol{s} = \partial_{\boldsymbol{q}}\omega$ are electron and phonon group velocities, and $\boldsymbol{E}$ is an external electric field. A finite magnetic field or temperature gradient will be added later in the text when we consider applications where this becomes necessary. In the steady-state regime the explicit time derivatives on the left hand sides vanish. In equilibrium, the fermionic and bosonic distributions are the usual Fermi-Dirac and Bose-Einstein functions

$$f_\varepsilon = [\exp[(\varepsilon_{\boldsymbol{p}} - \varepsilon_F)/T] + 1]^{-1}, \quad b_\omega = [\exp(\omega_{\boldsymbol{q}}/T) - 1]^{-1}. \tag{2.3}$$

The primary focus of our attention will be the electron-phonon ($\mathrm{St}_{\mathrm{ep}}$) and phonon-electron ($\mathrm{St}_{\mathrm{pe}}$) collision integrals. The other terms such as electron-impurity $\mathrm{St}_{\mathrm{ei}}$, and phonon-phonon collisions, both normal type via phonon nonlinearities $\mathrm{St}^{N}_{\mathrm{pp}}$ and umklapp type $\mathrm{St}^{U}_{\mathrm{pp}}$, are kept for generality but their explicit forms will not be needed.

The electron-phonon collision integral consists of two contributions corresponding to emission and absorption of a phonon:

$$\begin{aligned}\mathrm{St}_{\mathrm{ep}}\{n,N\} = &\int_{\boldsymbol{q}} W(\boldsymbol{p}|\boldsymbol{p}'\boldsymbol{q})\delta(\varepsilon_{\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}'} - \omega_{\boldsymbol{q}})[n_{\boldsymbol{p}'}(1-n_{\boldsymbol{p}})N_{\boldsymbol{q}} - n_{\boldsymbol{p}}(1-n_{\boldsymbol{p}'})(1+N_{\boldsymbol{q}})] \\ &+ \int_{\boldsymbol{q}} W(\boldsymbol{p}\boldsymbol{q}|\boldsymbol{p}')\delta(\varepsilon_{\boldsymbol{p}} + \omega_{\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}'})[n_{\boldsymbol{p}'}(1-n_{\boldsymbol{p}})(1+N_{\boldsymbol{q}}) - n_{\boldsymbol{p}}(1-n_{\boldsymbol{p}'})N_{\boldsymbol{q}}].\end{aligned} \tag{2.4}$$

These two terms take care of the out-scattering and reverse in-scattering processes. In equilibrium the difference between these processes is nullified as dictated by the detailed balance condition. The momentum conservation in the first term implies $\boldsymbol{p} = \boldsymbol{p}' + \boldsymbol{q} + \mathbf{g}$, while in the second $\boldsymbol{p} + \boldsymbol{q} = \boldsymbol{p}' + \mathbf{g}$ where $\mathbf{g}$ is reciprocal lattice vector. The phonon-electron collision integral counts the overall difference between the number of phonons emitted by electrons with momenta $\boldsymbol{p}$, as allowed by the conservation laws, and number of phonons absorbed by electron with momenta $\boldsymbol{p}'$:

$$\mathrm{St}_{\mathrm{pe}}\{n,N\} = 2\int_{\boldsymbol{p}} W(\boldsymbol{p}|\boldsymbol{p}'\boldsymbol{q})\delta(\varepsilon_{\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}'} - \omega_{\boldsymbol{q}})\left[n_{\boldsymbol{p}}(1-n_{\boldsymbol{p}'})(1+N_{\boldsymbol{q}}) - n_{\boldsymbol{p}'}(1-n_{\boldsymbol{p}})N_{\boldsymbol{q}}\right]. \tag{2.5}$$

A factor of two accounts for the electron spin in these processes, and momentum conservation is implicit and fixes the momentum $\boldsymbol{p}'$. At the level of the leading Born approximation, the probabilities of scattering for direct and reverse processes are equal to each other $W(\boldsymbol{p}|\boldsymbol{p}'\boldsymbol{q}) = W(\boldsymbol{p}\boldsymbol{q}|\boldsymbol{p}')$. Furthermore, for the deformation potential interaction and in the long-wavelength limit, the transition probability is linearly proportional to phonon momentum $W \propto |\boldsymbol{q}|$. In what follows, we will concentrate on low-temperature processes below the scale of Debye energy, namely $T < \omega_D$.

2.2. Linearized collision kernels and scattering rates

In general, it is not possible to solve the coupled nonlinear Boltzmann equations (2.1) and (2.2). An analytical analysis is often restricted to the linear-response regime and uses solely the linearized form of the collision terms. For this purpose we assume that the distribution functions are close to their equilibrium expressions with small corrections $n = f + \delta n$ and $N = b + \delta N$. To determine the collision terms in Eqs. (2.4) and (2.5) up to linear order in non-equilibrium corrections, it is customary to parametrize them as follows

$$\delta n = f(1-f)\psi = -T\frac{\partial f}{\partial \varepsilon}\psi, \quad \delta N = b(1+b)\phi = -T\frac{\partial b}{\partial \omega}\phi. \tag{2.6}$$

This form of δn and δN makes it convenient to employ the detailed balance conditions under the integral. In addition, the expression for the entropy production in the system becomes a symmetric quadratic form in terms of ψ and ϕ, which is very useful for the variational formulation of the Boltzmann equation.

We begin with $\mathrm{St}_{\mathrm{pe}}$ in Eq. (2.5) as it is simpler in structure, but the same sequence of steps will apply to the remaining collision terms. We follow the presentation given in Ref. [30] including the notation. In the brackets of Eq. (2.5) that account for statistical occupations we take out the product $(1-n_{\boldsymbol{p}})(1-n_{\boldsymbol{p}'})(1+N_{\boldsymbol{q}})$ and then perform a variation of this expression with respect to the equilibrium state, which gives

$$\delta\,\mathrm{St}_{\mathrm{pe}}\{n,N\}=2\int_{\boldsymbol{p}}W(\boldsymbol{p}|\boldsymbol{p}'\boldsymbol{q})\delta(\varepsilon_{\boldsymbol{q}}-\varepsilon_{\boldsymbol{p}'}-\omega_{\boldsymbol{q}})(1-f_{\varepsilon_{\boldsymbol{p}}})(1-f_{\varepsilon_{\boldsymbol{p}'}})(1+b_{\omega_{\boldsymbol{q}}})\delta\left[\frac{n_{\boldsymbol{p}}}{1-n_{\boldsymbol{p}}}-\frac{n_{\boldsymbol{p}'}}{1-n_{\boldsymbol{p}'}}\frac{N_{\boldsymbol{q}}}{1+N_{\boldsymbol{q}}}\right]. \tag{2.7}$$

Next we observe that

$$\delta\left(\frac{n}{1-n}\right)=\frac{\delta n}{(1-f)^2}=\frac{f}{1-f}\psi,\quad \delta\left(\frac{N}{1+N}\right)=\frac{\delta N}{(1+b)^2}=\frac{b}{1+b}\phi, \tag{2.8}$$

and use well-known properties between equilibrium Fermi and Bose functions (also making use of the energy-conserving delta function):

$$f_{\varepsilon_{\boldsymbol{p}}}(1-f_{\varepsilon_{\boldsymbol{p}}-\omega_{\boldsymbol{q}}})=[f_{\varepsilon_{\boldsymbol{p}}-\omega_{\boldsymbol{q}}}-f_{\varepsilon_{\boldsymbol{p}}}]b_{\omega_{\boldsymbol{q}}},\quad f_{\varepsilon_{\boldsymbol{p}}-\omega_{\boldsymbol{q}}}(1-f_{\varepsilon_{\boldsymbol{p}}})=[f_{\varepsilon_{\boldsymbol{p}}-\omega_{\boldsymbol{q}}}-f_{\varepsilon_{\boldsymbol{p}}}](1+b_{\omega_{\boldsymbol{q}}}). \tag{2.9}$$

As a result we find

$$\delta\,\mathrm{St}_{\mathrm{pe}}\{\psi,\phi\}=2\int_{\boldsymbol{p}}K_{-}(\boldsymbol{p},\boldsymbol{q})\left[\psi_{\boldsymbol{p}}-\psi_{\boldsymbol{p}'}-\phi_{\boldsymbol{q}}\right] \tag{2.10}$$

with the kernel

$$K_{\mp}(\boldsymbol{p},\boldsymbol{q})=W(\boldsymbol{p}|\boldsymbol{p}'\boldsymbol{q})b_{\omega_{\boldsymbol{q}}}(1+b_{\omega_{\boldsymbol{q}}})[f_{\varepsilon_{\boldsymbol{p}}\mp\omega_{\boldsymbol{q}}}-f_{\varepsilon_{\boldsymbol{p}}}]\delta(\varepsilon_{\boldsymbol{q}}-\varepsilon_{\boldsymbol{p}'}\mp\omega_{\boldsymbol{q}}). \tag{2.11}$$

In complete analogy we find for the linearized version of Eq. (2.4) the following expression

$$\delta\,\mathrm{St}_{\mathrm{ep}}\{\psi,\phi\}=\int_{\boldsymbol{q}}K_{-}(\boldsymbol{p},\boldsymbol{q})\left[\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}}+\phi_{\boldsymbol{q}}\right]-\int_{\boldsymbol{q}}K_{+}(\boldsymbol{p},\boldsymbol{q})\left[\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}}-\phi_{\boldsymbol{q}}\right]. \tag{2.12}$$

The important property of these collision kernels is that they preserve the parity $\boldsymbol{q}\to-\boldsymbol{q}$ or $\boldsymbol{p}\to-\boldsymbol{p}$ of the distribution functions. It then follows that even and odd modes of the non-equilibrium distributions are decoupled and relax on parametrically different time scales. To see this explicitly let us estimate these rates from Eqs. (2.10) and (2.12). The form of the out-scattering term in each of the linearized kernels suggests introducing the following rates:

$$\Gamma_{\mathrm{pe}}(T)=\int_{p}K(\boldsymbol{p},\boldsymbol{q})\sim\lambda_{\mathrm{ep}}\frac{s}{v_F}T,\qquad \Gamma_{\mathrm{ep}}(T)=\int_{q}K(\boldsymbol{p},\boldsymbol{q})\sim\lambda_{\mathrm{ep}}\frac{T^3}{\omega_D^2}. \tag{2.13}$$

We suppressed here plus/minus subscript in $K(\boldsymbol{p},\boldsymbol{q})$ as phonon absorption and emission processes have the same kinematics. Here, $\lambda_{\mathrm{ep}}=2D_0p_F/(sv_F)$ is the dimensionless electron-phonon coupling constant while D_0 is a constant related to the deformation potential. In what follows we estimate the given, rather distinct, T-dependencies of these two rates.

At low temperatures below the scale of the Debye temperature, $T\ll\omega_D$, we have $\omega_{\boldsymbol{q}}\sim T$ and $\varepsilon_{\boldsymbol{p}}-\varepsilon_F\sim T$, so that $f_\varepsilon\sim b_\omega\sim1$. Furthermore, the typical scale of the phonon momentum is $q\sim T/s$, which is small compared to electronic momentum p_F, where s is the sound velocity. For this reason, the delta-function in the kernel of the collision term can be simplified $\delta(\varepsilon_{\boldsymbol{p}}\pm\omega_{\boldsymbol{q}}-\varepsilon_{\boldsymbol{p}'})\approx\frac{1}{v_Fq}\delta(\cos\theta_{\boldsymbol{pq}}\pm s/v_F)$. Since $s/v_F\ll1$ it is clear that $\theta_{\boldsymbol{pq}}\sim\pi/2$ so that the phonon propagates in a direction that is almost perpendicular to the direction of the electronic momentum. In the phonon-electron scattering rate, the momentum d^3p integration is taken over the volume of a layer with thickness $\sim T/v_F$ along the Fermi surface, so that $\int_p\to\nu\int d\varepsilon d\Omega$ where the solid angle is $d\Omega=2\pi\sin\theta d\theta$ and ν is the density of states at the Fermi level ($\nu=mp_F$ for a

3D metal with spherical Fermi surface). The angular average brings a factor $1/(v_F q) \sim s/(Tv_F)$ from the delta function. Another factor of T comes from $d\varepsilon$ and another T from ω_q in the scattering probability $W \sim D_0(\omega_{\boldsymbol{q}}/\omega_D)$. As a result $\Gamma_{\text{pe}} \sim D_0\nu(s/v_F)(T/\omega_D)$. The electron-phonon relaxation rate is estimated in exactly the same fashion, the only difference is that the integration goes over the phase space of a phonon such that $\int_q$ gives a factor $(T/s)^3$. Combined with the factor $1/(v_F q)$ from the delta function, and a factor $\omega_{\boldsymbol{q}}/\omega_D$ from the scattering probability, this gives $\Gamma_{\text{ep}} \simeq D_0 T^3/(v_F s^2 \omega_D)$. The rate Γ_{ep} defines the typical relaxation scale for even modes e.g. the energy relaxation. With the above given definition of λ_{ep} this yields our estimates for the distinct relaxation rates of electrons and phonons given in Eq. (2.13).

The electronic momentum is relaxed on a different scale. This is not immediately clear from the form of Γ_{ep} itself but rather dictated by kinematic considerations. Indeed, during a given scattering event, the angle between the momenta of the incoming and outgoing electron is small, $\theta_{\boldsymbol{pp}'} \sim q/p_F \sim T/\omega_D$, and the change in electron momentum is $\delta p \sim q^2/p_F \ll p_F$. Thus electrons effectively diffuse in momentum space. We can easily estimate the corresponding diffusion coefficient B from the Einstein relation $\delta p^2 \sim B\tau$, where $\tau \sim \Gamma_{\text{ep}}^{-1}$ is the typical time scale between two consecutive collisions. This gives for $B \propto T^5$. The corresponding mean-free time for momentum relaxation, namely the time needed to change the momentum from to its initial value, is then $\tau_{\text{ep}}^{-1} \sim B/p_F^2 \sim \omega_D(T/\omega_D)^5$. We can estimate the frequency of collisions of phonons with electrons in the same manner, we only need to account for the ratio between the number of electrons and the number of phonons in the region of Fermi function smearing which is of the order $\sim (T/\varepsilon_F)(T/\omega_D)^{-3}$. This implies the collision frequency per phonon occurring with the rate $\tau_{\text{pe}}^{-1} \sim \omega_D(T/\varepsilon_F)$. Hence, phonons are short-lived compared to electrons which can, for example, be used as a justification to integrate out the lattice degrees of freedoms as fast intermediate excitations. Such phononic states are therefore tied to the out-of-equilibrium dynamics of the electrons.

2.3. Bloch law and its violation under complete drag

As a first step in our analysis, it is useful to revisit the solution of the linearized Boltzmann equation for the case of equilibrium phonons (namely neglecting the drag effect). This computation contains all the technical elements that appear in the general calculation and is helpful methodologically. When phonons are assumed to be in equilibrium, we can set $\phi_{\boldsymbol{q}}$ to zero in the linearized collision integral $\delta\,\mathrm{St}\{\psi,\phi\}$ of Eq. (2.12). Thus we are looking for a solution of the following linear integral equation

$$e\boldsymbol{vE}\frac{\partial f}{\partial\varepsilon} = \int_{\boldsymbol{q}} K_-(\boldsymbol{p},\boldsymbol{q})\,[\psi_{\boldsymbol{p}'} - \psi_{\boldsymbol{p}}] - \int_{\boldsymbol{q}} K_+(\boldsymbol{p},\boldsymbol{q})\,[\psi_{\boldsymbol{p}'} - \psi_{\boldsymbol{p}}]\,. \tag{2.14}$$

The fact that the left-hand-side is odd in momentum and that kernels preserve the parity of the function tells us that $\psi_{\boldsymbol{p}}$ must be odd as well. Since $\boldsymbol{Ev}$ contains only one (first) spherical harmonic we chose a trial solution of the form

$$\psi_{\boldsymbol{p}} = \frac{e\boldsymbol{vE}\tau_D}{T}\chi(\eta_{\boldsymbol{p}}), \quad \eta_{\boldsymbol{p}} = (\varepsilon_{\boldsymbol{p}} - \varepsilon_F)/T, \tag{2.15}$$

where time τ_D is introduced to have correct dimensionality which happens to be the characteristic relaxation time of electron-phonon collisions at $T \sim \omega_D$. The terms with $\psi_{\boldsymbol{p}}$ and $\psi_{\boldsymbol{p}'}$ have different angular structure because the electric field has to be projected onto the initial or final momentum respectively. To resolve this difficulty we proceed as follows. Let us choose the integration z-axis in momentum space to be along the initial momentum $\boldsymbol{p}$. Then in the terms $\psi_{\boldsymbol{p}'} \propto (\boldsymbol{p}'\boldsymbol{E})\chi(\eta_{\boldsymbol{p}'})$ we can rewrite $\boldsymbol{p}'\boldsymbol{E} = p'_z E_z + \boldsymbol{p}'_\perp \boldsymbol{E}_\perp$ which implies an angular decomposition

$$\cos\theta_{\boldsymbol{p}'\boldsymbol{E}} = \cos\theta_{\boldsymbol{pp}'}\cos\theta_{\boldsymbol{pE}} + \sin\theta_{\boldsymbol{pp}'}\sin\theta_{\boldsymbol{pE}}\cos\varphi_{\boldsymbol{p}'\boldsymbol{E}} \tag{2.16}$$

where $\varphi_{\boldsymbol{p}'\boldsymbol{E}}$ is the angle between projections of $\boldsymbol{p}'$ and $\boldsymbol{E}$ on the plane perpendicular to the direction of $\boldsymbol{p}$. Note that conservation of momentum and energy fixes the relationship between the angles $\theta_{\boldsymbol{pp}'}$ and $\theta_{\boldsymbol{pq}}$. Upon integration over the angle $\varphi_{\boldsymbol{p}'\boldsymbol{E}}$ the second term vanishes since we have assumed that kernels $K_\pm$ are isotropic and $\chi(\eta_{\boldsymbol{p}})$ does not depend on the direction of momentum by construction. As a result, we accumulate an extra term $\propto \cos\theta_{\boldsymbol{pp}'} \approx (1-\theta^2_{\boldsymbol{pp}'}/2)$ in the differential scattering cross-section. This is noting else but the usual angular factor in the transport scattering time. After the angular part of the integration is done, the integration over the absolute value of momentum $\boldsymbol{q}$ can be brought to a dimensionless form. Combining contributions from both K_- and K_+, we arrive at

$$\cosh^{-2}(\eta/2) \simeq -\vartheta_D^3 \int_{\eta'} K_0(\eta,\eta')\chi(\eta') + \vartheta_D^3\vartheta_F \int_{\eta'} K_1(\eta,\eta')\chi(\eta') + \vartheta_D^5 \int_{\eta'} K_2(\eta,\eta')\chi(\eta'), \tag{2.17}$$

where $\vartheta_D = T/\omega_D$, and $\vartheta_F = T/\varepsilon_F$, while

$$K_k(\eta,\eta') = (\eta-\eta')^k K(\eta,\eta') \tag{2.18}$$

for $k = 0, 1, 2$ and also

$$K(\eta,\eta') = \frac{(\eta-\eta')^2}{(1+e^{-\eta})(1+e^{-\eta'})|e^{\eta}-e^{\eta'}|}. \tag{2.19}$$

The semi-equality sign $\simeq$ in above equation (2.17) implies that we kept the main parametric and functional dependences on the right-hand-side, but we suppressed all the numerical pre-factors of the order of unity in each of the three terms. Retaining these numerical factors will be done in Appendix A.1. Without the last two terms in Eq. (2.17) this equation has no solution for $\chi(\eta)$. This is the consequence of the symmetry of the kernel and the fact that uniform solution is not orthogonal to the left hand-side which is easy to check. The solution can be then found by perturbation theory treating the last two terms as corrections. The term with K_1 does not contribute to the leading order, as it is odd, while the second term gives

$$\chi(\eta) = c/\vartheta_D^5 \tag{2.20}$$

with the constant c is determined by the double integral $c^{-1} = \frac{1}{4}\iint K_2(\eta,\eta')d\eta d\eta'$. With this solution at hand we can compute the electrical current

$$\boldsymbol{j} = \frac{e^2\tau_D}{T}\int_{\boldsymbol{p}} \boldsymbol{v}(\boldsymbol{v}\boldsymbol{E}) f_{\eta_{\boldsymbol{p}}}(1-f_{\eta_{\boldsymbol{p}}})\chi(\eta_{\boldsymbol{p}}) = \sigma_B \boldsymbol{E}, \tag{2.21}$$

with $\sigma_B = ne^2\tau_1/m$, and $\tau_1^{-1} \sim \lambda_{\text{ep}} T^5/\omega_D^4$, where dimensionless coupling constant of the electron-phonon interaction λ_{ep} was introduced earlier in the text below Eq. (2.13). An alternative derivation of the above formula based on the variational analysis of the functional corresponding to the Boltzmann equation (2.14) is presented in Appendix A.1. This approach rather easily allows to fix the numerical pre-factor in σ_B and can be naturally generalized for the calculation of other kinetic coefficients, such as thermal conductivity for example.

As the next methodological step, it is instructive to investigate an opposite extreme limit of complete drag when the non-equilibrium electronic and bosonic distributions are locked together. For this case we need to solve two coupled equations

$$e\boldsymbol{v}\boldsymbol{E}\frac{\partial f}{\partial\varepsilon} = \delta\,\mathrm{St}_{\text{ep}}\{\psi,\phi\}, \quad \delta\,\mathrm{St}_{\text{pe}}\{\psi,\phi\} = 0. \tag{2.22}$$

From the second of these equations we can find bosonic function explicitly as an integral over the fermionic function [see Eq. (2.10)]

$$\phi_{\boldsymbol{q}} = \frac{1}{\Gamma_{\text{pe}}}\int_{\boldsymbol{p}} K_-(\boldsymbol{p},\boldsymbol{q})\,[\psi_{\boldsymbol{p}} - \psi_{\boldsymbol{p}'}] \tag{2.23}$$

and insert it back into the first equation. Then repeating all the same steps as above we obtain instead of Eq. (2.17)

$$\cosh^{-2}(\eta/2) \simeq -\vartheta_D^3 \int_{\eta'} K_0(\eta,\eta')\chi(\eta') + \vartheta_D^5 \int_{\eta'} [K_2(\eta,\eta') - K_d(\eta,\eta')]\chi(\eta'), \tag{2.24}$$

where a contribution with ϑ_F was omitted for brevity as it only gives a sub-leading corrections. The crucial new piece is the drag kernel which has the following form

$$K_d = \frac{1}{2(e^{\eta}+1)}\int_{\zeta} |\zeta|^3 \frac{e^{\zeta}+1}{e^{\zeta}+e^{-\eta}}\frac{e^{\eta'}}{(e^{\eta'}+e^{-\zeta})(e^{\eta'}+e^{\zeta})}. \tag{2.25}$$

It can be shown that $\int [K_2 - K_d] = 0$ where the integration could be either over η or η'. We can now integrate both sides of Eq. (2.24) over η to demonstrate that it has no solution. Physically this is the regime of infinite conductivity that can only be stabilized by momentum-relaxing collisions.

2.4. Super-diffusive dynamics in phase space

The same technique that we used to analyze the resistivity can be applied to determine the viscosity, as we show in the subsequent section. The scattering time τ_1 listed below Eq. (2.21) is basically $\tau_{\rm ep}$ discussed earlier. The subscript $l = 1$ was introduced to emphasize that this time corresponds to the relaxation of a particular harmonic of the distribution function. We will determine below the viscosity that is determined by a parametrically similar time scale but that corresponds to a relaxation of the different harmonic $l = 2$. In Appendix A.4 we analyze in some detail the relaxation τ_l of arbitrary l by performing the angular momentum expansion of the collision term. We obtain

$$\tau_l^{-1} = \begin{cases} l\left(l+1\right) 240\zeta\left(5\right)\lambda_{\rm ep} T^5/\omega_D^4 & \text{if } T \ll \omega_D/l \\ 24\zeta\left(3\right)\lambda_{\rm ep} T^3/\omega_D^2 & \text{if } \omega_D/l \ll T \ll \omega_D \\ \left(1-\delta_{l,0}\right)\left(2-\delta_{l,1}\right)\lambda_{\rm ep} T & \text{if } T \gg \omega_D \end{cases} \tag{2.26}$$

where the intermediate regime only exists at large l. One has to be careful with the behavior above the Debye energy as we neglected drag corrections in the source terms and collision integrals that might correct for the numerical coefficient $2\lambda_{\rm ep}$ for $l \geq 2$.

These results tell us that at lowest temperatures or for small angular momentum the collision operator can be written as an angular Laplacian $\left(\frac{1}{2}\tau_1^{-1}\hat{\mathbf{L}}^2\right)$, which corresponds to a diffusion on the Fermi surface. However, at any finite temperature there are angular momentum modes where we get super-diffusion. The temperature dependence of the rate of super-diffusion for $T \gg \omega_D$ is given by the phonon scattering rate $\Gamma_{\rm ep} \propto T^3$, introduced in Eq. (2.13).

We remind that the term *super-diffusion* is commonly used in the literature to describe the anomalous diffusion equation

$$(\partial_t - D\left|\Delta_{\boldsymbol{p}}\right|^{\mu/2}) n(\boldsymbol{p}, t) = 0 \tag{2.27}$$

with the exponent $\mu < 2$, whereas the case $\mu > 2$ is typically termed *sub-diffusion.* The fractional derivative should be understood via the action of $|\boldsymbol{p}|^{\mu/2}$ in Fourier space. In our case we have $\mu = 2$ at lowest temperatures while highest angular momentum states ultimately behave as $\mu \to 0$. Notice, here diffusion takes place in phase space as a consequence of collisions. Such behavior is of importance if one analyzes the relaxation of focussed electron beams or the time dependence of heat pulses [61–64]. Related behavior was previously discussed in the context of two-dimensional Fermi liquids [61–63] and for graphene at the neutrality point [64]. With electron-phonon fluids we have identified three-dimensional systems that should display superdiffusive dynamics in phase space.

2.5. Emergent drift velocity and conservation laws

In this section we use the conservation laws of the system without umklapp and impurity scattering to establish that a joint drift velocity emerges as hydrodynamic variable. The reason for the joint drift velocity is rather transparent. Only the total momentum $\boldsymbol{P}_{\rm tot}$ is conserved, which gives rise to only one canonically conjugate hydrodynamic variable, the drift velocity $\boldsymbol{u}\left(\boldsymbol{r}\right)$.

We start from the second law of thermodynamics as it enters the Boltzmann theory in the context of the H-theorem. To this end, we consider the entropy per degree of freedom expressed in terms of the distribution functions:

$$s_{\boldsymbol{p}}^{\rm el} = -\left[n_{\boldsymbol{p}} \ln n_{\boldsymbol{p}} + \left(1-n_{\boldsymbol{p}}\right)\ln\left(1-n_{\boldsymbol{p}}\right)\right], \quad s_{\boldsymbol{q}}^{\rm ph} = -\left[N_{\boldsymbol{q}} \ln N_{\boldsymbol{q}} - \left(1+N_{\boldsymbol{q}}\right)\ln\left(1+N_{\boldsymbol{q}}\right)\right]. \tag{2.28}$$

This allows to determine the total entropy production

$$Q \equiv \frac{\partial S}{\partial t} = \frac{\partial}{\partial t}\left(\int_{\boldsymbol{p}} s_{\boldsymbol{p}}^{\rm el} + \int_{\boldsymbol{q}} s_{\boldsymbol{q}}^{\rm ph}\right) = \int_{\boldsymbol{p}} \ln\left(\frac{1}{n_{\boldsymbol{p}}} - 1\right)\frac{\partial n_{\boldsymbol{p}}}{\partial t} + \int_{\boldsymbol{q}} \ln\left(\frac{1}{N_{\boldsymbol{q}}} + 1\right)\frac{\partial N_{\boldsymbol{q}}}{\partial t}. \tag{2.29}$$

We can use the Boltzmann equations, Eqs. (2.1) and (2.2), to express $\frac{\partial n_{\boldsymbol{p}}}{\partial t}$ and $\frac{\partial N_{\boldsymbol{q}}}{\partial t}$. For closed systems it follows after a few steps that

$$Q=-\int_{\boldsymbol{p}}\ln\left(\frac{1}{n_{\boldsymbol{p}}}-1\right)\left(\mathrm{St}_{\mathrm{ep}}\left\{n,N\right\}+\mathrm{St}_{\mathrm{ei}}\left\{n\right\}\right)-\int_{\boldsymbol{q}}\ln\left(\frac{1}{N_{\boldsymbol{q}}}+1\right)\left(\mathrm{St}_{\mathrm{pe}}\left\{n,N\right\}+\mathrm{St}_{\mathrm{pp}}\left\{N\right\}\right)\geq 0, \quad (2.30)$$

where the last inequality reflects the fact that the entropy of the system cannot decrease. In addition we used $\mathrm{St}_{\mathrm{pp}}\left\{N\right\}=\mathrm{St}_{\mathrm{pp}}^{N}\left\{N\right\}+\mathrm{St}_{\mathrm{pp}}^{U}\left\{N\right\}$ which combines normal and umklapp phonon-phonon processes.

Next we summarize the well-known implications of conservation laws. For charge conservation we sum Eq. (2.1) over $\boldsymbol{p}$ and obtain the continuity equation

$$\frac{\partial\rho}{\partial t}+\nabla_{\mathbf{r}}\cdot\boldsymbol{j}=0 \quad (2.31)$$

since $\int_{\boldsymbol{p}}\left(\mathrm{St}_{\mathrm{ep}}\left\{n,N\right\}+\mathrm{St}_{\mathrm{ei}}\left\{n\right\}\right)=0$. Here we have the charge density $\rho\left(\mathbf{r},t\right)=e\int_{\mathbf{p}}n_{\boldsymbol{p}}\left(\mathbf{r},t\right)$ and the current density $\boldsymbol{j}\left(\boldsymbol{r},t\right)=e\int_{\boldsymbol{p}}\boldsymbol{v}_{\boldsymbol{p}}n_{\boldsymbol{p}}\left(\boldsymbol{p},t\right)$. To analyze energy conservation we introduce the energy density and energy current of the combined system:

$$\varepsilon\left(\boldsymbol{r},t\right)=2\int_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}n_{\boldsymbol{p}}\left(\boldsymbol{r},t\right)+\int_{\boldsymbol{q}}\omega_{\boldsymbol{q}}N_{\boldsymbol{q}}\left(\boldsymbol{r},t\right), \quad (2.32)$$

$$\boldsymbol{j}_{\varepsilon}\left(\boldsymbol{r},t\right)=2\int_{\boldsymbol{p}}\boldsymbol{v}_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}n_{\boldsymbol{p}}\left(\boldsymbol{r},t\right)+\int_{\boldsymbol{q}}\boldsymbol{s}_{\boldsymbol{q}}\omega_{\boldsymbol{q}}N_{\boldsymbol{q}}\left(\boldsymbol{r},t\right). \quad (2.33)$$

Multiplying the Boltzmann equations by the electron and phonon energies and integrating over momenta, we obtain

$$\frac{\partial\rho_{\varepsilon}}{\partial t}+\nabla_{\mathbf{r}}\cdot j_{\varepsilon}=2\int_{\boldsymbol{p}}(\dot{\boldsymbol{p}}\cdot\boldsymbol{v}_{\boldsymbol{p}})n_{\boldsymbol{p}}+\int_{\boldsymbol{q}}(\dot{\boldsymbol{q}}\cdot\boldsymbol{s}_{\boldsymbol{q}})N_{\boldsymbol{q}}. \quad (2.34)$$

If there is no work done by or at the system ($\dot{\boldsymbol{p}}\cdot\boldsymbol{v}_{\boldsymbol{p}}=\dot{\boldsymbol{q}}\cdot\boldsymbol{s}_{\boldsymbol{q}}=0$) this corresponds to the the continuity equation for the energy. It is a consequence of the fact that the sum of $\int_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}\left(\mathrm{St}_{\mathrm{ep}}\left\{n,N\right\}+\mathrm{St}_{\mathrm{ei}}\left\{n\right\}\right)$ and $\int_{\mathbf{q}}\omega_{\boldsymbol{q}}\left(\mathrm{St}_{\mathrm{pe}}\left\{n,N\right\}+\mathrm{St}_{\mathrm{pp}}\left\{N\right\}\right)$ vanishes. Finally we consider the momentum density and momentum current:

$$\begin{aligned}\boldsymbol{g}\left(\boldsymbol{r},t\right)&=2\int_{\boldsymbol{p}}\boldsymbol{p}n_{\boldsymbol{p}}\left(\boldsymbol{r},t\right)+\int_{\boldsymbol{q}}\boldsymbol{q}N_{\boldsymbol{q}}\left(\mathbf{r},t\right),\\ T_{\alpha\beta}\left(\boldsymbol{r},t\right)&=2\int_{\boldsymbol{p}}p_{\alpha}v_{\beta}n_{\boldsymbol{p}}\left(\boldsymbol{r},t\right)+\int_{\boldsymbol{q}}q_{\alpha}s_{\beta}N_{\boldsymbol{q}}\left(\boldsymbol{r},t\right).\end{aligned} \quad (2.35)$$

In the absence of impurity and umklapp scattering, i.e. for $\mathrm{St}_{\mathrm{ei}}\left\{n\right\}=\mathrm{St}_{\mathrm{ep}}^{U}\left\{n,N\right\}=\mathrm{St}_{\mathrm{pp}}^{U}\left\{N\right\}=0$. We obtain

$$\frac{\partial g_{\alpha}}{\partial t}+\frac{\partial T_{\alpha\beta}}{\partial x_{\beta}}=2\int_{\boldsymbol{p}}\dot{p}_{\alpha}n_{\boldsymbol{p}}+\int_{\boldsymbol{q}}\dot{q}_{\alpha}N_{\boldsymbol{q}}, \quad (2.36)$$

which becomes the momentum continuity equation in the absence of external forces ($\dot{\boldsymbol{p}}=\dot{\boldsymbol{q}}=0$). The continuity equation follows because the sum of $\int_{\boldsymbol{p}}\boldsymbol{p}\left(\mathrm{St}_{\mathrm{ep}}^{N}\left\{n,N\right\}\right)$ and $\int_{\boldsymbol{q}}\boldsymbol{q}\left(\mathrm{St}_{\mathrm{pe}}^{N}\left\{n,N\right\}+\mathrm{St}_{\mathrm{pp}}^{N}\left\{N\right\}\right)$ vanishes.

Let us now search for distribution functions that yield a constant entropy. Under the given conservation laws the entropy production Q, as given in Eq. (2.30), vanishes for the distributions

$$\begin{aligned}\ln\left(\frac{1}{n_{\boldsymbol{p}}}-1\right)&=-\beta\left(\boldsymbol{r}\right)\mu\left(\boldsymbol{r}\right)+\beta\left(\boldsymbol{r}\right)\varepsilon_{\boldsymbol{p}}-\beta\left(\boldsymbol{r}\right)\boldsymbol{u}\left(\boldsymbol{r}\right)\cdot\boldsymbol{p},\\ \ln\left(\frac{1}{N_{\boldsymbol{q}}}+1\right)&=\beta\left(\boldsymbol{r}\right)\omega_{\boldsymbol{q}}-\beta\left(\boldsymbol{r}\right)\boldsymbol{u}\left(\boldsymbol{r}\right)\cdot\boldsymbol{q},\end{aligned} \quad (2.37)$$

with same $\beta\left(\boldsymbol{r}\right)$ an $\boldsymbol{u}\left(\boldsymbol{r}\right)$ in the two equations. This gives rise to local equilibrium with Fermi-Dirac distribution function for the electrons

$$n_{\boldsymbol{p}}\left(\boldsymbol{r}\right) = \frac{1}{e^{\beta(\boldsymbol{r})(\varepsilon_{\boldsymbol{p}}-\mu(\boldsymbol{r})-\boldsymbol{u}(\boldsymbol{r})\cdot\boldsymbol{p})}+1}, \quad (2.38)$$

and Bose-Einstein distribution for the phonons

$$N_{\boldsymbol{q}}(\boldsymbol{r}) = \frac{1}{e^{\beta(\boldsymbol{r})(\omega_{\boldsymbol{q}} - \boldsymbol{u}(\boldsymbol{r})\cdot\boldsymbol{p})} - 1}. \tag{2.39}$$

Obviously we have the usual interpretation of $\beta(\boldsymbol{r})$, $\mu(\boldsymbol{r})$ and $\boldsymbol{u}(\boldsymbol{r})$ as local inverse temperature, chemical potential of the electrons, and flow velocity, respectively.

Just like the conservation of the total energy gives rise to a joint temperature of the electrons and phonons, does the conservation of the total momentum yield a joint drift velocity $\boldsymbol{u}(\boldsymbol{r})$. While the local equilibrium is only a solution of the Boltzmann equation in the limit where the collision terms dominate, they do represent a natural starting point in the limit of small Knudsen number – i.e. the ratio of the momentum conserving mean free path and the typical length scale of applied forces or geometric confinement – as employed by the Chapman-Enskog method [65].

The hydrodynamic flow is protected by the conservation of the total momentum

$$\boldsymbol{P}_{\mathrm{tot}} = \boldsymbol{P}_{\mathrm{el}} + \boldsymbol{P}_{\mathrm{ph}} \tag{2.40}$$

and must be understood as a combined electron-phonon fluid. However, in Appendix A.2 we demonstrate that the primary mechanism by which the flow gradient couples to the electron-phonon fluid is by directly affecting its electron component. In addition we show that while phonon drag is crucial to give the viscosity a true hydrodynamic interpretation, perhaps counterintuitively, it is not important for the actual value of the viscosity. Finally, because of the larger value of the Fermi velocity and because of the different phase space nature of degenerate electrons and acoustic phonons, it holds that the momentum current is also dominated by the electronic system.

2.6. Hydrodynamic electron-phonon drag viscosity

Provided that momentum-conserving electron-phonon collisions are the most frequent, the regime of phonon drag can be characterized by an emergent hydrodynamic mode, which is the drift velocity of electrons and phonons. Indeed, both collision terms $\mathrm{St}_{\mathrm{ep}}\{n, N\}$ and $\mathrm{St}_{\mathrm{pe}}\{n, N\}$ are simultaneously solved by a distribution function with the finite boost $n(\boldsymbol{p}, \boldsymbol{r}) = f(\varepsilon_{\boldsymbol{p}} - \boldsymbol{p}\boldsymbol{u}(\boldsymbol{r}))$ and $N(\boldsymbol{p}, \boldsymbol{r}) = b(\omega_{\boldsymbol{q}} - \boldsymbol{q}\boldsymbol{u}(\boldsymbol{r}))$. In the previous sub-section we discuss the origin of the joint drift velocity as conjugated variable to the conserved total momentum in some detail.

To determine the equation of motion for $\boldsymbol{u}(\boldsymbol{r})$ we follow the approach of Gurzhi [51] who solved the kinetic equations by the method of consecutive approximations. The accuracy of the method is controlled by the ratio between momentum-conserving and momentum-relaxing scattering lengths. We seek the non-equilibrium distribution functions in the form of a formal series expansion: $n = f + \delta n_1 + \delta n_2 + \ldots\ldots$ and $N = b + \delta N_1 + \delta N_2$. To the first order we obtain two equations:

$$\boldsymbol{v}\frac{\partial f}{\partial \boldsymbol{r}} = \delta\,\mathrm{St}_{\mathrm{ep}}\{\delta n_1, \delta N_1\}, \quad \boldsymbol{s}\frac{\partial b}{\partial \boldsymbol{r}} = \delta\,\mathrm{St}_{\mathrm{pe}}\{\delta n_1, \delta N_1\}, \tag{2.41}$$

where linearized collision kernels are given by Eqs. (2.10) and (2.12). The contribution from the collision term with normal phonon processes, governed by $\delta\,\mathrm{St}^{N}\{\delta N_1\}$, can be neglected as it has a subdominant temperature dependence in comparison with phonon-electron collisions. Since the spatial dependency of the distribution is contained in the velocity field $\boldsymbol{u}(\boldsymbol{r})$ we search for a solution of the form

$$\delta n_1 = -v_i p_j \frac{\partial u_j}{\partial r_i}\frac{\partial f}{\partial \varepsilon}\tau_D(\omega_D/T)^3\chi(\eta_{\boldsymbol{p}}), \tag{2.42}$$

$$\delta N_1 = -s_i q_j \frac{\partial u_j}{\partial r_i}\tau_D(T/ms^2)\frac{\partial b}{\partial \omega}\phi(\zeta_{\boldsymbol{q}}), \tag{2.43}$$

where $\zeta_{\boldsymbol{q}} = \omega_{\boldsymbol{q}}/T$. Again repeating all the same technical steps from the previous section, where we discussed Bloch's solution of the linearized kinetic equations, we find two coupled integral equations for the non-equilibrium distributions ψ and ϕ:

$$\cosh^{-2}(\eta/2) \simeq -\int_{\eta'} K_0(\eta,\eta')\chi(\eta') + \vartheta_D^2 \int_{\eta'} K_2(\eta,\eta')\chi(\eta') - \vartheta_D^2 \int_{\eta'} K_1(\eta,\eta')\phi(\eta-\eta'), \tag{2.44}$$

$$\zeta/(e^{\zeta}-1) \simeq \vartheta_D^2\zeta^2\phi(\zeta) + (1-\vartheta_D^2\zeta^2)\int_{\zeta'} Q(\zeta,\zeta')\chi(\zeta'), \tag{2.45}$$

where

$$Q(\zeta,\zeta') = \frac{(e^{\zeta}-1)(e^{\zeta'}-1)}{(e^{\zeta}+1)(e^{\zeta}+e^{\zeta'})(e^{\zeta}+e^{-\zeta'})}. \tag{2.46}$$

We were unsuccessful in finding an analytical solution of these equations. However, exploring the smallness of $\vartheta_D \ll 1$ it is possible to show that $\chi(\eta) \sim \vartheta_D^{-2}$.

At the second order of the expansion, the set of equations takes the form

$$\boldsymbol{v}\frac{\partial \delta n_1}{\partial \boldsymbol{r}} + e\boldsymbol{v}\boldsymbol{E}\frac{\partial f}{\partial \varepsilon} = \delta\,\mathrm{St}_{\mathrm{ep}}\{\delta n_2, \delta N_2\} + \delta\,\mathrm{St}_{\mathrm{ei}}\{f\}, \tag{2.47}$$

$$\boldsymbol{s}\frac{\partial \delta N_1}{\partial \boldsymbol{r}} = \delta\,\mathrm{St}_{\mathrm{pe}}\{\delta n_2, \delta N_2\} + \delta\,\mathrm{St}^U_{\mathrm{pp}}\{b\}. \tag{2.48}$$

It is important to emphasize at this point that $\delta\,\mathrm{St}_{\mathrm{ei}}\{f\} \propto \boldsymbol{u}(\boldsymbol{r})$ and similarly $\delta\,\mathrm{St}^U_{\mathrm{pp}}\{f\} \propto \boldsymbol{u}(\boldsymbol{r})$ as these two terms capture momentum-relaxing collisions and as such will define the relaxation of $\boldsymbol{u}$. Finally, we use the explicit form of δn_1 from Eq. (2.42), multiply Eq. (2.47) by $\boldsymbol{p}$ and integrate both sides over momentum. Similarly we use δN_1 from Eq. (2.43) in Eq. (2.48), multiply by $\boldsymbol{q}$ and integrate both sides. We then add together these equations and obtain the desired hydrodynamic equation for $\boldsymbol{u}(\boldsymbol{r})$ (see also Refs. [51, 57]):

$$\nu\nabla^2\boldsymbol{u} + e\boldsymbol{E}/m = \boldsymbol{u}/\tau_{\mathrm{MR}}. \tag{2.49}$$

Here momentum-relaxation time $\tau_{\mathrm{MR}}^{-1} = \tau_{\mathrm{ei}}^{-1} + \tau_U^{-1}$ is given by the sum of two terms due to electron-impurity and phonon umklapp scattering. While the former is temperature independent, the latter has steep exponential behavior $\tau_U^{-1} \propto (T/\omega_D)^4(\tau^U_{\mathrm{pp}})^{-1}$, with $(\tau^U_{\mathrm{pp}})^{-1} \propto \exp(-\gamma\omega_D/T)$ and $\gamma \sim 1$. The kinematic viscosity of the electron-phonon fluid $\nu = \eta_{\mathrm{ep}}/mn$ in Eq. (2.49) is expressed in terms of the corresponding shear viscosity in a standard way:

$$\eta_{\mathrm{ep}} = \frac{1}{5}mnv_F^2\tau_2, \quad \tau_2^{-1} = 1440\zeta(5)\lambda_{\mathrm{ep}}T^5/\omega_D^4. \tag{2.50}$$

For the detailed derivation of Eq. (2.50) see Appendix A.2. Notice that the functional form of Eq. (2.49) is formally identical to the equation of motion of an electron fluid where the hydrodynamic regime is established by electron-electron collisions. The difference is only in the temperature dependence of the viscosity, i.e. of the relaxation time τ_2. The electron-phonon collisions that give rise to a T^5 Bloch-Grüneisen law in the resistivity of the kinetic regime are the same processes that determine the viscosity $\eta_{\mathrm{ep}} \propto T^{-5}$ in the hydrodynamic regime. This parallels the electron-electron hydrodynamic regime where the T^2 term in the resistivity translates into $\eta_{\mathrm{ee}} \propto T^{-2}$ for the electron viscosity [66]. In closing this section we also wish to draw attention to an analogy between phonon drag viscosity and recently studied Coulomb drag viscosity contribution [67], and its relation to hydrodynamic drag resistivity in the transport properties of interactively coupled double-layers [68].

2.7. Thermal conductivity and the Lorentz ratio in a drag regime

The theory of thermal conductivity in the hydrodynamic regime of a phonon gas was put forward in pioneering works of Callaway [69] and Gurzhi [70] (the classical review on the topic can be found in Ref. [71], whereas a concise summary of the field with the modern perspective can be found in Ref. [72]). These authors carefully analyzed the interplay of various scattering processes including (i) sample boundary scattering, described by a constant relaxation time; (ii) three-phonon nonlinearities, whose relaxation time is a power-law of temperature; (iii) impurity scattering; (iv) umklapp processes with an exponential relaxation time. The resulting thermal conductivity was shown to exhibit fairly complicated non-monotonic behavior. Recently phonon-mediated heat diffusion in insulators received a renewed attention and interest triggered by a realization of apparently universal bound controlled by the Plankian time scale, $\tau_{\mathrm{Pl}} \sim (\hbar/k_B T)$, quantum mechanical bound on sound velocity [73, 74], and generalization of Fourier's law into viscous heat equations [75]. In this section we consider the problem of thermal conduction from the perspective of mutual electron-phonon drag and reveal its distinct properties. The corresponding electron-phonon bound on thermal diffusion can be analyzed in a similar spirit as it was done recently in the context of the Coulomb drag problem [76].

The starting point of our treatment is the same set of linearized coupled integro-differential Boltzmann equations as used in the case of conductivity and viscosity calculations in previous sections. The only difference is that we are looking now at the response to the temperature gradient $\nabla_{\boldsymbol{r}} T$, thus we have

$$-\frac{\varepsilon_{\boldsymbol{p}}}{T}\frac{\partial f}{\partial \varepsilon_{\boldsymbol{p}}}(\boldsymbol{v}_{\boldsymbol{p}}\nabla_{\boldsymbol{r}}T) = \delta\,\mathrm{St}_{\mathrm{ep}}\{\delta n_1, \delta N_1\}, \quad -\frac{\omega_{\boldsymbol{q}}}{T}\frac{\partial b}{\partial \omega_{\boldsymbol{q}}}(\boldsymbol{s}_{\boldsymbol{q}}\nabla_{\boldsymbol{r}}T) = \delta\,\mathrm{St}_{\mathrm{pe}}\{\delta n_1, \delta N_1\}. \tag{2.51}$$

It is clear that in the linear response analysis the non-equilibrium corrections to electron and phonon distribution functions are proportional to the thermal bias, namely $\{\delta n_1, \delta N_1\} \propto \nabla T$. Provided that a solution is found, the heat current can be computed in accordance with the usual kinetic formula

$$\boldsymbol{j}_\varepsilon = \int_{\boldsymbol{p}} \boldsymbol{v}_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}\delta n_1 + \int_{\boldsymbol{q}} \boldsymbol{s}_{\boldsymbol{q}}\omega_{\boldsymbol{q}}\delta N_1 = -\kappa_{\mathrm{ep}}\nabla_{\boldsymbol{r}}T, \tag{2.52}$$

that thus defines the electron-phonon drag thermal conductivity κ_{ep}. Just like in the case for the electron viscosity calculation, discussed in the Appendix A.2, we can first solve for the non-equilibrium phonon distribution δN_1 in terms of yet unknown δn_1, and insert the result into the Boltzmann equation for the electrons. This yields then purely electronic Boltzmann equation of the type

$$\boldsymbol{R}_{\boldsymbol{p}} \cdot \nabla_{\boldsymbol{r}} T = \delta\,\mathrm{St}_{\mathrm{el}}\{\delta n_1\}. \tag{2.53}$$

The source term on the right-hand-side $\boldsymbol{R}_{\boldsymbol{p}} = -\boldsymbol{v}_{\boldsymbol{p}}(\varepsilon_{\boldsymbol{p}}/T)(\partial f/\partial\varepsilon_{\boldsymbol{p}}) + \delta\boldsymbol{R}_{\boldsymbol{p}}$ is renormalized by the drag effect. The collision term $\delta\,\mathrm{St}\{\delta n_1\}$ also contains an additional correction. The analysis of the second term $\boldsymbol{\delta R_p}$ yields the conclusions that it can be neglected at temperatures $T \ll \omega_D$. The subsequent analysis of the collision term is analogous to the one for the viscosity and yields for the thermal conductivity the result (see Appendix A.3 for further details)

$$\kappa_{\mathrm{ep}} = \frac{1}{3}v_F^2 c_{\mathrm{el}}(T)\tau_E, \quad \tau_E^{-1} = 480\zeta(5)\lambda_{\mathrm{ep}}T^3/\omega_D^2, \quad T \ll \omega_D, \tag{2.54}$$

with the electronic heat capacity $c_{\mathrm{el}}(T)$. For higher temperatures it holds that $\tau_E^{-1} \simeq \lambda_{\mathrm{ep}}T$. With the linear low-$T$ heat capacity $c_{\mathrm{el}} \simeq \gamma_s T$, where γ_s is the usual Sommerfeld coefficient, it follows for the thermal conductivity $\kappa_{\mathrm{ep}} \propto 1/T^2$. This is distinct from the thermal conductivity of a Fermi liquid $\kappa_{\mathrm{ee}} \propto 1/T$ [66] and would naturally lead to a temperature dependent Lorentz ratio, $L(T) = \kappa/\sigma T$, quite distinct from the universal Sommerfeld bound of $\pi^2/3e^2$ in the Wiedemann-Franz law. We note that thermal conductivity has been measured recently in Refs. [16, 21] in a phonon drag regime driven by normal electron-phonon scattering processes. The scaling consistent with T^{-2}-behavior in the intermediate range of temperatures was indeed observed in $\mathrm{PtSn_4}$ [21]. Additionally, hydrodynamic features due to electron viscosity accompanied by the size-dependent departure from the Wiedemann-Franz law, expected in the hydrodynamic picture, were observed in recent thermal resistivity measurements in semi-metallic antimony Sb [24]. Similar thermal transport anomalies were also reported in $\mathrm{WP_2}$ [22] and analyzed theoretically in Ref. [77].

3. Applications

3.1. Gurzhi resistance at arbitrary slip length

As a first application let us consider hydrodynamic flow in a two-dimensional slab geometry of width d where the flow occurs in the x-direction such that the velocity field $\boldsymbol{u}(\boldsymbol{r}) = (u(y), 0)$ has a nontrivial profile

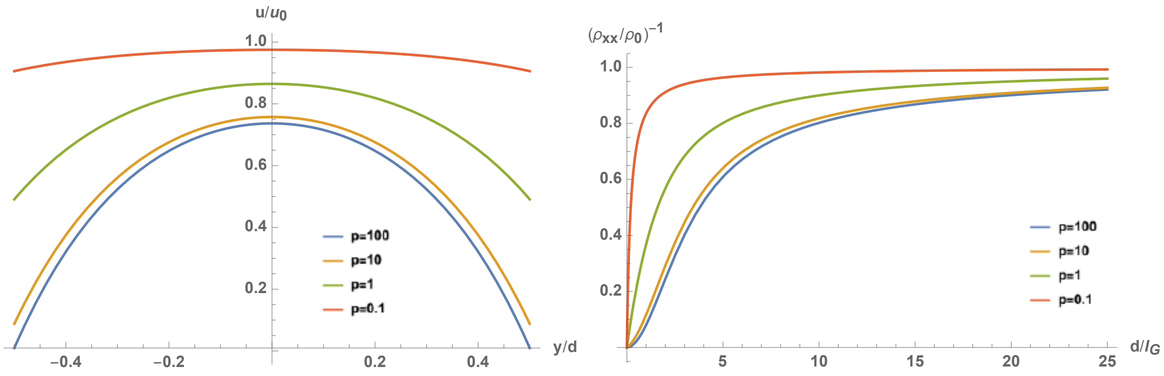


Figure 2: [Left panel]: Spatial profile of the hydrodynamic flow field in the slab geometry for $d/l_\mathrm{G} = 4$ in the crossover regime from no-slip to no-stress boundary conditions. [Right panel]: Dependence of the resistivity as a function of the channel width normalized to the Gurzhi length plotted for several different values between Gurzhi length and slip length.

along the y-direction, where the electric field $\boldsymbol{E} = (E_x, 0)$ is directed along x-direction. The equation of motion (2.49) then becomes

$$\nu \frac{d^2u}{dy^2} + eE_x/m = u/\tau_\mathrm{MR}. \tag{3.1}$$

This equation should be supplemented by a boundary condition. We use a generic one allowing for an arbitrary slip length l_S [78, 79]

$$\left(\frac{du}{dy}\right)_{y=\pm d/2} = \mp \frac{u(\pm d/2)}{l_\mathrm{S}}. \tag{3.2}$$

Solving this linear differential equation we find a flow profile

$$u(y) = u_0 \left[1 - 2p\frac{(1+p)e^{w/2} + (1-p)e^{-w/2}}{(1+p)^2e^{w} - (1-p)^2e^{-w}} \cosh \frac{y}{l_\mathrm{G}}\right]. \tag{3.3}$$

Here we introduced the characteristic steady state velocity u_0, the Gurzhi length l_G, and two dimensionless parameters p, w:

$$u_0 = \frac{eE_x\tau_\mathrm{MR}}{m}, \quad l_\mathrm{G} = \sqrt{\nu\tau_\mathrm{MR}}, \quad p = l_\mathrm{G}/l_\mathrm{S}, \quad w = d/l_\mathrm{G}. \tag{3.4}$$

The no-slip boundary condition corresponds to the limit where $p \to \infty$, whereas the opposite limit $p \to 0$ defines the no-stress regime. The flow profiles at different values of p are illustrated in Fig. 2. We introduce the average flow velocity across the channel

$$\bar{u} = \frac{1}{d}\int_{-d/2}^{d/2} u(y)dy. \tag{3.5}$$

This expression enables us to find current density $j_x = en\bar{u}$ and consequently resistance along the channel

$$\rho_{xx}^{-1} = \rho_0^{-1}\left[1 - \frac{4p}{w}\frac{(1+p)e^{w/2} + (1-p)e^{-w/2}}{(1+p)^2e^{w} - (1-p)^2e^{-w}} \sinh\frac{w}{2}\right], \tag{3.6}$$

where $\rho_0^{-1} = e^2n\tau_\mathrm{MR}/m$ is the familiar formula of the Drude resistivity. This result simplifies in the limit of no slip $p \to \infty$ [60, 80]

$$\rho_{xx}^{-1} = \rho_0^{-1}\left[1 - \frac{2}{w}\tanh\frac{w}{2}\right]. \tag{3.7}$$

For a wide channel, $d \gg l_\mathrm{G}$, the resistance saturates to its bulk value ρ_0 which is governed by the momentum-relaxing time. In contrast, for a narrow channel, $d \ll l_\mathrm{G}$, the resistivity is determined by momentum

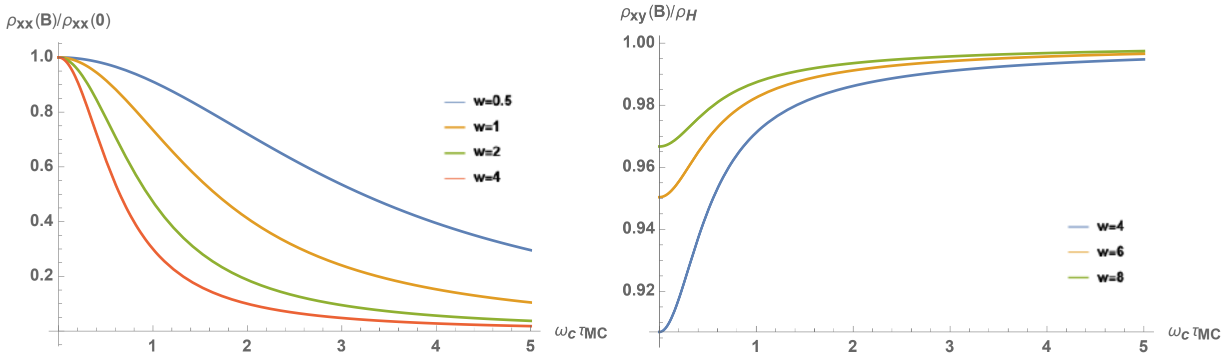


Figure 3: [Left panel]: Field dependence of the diagonal resistivity for different channel width aspect ratios. [Right panel]: Field dependence for the Hall resistivity normalized to the classical Hall resistance.

conserving electron-phonon collisions and inversely proportionally to the channel width as expected for the Poiseuille flow $\rho_{xx} \simeq (p_F/e^2n)(l_{\mathrm{MC}}/d)^2$. This defines the regime of the Gurzhi effect [60] when the resistance drops with increasing temperature as controlled by the momentum conserving length scale $l_{\mathrm{MC}} = v_F\tau_{\mathrm{ep}}$.

As the next step, we briefly investigate the sensitivity of these results to the geometry of the conducting channel. For this purpose we look at the quantum wire (cylindrical geometry) of radius d. Using the Laplacian in radial coordinates the equation of motion and boundary condition take the form

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) - \frac{u}{l_{\mathrm{G}}^2} = -eE_x/m\nu, \qquad (du/dr)_{r=d} = -u(d)/l_{\mathrm{S}}, \tag{3.8}$$

where $\boldsymbol{u}(\boldsymbol{r}) = (u(r), 0, 0)$. This equation is solved in terms of the modified Bessel functions of zero index. However, for a bounded solution at the origin we must retain only I_0 function but not K_0. Recalling then the property of the derivative that $\mathrm{I}_0'(z) = \mathrm{I}_1(z)$ we find

$$u(r) = u_0\left[1 - \frac{p\mathrm{I}_0(r/l_{\mathrm{G}})}{\mathrm{I}_1(w) + p\mathrm{I}_0(w)}\right]. \tag{3.9}$$

Averaging this expression over the wire cross-section and recalling the integral property

$$\int_0^z r\mathrm{I}_0(r)dr = z\mathrm{I}_1(z) \tag{3.10}$$

we find wire resistivity in the form

$$\rho_{xx}^{-1} = \rho_0^{-1}\left[1 - \frac{2p}{w}\frac{\mathrm{I}_1(w)}{\mathrm{I}_1(w) + p\mathrm{I}_0(w)}\right]. \tag{3.11}$$

The flow profile is analogous to that of a slab presented in Fig. (2) with the only difference that it looks flatter at the center of the wire. The resistance also exhibits the same dependency on the ratio d/l_{G}. The only difference is numerical coefficients of the order unity that occur in the respective asymptotic limits.

3.2. Magnetoresistance and Hall resistance

In the presence of an external magnetic field we need to add the Lorentz force as well as a Hall viscosity ν_{H} into the equation of motion. For a steady flow we thus have [80–84]

$$\nu\nabla^2\boldsymbol{u} + \nu_{\mathrm{H}}[\nabla^2\boldsymbol{u} \times \boldsymbol{e}_B] + e(\boldsymbol{E} + [\boldsymbol{u} \times \boldsymbol{B}])/m = \boldsymbol{u}/\tau_{\mathrm{MR}}, \tag{3.12}$$

where $\boldsymbol{e}_B$ is the unit vector along the magnetic field. In the semiclassical approximation, the Hall viscosity can be derived in a manner similar to the classical work of Steinberg [56] (see also the recent discussions by

Alekseev [80] and Scaffidi *et al.* [81]). The difference is that for our case it is assumed that the kinematic viscosity is dominated by electron-phonon collisions instead of electron-electron collisions.

For a Hall bar strip geometry with magnetic field along the z-axis there is no flow in the y-direction due to the build up of an electric field that compensates for the Lorentz force in the classical Hall effect. The resulting equations of motion read

$$\nu\frac{d^2u_x}{dy^2}+eE_x/m=u_x/\tau_{\mathrm{MR}},\quad -\nu_{\mathrm{H}}\frac{d^2u_x}{dy^2}+eE_y/m=\omega_c u_x, \tag{3.13}$$

where $\omega_c=eB/m$ is the cyclotron frequency. The first equation is structurally unchanged as compared to the case of no field, so is solved exactly as in the previous section. To find E_y in the second equation, we integrate this equation over the strip width and get

$$-\frac{\nu_{\mathrm{H}}}{d}\left[\left(\frac{du_x}{dy}\right)_{y=d/2}-\left(\frac{du_x}{dy}\right)_{y=-d/2}\right]+eE_y/m=\omega_c\bar{u}_x \tag{3.14}$$

From the boundary conditions, we can express derivatives of the velocity field in terms of the velocity itself and the slip length

$$\frac{2\nu_{\mathrm{H}}}{dl_{\mathrm{S}}}u_x(d/2)+\frac{eE_y}{m}=\omega_c\bar{u}_x \tag{3.15}$$

This equation yields the Hall field (and voltage) and thus gives us the transverse resistivity

$$\rho_{xy}=\rho_{\mathrm{H}}\left[1-\frac{\nu_{\mathrm{H}}}{d^2\omega_c}H(p,w)\right],\quad \rho_{\mathrm{H}}=\frac{B}{en}. \tag{3.16}$$

The dimensionless function $H(p,w)=2pwu_x(d/2)/\bar{u}_x$ can be found from the longitudinal flow profile of the velocity field and is given by

$$H(p,w)=2pw\frac{1-2pP(p,w)\cosh(w/2)}{1-(4p/w)P(p,w)\sinh(w/2)}, \tag{3.17}$$

where

$$P(p,w)=\frac{(1+p)e^{w/2}+(1-p)e^{-w/2}}{(1+p)^2e^{w}-(1-p)^2e^{-w}}. \tag{3.18}$$

The Hall resistance takes a particularly simple form in the no-slip limit where

$$H=\frac{2w\tanh(w/2)}{1-(2/w)\tanh(w/2)}. \tag{3.19}$$

In the weak-field limit, taking $\nu_{\mathrm{H}}\simeq\nu(\omega_c\tau_{\mathrm{MC}})$, where τ_{MC} is the momentum conserving time scale, given by electron-phonon collisions in our case, we estimate the correction to the Hall resistivity to be of the form

$$\frac{\delta\rho_{xy}}{\rho_{\mathrm{H}}}\simeq-\left(\frac{l_{\mathrm{MC}}}{d}\right)^2\frac{(2d/l_{\mathrm{G}})\tanh(d/2l_{\mathrm{G}})}{1-(2l_{\mathrm{G}}/d)\tanh(d/2l_{\mathrm{G}})}. \tag{3.20}$$

We remind that the underlying assumption for the length scales is such that $l_{\mathrm{MC}}\ll d\ll l_{\mathrm{MR}}$. Note that the Gurzhi length can be equivalently presented as $l_{\mathrm{G}}=\sqrt{l_{\mathrm{MC}}l_{\mathrm{MR}}}$ such that, in principle, the relationship between d and l_{G} can be arbitrary. Having this in mind we conclude that the correction $\delta\rho_{xy}$ is universal in the narrow channel when $d\ll l_{\mathrm{G}}$ where $\delta\rho_{xy}/\rho_{\mathrm{H}}\simeq-(l_{\mathrm{MC}}/d)^2$ while it scales as $\delta\rho_{xy}/\rho_{\mathrm{H}}\simeq-l^2_{\mathrm{MC}}/dl_{\mathrm{G}}$ in the opposite limit. The field dependence of both, the diagonal and the Hall resistivities in the semiclassical limit is illustrated in Fig. (3) for different aspect ratios of the Hall bar channel and different ratios of the channel width and the Gurzhi length, respectively.

3.3. *Stokes-to-Ohm crossover in a swirling magneto-flow*

The 2D cylindrical geometry of a Corbino disk with inner radius r_1 and outer radius r_2 also attracts considerable attention. It was recently suggested that the electronic shear viscosity can be measured with this device in the response to an alternating magnetic flux that generates a measurable (dc) potential drop, induced between the inner and the outer edge of the disk [85]. It also offers new opportunities to experimentally determine the Hall viscosity [86] and the hydrodynamic magnetoresistance that is dominated by the field-induced vorticity of the flow rather than by the field dependence of the kinetic coefficients [87]. Here we elaborate on the latter example focusing on the magnetoresistance in the crossover region of the Gurzhi effect from the Stokes-to-Ohmic flow.

The centro-symmetry of the Corbino disk suggests the use of polar coordinates. For the purpose of MR calculation we need to project the Navier-Stokes equation (3.12) into the radial (u_r) and azimuthal (u_ϕ) components of the flow field. The corresponding components of the Laplacian operator are given by [88]

$$(\nabla^2\boldsymbol{u})_r = \nabla^2 u_r - \frac{u_r}{r^2} - \frac{2}{r^2}\frac{\partial u_\phi}{\partial\phi}, \qquad (\nabla^2\boldsymbol{u})_\phi = \nabla^2 u_\phi - \frac{u_\phi}{r^2} + \frac{2}{r^2}\frac{\partial u_r}{\partial\phi}. \tag{3.21}$$

For an isotropic system with magnetic field perpendicular to the plane of the flow, both components of the flow velocity depend only on the radial coordinate such that terms like $\partial_\phi u_{r,\phi}$ vanish. Consequently, for the corresponding components of the electrical current we find the two equations

$$\frac{\eta}{ne}\Delta j_r + enE_r + j_\phi B = \rho_0 ne j_r, \quad \frac{\eta}{ne}\Delta j_\phi + enE_\phi - j_r B = \rho_0 ne j_\phi, \tag{3.22}$$

where we introduced radial operator $\Delta = \nabla^2 - 1/r^2$ and expressed the kinematic viscosity $\nu = \eta/(mn)$ in terms of shear viscosity η. In the current setup there is no azimuthal component of the electric field $E_\phi = 0$, but there is a freely circulating current j_ϕ. The situation here is opposite to that of the Hall bar, with a transversal field but no current. Furthermore, from the continuity equation, current conservation in the radial direction implies

$$j_r(r) = I/(2\pi r), \tag{3.23}$$

which gives an equation for the azimuthal current

$$(\Delta - l_{\mathrm{G}}^{-2})j_\phi = \left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \left(\frac{1}{r^2} + \frac{1}{l_{\mathrm{G}}^2}\right)\right] j_\phi = \frac{ne}{\eta}\frac{IB}{2\pi r}. \tag{3.24}$$

This equation coincides with the canonical form of the differential equation for the modified Bessel function of the first order, which thus gives us two linearly independent solutions $\mathrm{I}_1(r/l_{\mathrm{G}})$ and $\mathrm{K}_1(r/l_{\mathrm{G}})$. The special solution due to the right-hand-side can be tried in the form $j_\phi = C(I/r)$ where C is a yet unknown constant. By observing that $\Delta(1/r) = 0$ we easily deduce that $C = -B/(2\pi\rho_0 ne)$. As a result, the general solution takes the form

$$j_\phi(r) = \frac{IB}{2\pi\rho_0 ne}\left[C_1\,\mathrm{I}_1(r/l_{\mathrm{G}}) + C_2\,\mathrm{K}_1(r/l_{\mathrm{G}}) - \frac{1}{r}\right]. \tag{3.25}$$

The integration constants C_1 and C_2 can be determined from the boundary conditions. For simplicity, we apply no-slip boundary conditions $j_\phi(r_1) = j_\phi(r_2) = 0$. To visualize viscous effects we deduced the flow pattern from the obtained solution and plotted $\boldsymbol{u}(\boldsymbol{r})$ in Fig. 4.

As next step in our analysis we use the components of the stress tensor [88]

$$\sigma_{rr} = 2\eta\frac{\partial u_r}{\partial r}, \quad \sigma_{r\phi} = \eta\left(\frac{1}{r}\frac{\partial u_r}{\partial\phi} + \frac{\partial u_\phi}{\partial r} - \frac{u_\phi}{r}\right), \quad \sigma_{\phi\phi} = 2\eta\left(\frac{1}{r}\frac{\partial u_\phi}{\partial\phi} + \frac{u_r}{r}\right), \tag{3.26}$$

to determine energy dissipation rate due to viscous friction

$$W = \frac{1}{2\eta}\sum_{ij}\sigma_{ij}^2 dV. \tag{3.27}$$

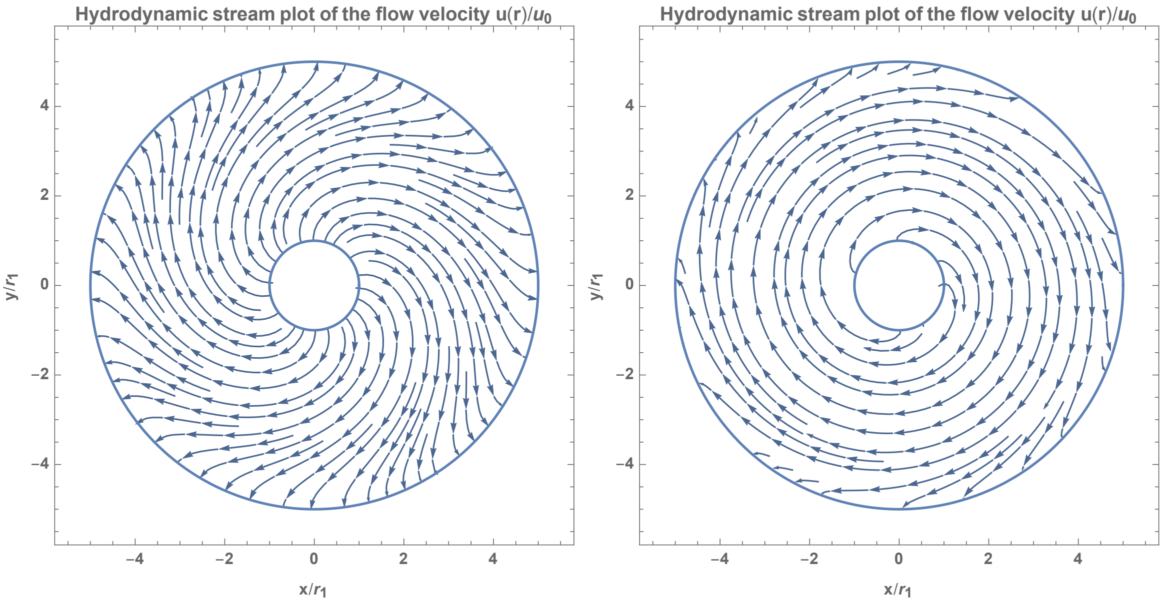


Figure 4: The stream plot of the viscous flow was generated in polar coordinates for $u_x(r) = u_r(r)\cos\phi - u_\phi(r)\sin\phi$ and $u_y(r) = u_r(r)\sin\phi + u_\phi(r)\cos\phi$ based on the solution from Eqs. (3.23) and (3.25). The velocity field was normalized in units of $u_0 = I/(2\pi e n r_1)$ for the aspect ratio $a = r_2/r_1 = 5$. The strength of the external field that controls the distribution of the flow pattern between electrodes in the bulk is characterized by a dimensionless parameter $q = nr_1^2/(2\eta l_B^2)$, where $l_B = \sqrt{1/eB}$ is the magnetic length. This parameter measures the relative strength of the Lorentz and viscous Stokes forces and determines the number of turns the flow makes between the electrodes. On the left panel we took $q = 0.55$ while on the right panel $q = 3.55$ for comparison, the aspect ratio was kept the same in both cases.

The latter gives us resistance $R = W/I^2$. As a result we find

$$R = R_0 + R_B. \tag{3.28}$$

The zero field part of the resistance R_0 comprises of Ohmic and Stokes contributions. The Ohmic part is determined by the momentum-relaxing scattering time in the bulk of the flow and is given by a standard expression

$$R_0^{\mathrm{Ohm}} = \frac{\rho_0}{2\pi}\ln(r_2/r_1). \tag{3.29}$$

This form of the resistance can be readily seen from the Navier-Stokes equation itself by noticing that $\Delta j_r = 0$ yields for the radial component of the electric field $E_r = \rho_0 j_r$, with the corresponding voltage drop $V = \int_{r_1}^{r_2} E_r dr$. This immediately yields Eq. (3.29). The viscous, Stokes contribution to the resistance is given by

$$R_0^{\mathrm{Stokes}} = \frac{\eta}{\pi(ne)^2}\left(\frac{1}{r_1^2} - \frac{1}{r_2^2}\right), \tag{3.30}$$

but its physical origin is much more subtle and to some extent paradoxical as explained in the recent insightful work [87]. To gauge the relative importance of these two terms one should notice that for the large disk, $r_2 \gg r_1$, the viscous term saturates. The Ohmic part, however, grows in this limit very slowly and the ratio between the two is $R_0^{\mathrm{Ohm}}/R_0^{\mathrm{Stokes}} \sim (r_1/l_{\mathrm{G}})^2\ln(r_2/r_1)$, which means that the Ohmic part could in principle dominate, even when the Gurzhi length is large. As explained in Ref. [87] the result for R_0^{Stokes} originates from the voltage drop at the electrodes. In the Ohmic regime, the impact of contact resistance was analyzed in the context of the electronic thermal transport: Lorenz number measurements and Wiedemann-Franz law in particular [89].

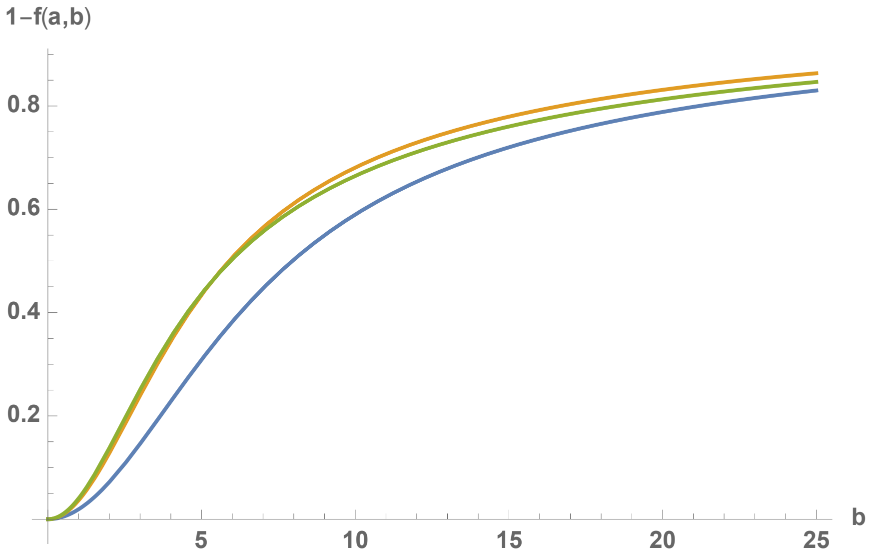


Figure 5: A dimensionless scaling function Eq. (3.32) that describes Stokes-to-Ohm crossover in the magnetoresistance Eq. (3.31) for Corbino device with different choice of aspect ratios $a = 2, 4, 8$.

The field dependent part of the resistance can be presented in the form

$$R_B = \frac{B^2 \ln a}{2\pi\rho_0 (ne)^2}[1 - f(a,b)], \quad a = r_2/r_1, \quad b = r_2/l_{\mathrm{G}}. \tag{3.31}$$

The dimensionless function

$$\begin{aligned} f(a,b) = 1 - \frac{1}{\ln a}\Bigg\{ &\frac{[\mathrm{I}_0(b) - \mathrm{I}_0(b/a)][(a/b)\mathrm{K}_1(b) - (1/b)\mathrm{K}_1(b/a)]}{\mathrm{I}_1(b/a)\mathrm{K}_1(b) - \mathrm{I}_1(b)\mathrm{K}_1(b/a)} \\ &+ \frac{[\mathrm{K}_0(b) - \mathrm{K}_0(b/a)][(a/b)\mathrm{I}_1(b) - (1/b)\mathrm{I}_1(b/a)]}{\mathrm{I}_1(b/a)\mathrm{K}_1(b) - \mathrm{I}_1(b)\mathrm{K}_1(b/a)}\Bigg\} \end{aligned} \tag{3.32}$$

describes the crossover from the Stokes to the Ohmic regime. This function is plotted in Fig. 5 for several different values of the aspect ratio a. Asymptotic limits of this function can be relatively easily extracted. In the Ohmic regime, $b \gg 1$, f is a decaying function of b such that to leading order holds:

$$R_B^{\mathrm{Ohm}} = \frac{B^2 \ln a}{2\pi\rho_0 (ne)^2} \propto B^2 \tau_{\mathrm{MR}}, \quad r_2 \gg l_{\mathrm{G}}. \tag{3.33}$$

In the opposite, viscosity-dominated limit, where $l_{\mathrm{G}} \gg r_{1,2}$, we can expand the Bessel functions at small argument $b \ll 1$ such that

$$R_B^{\mathrm{Stokes}} = \frac{B^2 r_2^2}{16\pi\eta}\left[1 - \frac{1}{a^2}\right]\left[1 - \frac{4a^2\ln^2 a}{(a^2-1)^2}\right] \propto \frac{B^2}{\tau_{\mathrm{MC}}} \quad r_2 \ll l_{\mathrm{G}}. \tag{3.34}$$

This result coincides with the earlier conclusion of Refs. [35, 87] that in the hydrodynamic regime the MR is inversely proportional to the viscosity. This, in principle, enables measurements of the temperature and density dependence of the viscosity from magneto-transport experiments. On the theoretical side it should be possible to extend these results to cover the ballistic-to-hydrodynamic crossover in the magneto-transport, as was recently done for the geometry of narrow channels [90]. It is also of a special interest to consider magneto-thermo-electric phenomena in Corbino geometry, and Nernst effect in particular [91].

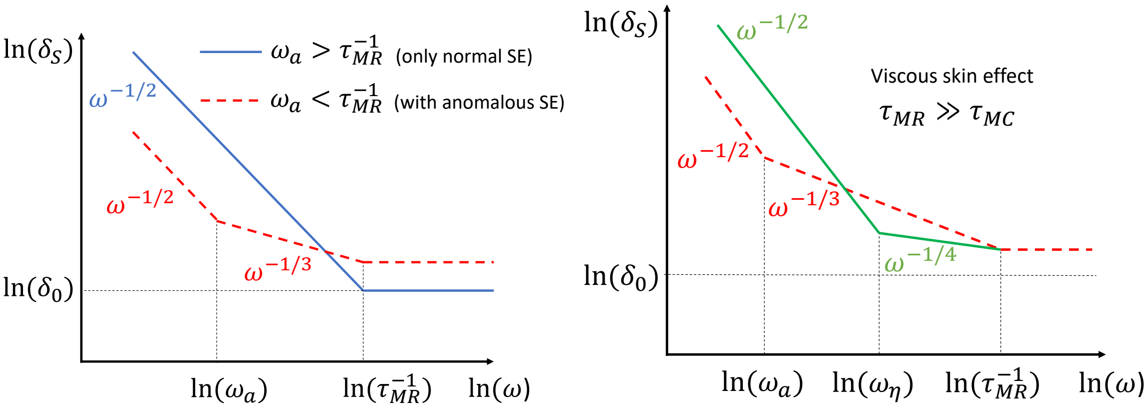


Figure 6: [Left]: Sketch of the frequency dependence of the normal and anomalous skin effect in the log-log scale where ω_a is the crossover frequency between the normal and anomalous skin effect. One only expects the anomalous skin effect in sufficiently clean samples. On the plot $\delta_0 = c/\omega_{\rm pl}$ and $\omega_a = (c/v_F)^2\tau_{\rm MR}^{-1}/(\omega_{\rm pl}\tau_{\rm MR})^2$ where $\omega_{\rm pl}$ is the plasma frequency. [Right]: Sketch of the skin effect including the intermediate regime of viscous skin effect behavior with $\omega_\eta \simeq (\tau_{\rm MR}/\tau_{\rm MC})\omega_a$

3.4. Hydrodynamic surface impedance in a viscous skin effect

In terms of the response to an electromagnetic field, the hydrodynamic regime of an electron-phonon fluid is not limited to (*dc*) transport properties but occupies a finite domain of the frequency-momentum (ω, q) parameter space which is bound by the conditions $\omega\tau_{\rm ep} \ll 1$ and $ql_{\rm ep} \ll 1$. Finite frequency properties of viscous electrons have attracted significant theoretical interest in recent years with interesting predictions ranging from nonlinear electrodynamics [92, 93] (e.g. second-harmonic generation), resonant phenomena [94] (e.g. viscous cyclotron motion) to nonlocal effects in pulsating flows [95, 96]. The optical conductivity and the transmission of electromagnetic waves through thin ultra-pure metals have been considered in Refs. [97, 98] under the condition that hydrodynamic regime is governed by fast electron-electron collisions. Quantum critical hydrodynamics in the *dc* conductivity of graphene at the neutrality point was predicted in Ref. [41] and recently observed experimentally in Ref. [99]. In this section we briefly consider a related problem of the skin-effect (SE) for the strongly coupled electron-phonon liquids. An observables of interest, discussed in the context of electron hydrodynamics already by Gurzhi in Ref. [51], is the frequency-dependent surface impedance [31].

Consider a skin-effect geometry when a monochromatic electromagnetic wave of frequency ω is incident on a metal surface (xy-plane). It is assumed that the metal occupies a semi-infinite volume $z > 0$ with the vacuum on the other side $z < 0$. From a pair of Maxwell equations

$$[\nabla\times\boldsymbol{E}] = -(1/c)\partial_t\boldsymbol{H}, \quad [\nabla\times\boldsymbol{H}] = (4\pi/c)\boldsymbol{j} \tag{3.35}$$

we can establish a self-consistent relation between the electrical field and the induced current in the medium

$$\nabla^2\boldsymbol{E} = (4\pi/c^2)\partial_t\boldsymbol{j}. \tag{3.36}$$

In the linear regime, the current is proportional to the drift velocity of the liquid

$$\boldsymbol{j} = en\boldsymbol{u}, \tag{3.37}$$

which obeys our hydrodynamic equation of motion

$$\partial_t\boldsymbol{u} = \nu\nabla^2\boldsymbol{u} + e\boldsymbol{E}/m - \boldsymbol{u}/\tau_{\rm MR} \tag{3.38}$$

that includes a time-dependent inertia term. By passing to Fourier space in frequency $\boldsymbol{E}(\boldsymbol{r},t) = \mathrm{Re}\left\{\boldsymbol{E}_\omega e^{i\omega t}\right\}$ and eliminating $\boldsymbol{u}(\boldsymbol{r},t)$, one easily obtains a single linear differential equation for the spatial dependence of the field. For the described geometry one finds

$$\partial_z^4E_\omega - l_{\rm G}^{-2}(\omega)\partial_z^2E_\omega + il_\omega^{-4}E_\omega = 0. \tag{3.39}$$

Here we introduced the frequency-dependent Gurzhi length

$$l_{\rm G}(\omega)=l_{\rm G}/\sqrt{1+i\omega\tau_{\rm MR}} \tag{3.40}$$

and also another frequency-dependent length scale

$$l_\omega=\sqrt[4]{v_F\delta_0^2 l_{\rm MC}/\omega},\quad \delta_0^2=mc^2/(4\pi ne^2), \tag{3.41}$$

where δ_0 is the familiar London penetration depth in the clean limit. The surface impedance is defined as the ratio between the electric field on the metal surface and the current density, integrated over the volume

$$Z(\omega)=\frac{E_\omega(0)}{(4\pi/c)\int j(z)dz}=-(i\omega/c)[E_\omega(0)/\partial_z E_\omega(0)]. \tag{3.42}$$

The impedance is a complex function of frequency and its real part determines the energy dissipated by the field. To find Z we look at the characteristic equation of Eq. (3.39), $E_\omega\propto e^{kz}$, whose roots follow as solutions of a bi-quadratic equation

$$k^2_\pm(\omega)=\frac{1}{2}\left[l_{\rm G}^{-2}\pm\sqrt{l_{\rm G}^{-4}-4il_\omega^{-4}}\right]. \tag{3.43}$$

This equation gives four different roots and one needs to select two of them k_1 and k_2 that have negative real part. These solutions correspond to a decaying field into the bulk of the sample. The spatial profile of the field is then given by a linear superposition of two exponentials: $E_\omega(z)=A_0e^{k_1z}+B_0e^{k_2z}$. Two coefficients are determined by the boundary conditions $E_\omega(0)=A_0+B_0$ and $[\partial^3_zE_\omega(z)]_{z=0}=-(1/l_{\rm S})[\partial^2_zE_\omega(z)]_{z=0}$, where $l_{\rm S}$ is again the slip length [78]. The second boundary condition corresponds to the linear relationship between $\boldsymbol{E}$ and $\boldsymbol{u}$ and follows directly from Eqs. (3.2) and (3.36). Solving the linear algebraic equations we find

$$A_0=E_\omega(0)\frac{k_2^2}{k_2^2-\beta k_1^2},\quad B_0=-E_\omega(0)\frac{\beta k_1^2}{k_2^2-\beta k_1^2} \tag{3.44}$$

where $\beta=(1+k_1l_{\rm S})/(1+k_2l_{\rm S})$. For the case of no-slip boundary condition ($\beta\to1$) one can expresses Z in terms of roots $k_{1,2}$ as follows

$$Z(\omega)=-\frac{i\omega}{c}\frac{k_1+k_2}{k_1k_2}. \tag{3.45}$$

In the opposite case of no-stress ($\beta\to k_1/k_2$) surface impedance takes the form

$$Z(\omega)=-\frac{i\omega}{c}\frac{k_1^2+k_1k_2+k_2^2}{k_1k_2(k_1+k_2)}. \tag{3.46}$$

It turns out that both limits exhibit the same frequency dependence (modulo numerical factors of the order of unity). Indeed, there are two special cases of interest that one can analyze. First is the regime when $l_\omega\gg l_{\rm G}$, which implies a bound on the range of frequencies $\omega<\omega_\eta$, where $\omega_\eta=(\tau_{\rm MR}/\tau_{\rm MC})\,\omega_a$ is determined by the frequency $\omega_a\sim\tau^{-1}_{\rm ep}(\delta_0/l_{\rm MR})^2$, where usually the skin effect crosses over to the anomalous skin effect. For $\omega<\omega_\eta$ it is easy to see from Eq. (3.43) that one of the roots is parametrically larger than the other: for example $k_1\gg k_2$, with $k_1\sim l_{\rm G}^{-1}$ and $k_2\sim\delta_{\rm S}^{-1}$. The length scale $\delta_{\rm S}=l^2_\omega/l_{\rm G}=\delta_0/\sqrt{\omega\tau_{\rm MR}}$ emerges, which is nothing else but the usual skin penetration depth, since $E_\omega(z)\propto e^{-(1+i)z/\sqrt{2}\delta_{\rm S}}$. The impedance in this frequency range is identical to the one in the normal skin effect

$$Z(\omega)\approx\frac{\delta_0}{c}\sqrt{\frac{\omega}{\tau_{\rm MR}}}e^{i\pi/4},\quad \omega<\omega_\eta. \tag{3.47}$$

In the opposite, viscous regime $\omega>\omega_\eta$ the Gurzhi length is large compared to l_ω. Now there are two parametrically identical roots $k_1=-ik_2=-l_\omega^{-1}e^{-i\pi/8}$ of Eq. (3.43), and the scale of skin penetration depth is controlled by l_ω only, such that $\delta_{\rm S}\propto1/\sqrt[4]{\omega}$. In this case the impedance is given by

$$Z(\omega)\approx\frac{\delta_0}{c}\sqrt[4]{\frac{\omega^3v_Fl_{\rm MC}}{\delta_0^2}}e^{3i\pi/8},\quad \omega_\eta<\omega, \tag{3.48}$$

which is solely determined by momentum-conserving electron-phonon collisions. This is the result for no-slip boundary conditions. In the opposite limit, with no stress boundary conditions, one obtains a result where $Z(\omega)$ of Eq. (3.48) is multiplied by a factor $i/2$. This gives rise to a measurable phase shift in the impedance. Whether no-slip or no-stress boundary conditions are appropriate depends on frequency. The former is correct for $\omega < \omega_\eta\,(l_{\mathrm{G}}/l_{\mathrm{S}})^4$ while the latter is appropriate in the opposite limit. In the regime where the Gurzhi length l_{G} is larger than the slip length l_{S}, which is clearly fulfilled for diffuse scattering at the interface [78], this frequency-dependent crossover between distinct boundary-scattering effects should be observable and may serve as tool to determine the slip length.

In complete analogy with the Gurzhi effect in the resistance, where the momentum-relaxing scattering rate drops out from the expression for the resistivity, this regime can be termed as *hydrodynamic skin effect* [51]. The upper bound on frequency that determines the regime of the viscous skin effect is set by the usual hydrodynamic condition $l_\omega > l_{\mathrm{MC}}$. It is worth emphasizing that this hydrodynamic limit is conceptually different from the high-frequency anomalous skin effect where $\delta_{\mathrm{S}} \propto 1/\sqrt[3]{\omega}$ and $Z \propto \omega^{2/3}$. Figure 6 summarizes the frequency dependence of the surface skin depth in different regimes.

3.5. Noise thermometry of electron-phonon scattering

Johnson noise thermometry provides fruitful experimental tools to study electronic thermoelectric conductivity in solids. Most recently these methods were applied to study electronic conduction of a monolayer graphene over a wide range of temperatures, charge densities, and magnetic fields [100]. In this section we discuss the role of strong electron-phonon scattering on the noise spectra of current fluctuations in mesoscopic conductors. The question itself is not new and has been discussed by multiple authors employing various approximations and methods of kinetic theory. The comprehensive summary of known results is given in the review article by Blanter and Büttiker [101], see specifically section 6.3.2 page 122. Perhaps the most concise and elegant summary of work that has been done on this topic is presented in the experimental paper of Steinbach *et al.* [102], see specifically their Fig. 1. To place our approach in the context of existing studies we first briefly summarize key results and acknowledge main contributions.

The interest in the problem of current noise in mesoscopic conductors was triggered by works of Beenakker and Büttiker [103] based on scattering matrix formalism, and Nagaev [104] who employed the stochastic Boltzmann-Langevin kinetic equation (see also book of Kogan [105] on electronic noise and fluctuations in solids for an in-depth overview). These authors showed that the celebrated result of Schottky for a Poisson process of the shot noise, namely the zero-frequency current power spectrum of fluctuations, $S = 2eIF$ is suppressed by a Fano factor $F = 1/3$. This is a single-particle effect that can be understood from the Dorokhov statistics of transmission eigenvalues in disordered conductors. In the current literature this regime is called shot noise of cold electrons. The subsequent studies focused on the role of inelastic processes. Frequent electron-electron collisions lead to rapid equilibration. Shot noise survives in this limit and has the same structural form as in the case of noninteracting particles but is described by a different Fano factor $F = \sqrt{3}/4$. This result was demonstrated by Kozub and Rudin [106], and de Jong and Beenakker [107] using a semiclassical approach. These authors assumed that inelastic processes lead to a local equilibrium, described by a Fermi distribution with spatially varying temperature $T(\boldsymbol{r})$ and derived an effective diffusion-like equation for the non-equilibrium (voltage-dependent) profile of $T(\boldsymbol{r})$. This regime is called shot noise of hot electrons. The crossover between the two and extensions to full-counting statistics was developed by Bagrets [108] and Gutman *et al.* [109]. The influence of strong electron-phonon collisions was addressed by Gurevich and Rudin [110], Nagaev [111], and Naveh *et al.* [112]. In the first of these papers the electron-phonon collision integral was treated perturbatively, whereas in the other two papers a two-temperature model of the electron-phonon out-of-equilibrium state was assumed and an equation for the electronic temperature profile derived. Naveh [113] also performed direct numerical calculations of the integral equation with a phenomenological ansatz for the distribution functions.

Unlike the calculation of the electron-phonon drag viscosity, where diffusion in momentum space is important, noise is primarily affected by the energy relaxation. For this reason it will be convenient and technically advantageous to average the distribution function over the Fermi surface such that it will depend on the energy and real-space coordinate

$$n_\varepsilon(\boldsymbol{r}) = \frac{1}{\nu}\int_{\boldsymbol{p}} n_{\boldsymbol{p}}(\boldsymbol{r})\delta(\varepsilon-\varepsilon_{\boldsymbol{p}}), \tag{3.49}$$

where ν is the density of states. With this notation, the collision integral due electron-phonon scattering in Eq. (2.4) can be rewritten as follows

$$\begin{aligned}\mathrm{St}_{\mathrm{ep}}\{n,N\} = &\int_\omega M(\varepsilon,\varepsilon',\omega)[n_{\varepsilon-\omega}(1-n_\varepsilon)N_\omega - n_\varepsilon(1-n_{\varepsilon-\omega})(1+N_\omega)]\\ &+\int_\omega M(\varepsilon,\varepsilon',\omega)[n_{\varepsilon+\omega}(1-n_\varepsilon)(1+N_\omega) - n_\varepsilon(1-n_{\varepsilon+\omega})N_\omega],\end{aligned} \tag{3.50}$$

where the Eliashberg kernel is of the form

$$M(\varepsilon,\varepsilon',\omega) = \frac{1}{\nu}\int_{\boldsymbol{pq}} W(\boldsymbol{p}|\boldsymbol{p'q})\delta(\varepsilon-\varepsilon_{\boldsymbol{p}})\delta(\varepsilon'-\varepsilon_{\boldsymbol{p'}})\delta(\omega-\omega_{\boldsymbol{q}}). \tag{3.51}$$

Its ε,ε' dependence is pinned to energies at the Fermi level, whereas the energy transfer dependence on ω is strong. Apparently, its functional form in the disordered conductors at frequencies below the scale of Debye energy was subject of certain controversy with multiple conflicting results (this is discussed by Belitz [114]). We will discuss a generic model

$$M(\omega) = \lambda_{\mathrm{ep}} k(\omega/\omega_D)^k/2, \quad k>1 \tag{3.52}$$

and show that main results are only weakly dependent on the exponent k. Here we use the same convention for the dimensionless coupling constant of electron-phonon interaction λ_{ep} as introduced below Eq. (2.13). To proceed we regroup terms in the collision integral by separating spontaneous emission contributions, namely pieces independent of the bosonic occupation function, and terms proportional to N_ω. Thus we have

$$\mathrm{St}_{\mathrm{ep}}\{n,N\} = \int_\omega M(\omega)\big\{[n_{\varepsilon+\omega}(1-n_\varepsilon) - n_\varepsilon(1-n_{\varepsilon-\omega})] + N_\omega[n_{\varepsilon+\omega}+n_{\varepsilon-\omega}-2n_\varepsilon]\big\}. \tag{3.53}$$

At this point we apply a Fokker-Planck approximation to this integral operator by expanding fermionic occupation factors over the frequency transfer up to quadratic order

$$n_{\varepsilon\pm\omega} \approx n_\varepsilon \pm \omega\partial_\varepsilon n_\varepsilon + (\omega^2/2)\partial^2_\varepsilon n_\varepsilon. \tag{3.54}$$

Inserting this back into the collision integral we find

$$\mathrm{St}_{\mathrm{ep}}\{n,N\} \approx \mathcal{A}(1-2n_\varepsilon)\partial_\varepsilon n_\varepsilon + \frac{\mathcal{B}}{2}\partial^2_\varepsilon n_\varepsilon, \tag{3.55}$$

where the expansion coefficients are

$$\mathcal{A} = \lambda_{\mathrm{ep}}\int_\omega \omega M(\omega) \simeq a\lambda_{\mathrm{ep}}\omega_D^2, \qquad \mathcal{B} = \lambda_{\mathrm{ep}}\int_\omega \omega^2 M(\omega)(1+N_\omega) \simeq b\lambda_{\mathrm{ep}}\omega_D^3, \tag{3.56}$$

with $a\sim b$ being model-specific numerical pre-factors of order of unity. In this estimation we assumed $T\ll\omega_D$ so that $N_\omega \ll 1$ and cut off the integral at the Debye energy. In general, $\mathcal{B}(\boldsymbol{r})$ is weakly coordinate dependent which is ignored in the analysis below. The advantage of the Fokker-Planck approximation is threefold: (i) it is not limited to near-equilibrium problems; (ii) it reduces the collision term to a local differential form; (iii) it preserves the nonlinearity of the collision operator. Curiously, the nonlinearity of Eq. (3.55) is of the Burgers type and known in the theory of nonlinear waves [115, 116].

Consider a quasi-1D geometry of a wire of length L subject to the voltage bias V. In the diffusive approximation, the distribution function obeys the following kinetic equation (see Eq. 221 in Ref. [101])

$$D\nabla^2 n_\varepsilon(x) + \mathrm{St}_{\mathrm{ep}}\{n\} = 0 \tag{3.57}$$

with the collision term taken from Eq. (3.55). Provided that $n_\varepsilon(x)$ is known the general semiclassical expression for the shot noise power of current fluctuations can be expressed in terms of a non-equilibrium steady-state distribution function as follows:

$$S = \frac{4}{RL}\int_{-L/2}^{L/2} dx \int_{-\infty}^{+\infty} n_\varepsilon(x)[1-n_\varepsilon(x)]d\varepsilon, \tag{3.58}$$

where R is the wire resistance. It will be useful to rescale the coordinate $l = x/L$ and energy $\epsilon = \varepsilon/\omega_D$, and introduce the Thouless energy $E_{\mathrm{Th}} = D/L^2$. In these dimensionless variables it follows

$$\frac{\partial^2 n}{\partial l^2} + \lambda_{\mathrm{ep}}\frac{\omega_D}{E_{\mathrm{Th}}}\left[a(1-2n)\frac{\partial n}{\partial \epsilon} + \frac{b}{2}\frac{\partial^2 n}{\partial \epsilon^2}\right] = 0. \tag{3.59}$$

This non-linear partial differential equation is of the Burgers type [115, 116], which is a prototypical equation to develop discontinuities such as shock waves. Recall that the Fokker-Planck approximation implies strong local equilibration, thus in the current context this means a short relaxation length scale as compared to the wire length $L \gg l_{\mathrm{ep}}$. This practically corresponds to an infinite wire limit. Exploring an analogy to nonlinear waves we can attempt searching for a solution in the form of a “propagating soliton” $n_\epsilon(l) \to n(\epsilon - ul)$, where the speed is governed by the voltage, namely $u = eV/\omega_D$. This is also physically justified; we simply assume that the energy dependence is governed by the local electrochemical potential. The result reads

$$n_\varepsilon(x) = \left[\exp\left(\beta(\varepsilon - eVx/L)/\omega_D\right) + 1\right]^{-1}, \quad \beta^{-1} = \frac{(eV/\omega_D)^2 E_{\mathrm{Th}}}{a\lambda_{\mathrm{ep}}\omega_D} + b/2a, \tag{3.60}$$

and corresponds to a highly non-thermal state with voltage dependent temperature. This is also the point where, perhaps, the Fokker-Planck approach overlaps with previous approximation, in particular a model with a coordinate and voltage dependent electronic temperature. From Eq. (3.58) it then follows that the current noise in this regime is described by the voltage-dependent Fano factor

$$S = 2eIF, \quad F \simeq eVE_{\mathrm{Th}}/\lambda_{\mathrm{ep}}\omega_D^2. \tag{3.61}$$

The Fano factor drops as $F \propto 1/L^2$ in this regime that corresponds to a suppression of shot noise by inelastic processes. This is in qualitative agreement with Fig. 1 of Ref. [102] in the long L asymptote. It is also in a qualitative agreement with other previous conclusions [112, 113] albeit obtained under different approximations.

4. Summary and outlook

In this work we have considered several examples of hydrodynamic effects that can occur in electron liquids under the condition of strong phonon drag. Electrons and phonons form a combined fluid with an emergent joint drift velocity as hydrodynamic variable. The effect is caused by the fact that the relaxation of the total momentum $\boldsymbol{P}_{\mathrm{el}} + \boldsymbol{P}_{\mathrm{ph}}$ is much slower than the momenta $\boldsymbol{P}_{\mathrm{el}}$ or $\boldsymbol{P}_{\mathrm{ph}}$ of electrons or phonons alone. This is guaranteed for clean samples with weak or kinematically forbidden umklapp scattering processes. We have studied coupled kinetic equations for electrons and phonons, and inferred the effective viscosity of this strongly-coupled transport regime as well as its thermal conductivity. The situation happens to be analogous to the viscous flows in the regime of electron-electron dominated collisions with the only difference that momentum-conserving mean free path has a different temperature dependence. This difference propagate to numerous observables such as the viscous resistance, the Hall resistance, or the surface impedance.

While our work was primarily motivated by recent experiments, the delafossite metals $\mathrm{PdCoO_2}$ and $\mathrm{PtCoO_2}$ studied in Refs. [16–18] and $\mathrm{PtSn_4}$ of Refs. [20, 21] in particular, we have not yet tried to tailor this analysis to the case of a multi-band conductors or systems with complex Fermi surfaces. Hydrodynamic transport theory of electron-phonon liquids in 3D Weyl or Dirac semimetals is yet to be fully developed. The first required step towards this direction would be to consider a minimal two-band model of a non-compensated metal. The generalized kinetic scheme has to be developed then for a coupled kinetic equations for electron, holes, and phonons. Another interesting possibility is to consider the possibility of a hydrodynamic regime in Luttinger semimetals [117, 118] with the inclusion of electron-phonon scattering. In addition, in these systems an electron-hole imbalance mode is not restricted so severely like in graphene so that an unusual transport regime is possible. To the best our knowledge, electron-phonon drag of imbalanced liquids has not been addressed in the previous studies.

5. Acknowledgments

We thank Graham Baker, Kamran Behnia, Douglas Bonn, Jennifer Coulter, Cory Dean, Igor Gornyi, Philip Kim, Egor Kiselev, Leonid Levitov, Julia Link, Andrew Lucas, Andrew Mackenzie, Dmitrii Maslov, Alexander Mirlin, Roderich Moessner, Nabhanila Nandi, Boris Narozhny, Thomas Scaffidi, Davide Valentinis, and especially Boris Shklovskii for valuable comments and discussions. This work was supported by the National Science Foundation Grant No. DMR-1653661, the Binational Science Foundation Grant No. 2016317, and the European Commission's Horizon 2020 RISE program Hydrotronics (Grant No. 873028). This work was performed in part at the Max Planck Institute for the Physics of Complex Systems, and the Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1607611.

A. Appendix

A.1. Variational solution of the Boltzmann equation for Bloch-Grüneisen conductivity

In this section we present a method to solve the linearized Boltzmann equation, which is based on the variational formulation of the kinetic theory. We begin from Eq. (2.14) and rewrite it by combining both terms on the right-hand-side together, which gives

$$\boldsymbol{v}_{\boldsymbol{p}}\frac{\partial f}{\partial \varepsilon_{\boldsymbol{p}}}=\frac{1}{T}\int_{\boldsymbol{p}'}D(\boldsymbol{p},\boldsymbol{p}')[\boldsymbol{v}_{\boldsymbol{p}'}g_{\boldsymbol{p}'}-\boldsymbol{v}_{\boldsymbol{p}}g_{\boldsymbol{p}}]\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'}), \tag{A.1}$$

where we took a parametrization of the form

$$\psi_{\boldsymbol{p}}=e\boldsymbol{v}_{\boldsymbol{p}}\boldsymbol{E}\frac{g(\varepsilon_{\boldsymbol{p}})}{T}, \tag{A.2}$$

and, after few algebraic steps, reorganized kernels $K_\pm$ to make the result manifestly symmetric with respect to interchange of momenta. This way we arrived at

$$D(\boldsymbol{p},\boldsymbol{p}')=D_0|\boldsymbol{p}-\boldsymbol{p}'|\frac{f(\varepsilon_{\boldsymbol{p}})f(\varepsilon_{\boldsymbol{p}'})}{|e^{-\varepsilon_{\boldsymbol{p}'}/T}-e^{-\varepsilon_{\boldsymbol{p}}/T}|}. \tag{A.3}$$

It is easy to see that Eq. (A.1) can be obtained from the variational derivative of the following auxiliary functional

$$Q_P[g]=\frac{1}{4T}\int_{\boldsymbol{p}\boldsymbol{p}'}D(\boldsymbol{p},\boldsymbol{p}')[\boldsymbol{v}_{\boldsymbol{p}'}g_{\boldsymbol{p}'}-\boldsymbol{v}_{\boldsymbol{p}}g_{\boldsymbol{p}}]^2\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'})-\int_{\boldsymbol{p}}\boldsymbol{v}^2_{\boldsymbol{p}}g_{\boldsymbol{p}}\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}. \tag{A.4}$$

Thus solving Eq. (A.1) is fully equivalent to minimizing Eq. (A.4). Of course, this is not an easy task either. However, one can try to postulate a variational ansatz for $g_{\boldsymbol{p}}$ and then extremize the functional, which is often a simpler computation. To this end, suppose that $g_{\boldsymbol{p}}=g$ is a constant, which is the leading contribution for temperatures small compared to the Fermi energy, we have

$$Q_P[g]=\frac{1}{2}A_Pg^2-B_Pg, \tag{A.5}$$

where

$$A_P=\frac{1}{2T}\int_{\boldsymbol{p}\boldsymbol{p}'}D(\boldsymbol{p},\boldsymbol{p}')[\boldsymbol{v}_{\boldsymbol{p}'}-\boldsymbol{v}_{\boldsymbol{p}}]^2\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'}),\quad B_P=\int_{\boldsymbol{p}}\boldsymbol{v}^2_{\boldsymbol{p}}\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}=v^2_F\nu. \tag{A.6}$$

Here in the integral for B_P we introduced density of states ν at the Fermi energy. The extremal Q_P is determined by $g = B_P/A_P$. This allows us to determine the conductivity as

$$\sigma_{\alpha\beta} = 2e^2 \frac{B_P}{A_P} \int_{\boldsymbol{p}} \boldsymbol{v}_{\boldsymbol{p}\alpha} \boldsymbol{v}_{\boldsymbol{p}\beta} \left(-\frac{\partial f}{\partial \varepsilon_{\boldsymbol{p}}}\right) = \frac{2e^2}{3} \frac{(v_F^2 \nu)^2}{A_P} \delta_{\alpha\beta} = \sigma_B \delta_{\alpha\beta}, \tag{A.7}$$

thus finding temperature dependence of the Bloch-Grüneisen conductivity $\sigma_B(T)$ is reduced to the computation of the $A_P(T)$. For the latter we have

$$A_P(T) = \frac{v_F^2}{sT} \int d\varepsilon d\varepsilon' d\omega F_P(\varepsilon, \varepsilon', \omega) \frac{\omega f(\varepsilon) f(\varepsilon')}{|e^{-\varepsilon/T} - e^{-\varepsilon'/T}|} \sum_{\sigma=\pm} \delta(\varepsilon' - \varepsilon + \sigma\omega), \tag{A.8}$$

$$F_P(\varepsilon, \varepsilon', \omega) = \frac{D_0}{v_F^2} \int_{\boldsymbol{p}\boldsymbol{p}'} (\boldsymbol{v}_{\boldsymbol{p}} - \boldsymbol{v}_{\boldsymbol{p}'})^2 \delta(\omega - \omega_{\boldsymbol{p}-\boldsymbol{p}'}) \delta(\varepsilon - \varepsilon_{\boldsymbol{p}}) \delta(\varepsilon' - \varepsilon_{\boldsymbol{p}'}). \tag{A.9}$$

Since electronic momenta are close to Fermi momentum, and the phonon momentum is small, the following approximations apply: $(\boldsymbol{v}_{\boldsymbol{p}} - \boldsymbol{v}_{\boldsymbol{p}'})^2 \approx 2v_F^2(1 - \cos\theta_{\boldsymbol{p}\boldsymbol{p}'})$ and $\omega_{\boldsymbol{p}-\boldsymbol{p}'} \approx \sqrt{2} s p_F \sqrt{1 - \cos\theta_{\boldsymbol{p}\boldsymbol{p}'}}$. This implies that to the leading order $F(\varepsilon, \varepsilon', \omega)$ is independent of $\varepsilon, \varepsilon'$ so that

$$F_P(\varepsilon, \varepsilon', \omega) \approx D_0 \nu^2 \int_0^\pi d\theta \sin\theta (1 - \cos\theta) \delta(\omega - \sqrt{2} s p_F \sqrt{1 - \cos\theta}) = \frac{D_0 \nu^2}{2 s p_F} \Theta(2 s p_F - \omega) \left(\frac{\omega}{s p_F}\right)^3, \tag{A.10}$$

where $\Theta(x)$ is the Heaviside step function. Next, we observe that under the approximation that F_P only depends on ω, the energy integrations in $A_P(T)$ can be performed in the closed form. Indeed, it can be readily verified that

$$\sum_{\sigma=\pm} \int d\varepsilon d\varepsilon' \frac{f(\varepsilon) f(\varepsilon')}{|e^{-\varepsilon/T} - e^{-\varepsilon'/T}|} \delta(\varepsilon' - \varepsilon + \sigma\omega) = \frac{\omega}{\cosh(\omega/T) - 1}. \tag{A.11}$$

Finally, combining everything together as a result we obtain with $\omega_D \approx 2 s p_F$

$$A_P(T) = 16 D_0 \nu^2 v_F^2 p_F G\left(\frac{T}{\omega_D}\right), \quad G(t) = t^5 \int_0^{t^{-1}} \frac{x^5 dx}{\cosh x - 1}. \tag{A.12}$$

As it was done in the main text, we can define the electron-phonon scattering time of momentum relaxation τ_1 via Bloch-Grüneisen conductivity $\sigma_B = e^2 n \tau_1 / m$ with

$$\tau_1^{-1} = 2\omega_D \lambda_{\mathrm{ep}} G(t) = \begin{cases} 480\zeta(5) \lambda_{\mathrm{ep}} \frac{T^5}{\omega_{\mathrm{D}}^4} & t \ll 1 \\ \lambda_{\mathrm{ep}} T & t \gg 1 \end{cases} \tag{A.13}$$

This is the well-known Bloch-Grüneisen behavior. As we saw, the implicit assumption of the analysis is that the phonons remain in equilibrium such that the total momentum conservation is violated.

A.2. Detailed calculation of the electron-phonon drag viscosity

In the phonon-drag regime, where the total momentum conservation is respected, the conductivity is infinite (without umklapp and impurity scattering), yet the joint electron-phonon fluid has a common flow viscosity. To this end, we analyze the problem for a finite shear flow with velocity gradient such that

$$T_{xy} = \eta \frac{\partial u_x}{\partial y}. \tag{A.14}$$

By starting out from the linearized coupled Boltzmann equations

$$-\frac{\partial f}{\partial \varepsilon_{\boldsymbol{p}}} v_y p_x \frac{\partial u_x}{\partial y} = \delta\, \mathrm{St}_{\mathrm{ep}}\{\psi, \phi\}, \quad -\frac{\partial b}{\partial \omega_{\boldsymbol{q}}} s_y q_x \frac{\partial u_x}{\partial y} = \delta\, \mathrm{St}_{\mathrm{pe}}\{\psi, \phi\}, \tag{A.15}$$

we first solve for the phonon distribution

$$\phi_{\boldsymbol{q}}=\frac{1}{\gamma_{\boldsymbol{q}}}\frac{\partial b}{\partial\omega_{\boldsymbol{q}}}s_yq_x\frac{\partial u_x}{\partial y}-\frac{1}{2\gamma_{\boldsymbol{q}}}\int_{\boldsymbol{pp}'}D(\boldsymbol{p},\boldsymbol{p}')(\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}})\sum_{\sigma=\pm}\sigma\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{q}})\delta_{\boldsymbol{p}'-\boldsymbol{p}+\sigma\boldsymbol{q}} \tag{A.16}$$

where

$$\gamma_{\boldsymbol{q}}=\frac{1}{2}\sum_{\sigma=\pm}\int_{\boldsymbol{pp}'}D(\boldsymbol{p},\boldsymbol{p}')\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{q}})\delta_{\boldsymbol{p}'-\boldsymbol{q}+\sigma\boldsymbol{q}}. \tag{A.17}$$

It holds that $\phi_{\boldsymbol{q}}=\phi_{-\boldsymbol{q}}$. We can now insert this solution into the expression for the electronic collision operator and obtain the effective purely electronic Boltzmann equation

$$\begin{aligned}R_{\boldsymbol{p}}\frac{\partial u_x}{\partial y}&=\int_{\boldsymbol{pp}'}D(\boldsymbol{p},\boldsymbol{p}')(\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}})\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'})\\&-\int_{\boldsymbol{kk}'\boldsymbol{p}'}\frac{D(\boldsymbol{p},\boldsymbol{p}')D(\boldsymbol{k},\boldsymbol{k}')}{2\gamma_{\boldsymbol{p}-\boldsymbol{p}'}}(\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}})\sum_{\sigma\sigma'=\pm}\sigma\sigma'\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'})\delta(\varepsilon_{\boldsymbol{k}'}-\varepsilon_{\boldsymbol{k}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'})\delta_{\boldsymbol{p}'-\boldsymbol{p}+\boldsymbol{k}-\boldsymbol{k}'}.\end{aligned} \tag{A.18}$$

It contains now the renormalized source term

$$R_{\boldsymbol{p}}=-\left(\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}v_yp_x+\int_{\boldsymbol{p}'}\frac{D(\boldsymbol{p},\boldsymbol{p}')}{\gamma_{\boldsymbol{p}-\boldsymbol{p}'}}\frac{\partial b(\omega_{\boldsymbol{p}-\boldsymbol{p}'})}{\partial\omega_{\boldsymbol{p}-\boldsymbol{p}'}}s_{\boldsymbol{p}-\boldsymbol{p}',y}(p_x-p'_x)\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'})\right) \tag{A.19}$$

and the collision term captured by the second contribution on the right-hand-side of Eq. (A.18). Let us estimate the second (integral) term of $R_{\boldsymbol{p}}$ that we denote in the following as $\delta R_{\boldsymbol{p}}$. First we notice that with the help of Eq. (A.11) $\gamma_{\boldsymbol{q}}$ defined in Eq. (A.17) can be reduced to the following form

$$\gamma_{\boldsymbol{q}}=\left(\frac{D_0\nu}{4v_F}\right)\Theta(2k_F-q)\frac{\omega_{\boldsymbol{q}}}{\cosh(\omega_{\boldsymbol{q}}/T)-1}. \tag{A.20}$$

Next we notice that due to kinematics $\omega_{\boldsymbol{k}-\boldsymbol{k}'}=\sqrt{2}sp_F\sqrt{1-\cos\theta_{\boldsymbol{kk}'}}\approx sp_F\theta_{\boldsymbol{kk}'}$ so that

$$\begin{aligned}\delta R_{\boldsymbol{p}}=-4vp_F\int d\theta\sin\theta[\sin\varphi-\sin(\theta+\varphi)][\cos\varphi-\cos(\theta+\varphi)]\\\times[\cosh(\omega_\theta/T)-1]\frac{\partial b(\omega_\theta)}{\partial\omega_\theta}\sum_{\sigma=\pm}\frac{f(\varepsilon_{\boldsymbol{p}})f(\varepsilon_{\boldsymbol{p}}-\sigma\omega_\theta)}{|e^{-\varepsilon_{\boldsymbol{p}}/T}-e^{-(\varepsilon_{\boldsymbol{p}}-\sigma\omega_\theta)/T}|},\end{aligned} \tag{A.21}$$

where we took $\boldsymbol{p}=p_F(\cos\varphi,\sin\varphi)$ and $\boldsymbol{p}'=p_F(\cos(\varphi+\theta),\sin(\varphi+\theta))$. The integral is dominated by the small angle of scattering $\theta=\omega/sp_F\ll1$, so that recalling that $e^{\varepsilon/T}f(\varepsilon)=1-f(\varepsilon)$, summing over $\sigma=\pm$, using Eq. (2.9), and expanding over ω to leading order we get

$$\delta R_{\boldsymbol{p}}\approx4vp_F\sin\varphi\cos\varphi\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}\int_0^\infty\frac{\omega^3d\omega}{(sp_F)^4}[\cosh(\omega/T)-1]\frac{\partial b_\omega}{\partial\omega}\frac{\omega[1+\coth(\omega/2T)]}{2\sinh(\omega/2T)}e^{-\omega/2T} \tag{A.22}$$

which yields

$$\delta R_{\boldsymbol{p}}=-v_xp_y\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}\frac{16\pi^4}{15}\left(\frac{T}{sp_F}\right)^4. \tag{A.23}$$

It is clear that at low temperatures we can ignore the second term in $R_{\boldsymbol{p}}$ compared to the first one. The primary mechanism by which the flow gradient couples to the electron-phonon fluid is by directly affecting its electron component. By the same token one can estimate the renormalization piece of the collision integral, namely the second integral term on the right-hand-side of Eq. (A.18). It happens to be smaller than the first term and can be also dropped. In the end, we arrive at the much simplified Boltzmann equation

$$-\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}v_yp_x\frac{\partial u_x}{\partial y}=\int_{\boldsymbol{p}'}D(\boldsymbol{p},\boldsymbol{p}')(\psi_{\boldsymbol{p}'}-\psi_{\boldsymbol{p}})\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'}), \tag{A.24}$$

which is essentially the Boltzmann equation without taking into account that the phonons are not equilibrated. Hence, momentum conservation, while important for the hydrodynamic interpretation of the viscosity is not important for its actual value. To proceed with the solution of Eq. (A.24) we can follow an analysis that is essentially the same as the one we used to determine the resistivity within the Bloch-Grüneisen limit. We can in fact perform this analysis for a distribution function $\propto \cos(l\theta)$, where l is the angular momentum. The resistivity corresponds to $l=1$ while the viscosity to $l=2$. This yields the scattering rate for arbitrary l, and for viscosity in particular

$$\tau_2^{-1} = 6\lambda_{\rm ep}\frac{T^5}{\omega_D^4}\int_0^{\omega_D/T}\frac{x^2\left(1-\frac{T^2}{2\omega_D^2}x^2\right)}{\cosh(x)-1}dx \tag{A.25}$$

The asymptotic behavior in the low-temperature regime gives Eq. (2.50) in the main text.

A.3. Detailed calculation of the electron-phonon drag thermal conductivity

For the sake of thermal conductivity calculation we can make the following ansatz for the non-equilibrium distribution function of electrons

$$\delta n_1 = -T\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}\psi_{\boldsymbol{p}}, \quad \psi_{\boldsymbol{p}} = (\boldsymbol{v}\nabla_{\boldsymbol{r}}T)\frac{\varepsilon_{\boldsymbol{p}}}{T}\frac{g(\varepsilon_{\boldsymbol{p}})}{T}. \tag{A.26}$$

Then, the Boltzmann equation for $g_{\boldsymbol{p}}$ be obtained from the variational analysis of the functional

$$Q_E[g] = \frac{1}{4T}\int_{\boldsymbol{p}\boldsymbol{p}'} D(\boldsymbol{p},\boldsymbol{p}')\left[\boldsymbol{v}_{\boldsymbol{p}'}g_{\boldsymbol{p}'}\frac{\varepsilon_{\boldsymbol{p}'}}{T} - \boldsymbol{v}_{\boldsymbol{p}}g_{\boldsymbol{p}}\frac{\varepsilon_{\boldsymbol{p}}}{T}\right]^2\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'}) - \int_{\boldsymbol{p}}\boldsymbol{v}^2_{\boldsymbol{p}}g_{\boldsymbol{p}}\frac{\varepsilon^2_{\boldsymbol{p}}}{T^2}\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}. \tag{A.27}$$

The analysis here parallels with that of Bloch-Grüneisen calculations with the only difference that we have now some extra energy factors as we seek the results for the thermal current in response to applied temperature gradient. At temperatures small compared to the Fermi energy we can assume that $g(\varepsilon_{\boldsymbol{p}}) = g$ is a constant and we obtain

$$Q_E[g] = \frac{1}{2}A_E g^2 - B_E g, \tag{A.28}$$

where

$$A_E = \frac{1}{2T^3}\int_{\boldsymbol{p}\boldsymbol{p}'} D(\boldsymbol{p},\boldsymbol{p}')[\boldsymbol{v}_{\boldsymbol{p}'}\varepsilon_{\boldsymbol{p}'} - \boldsymbol{v}_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}]^2\sum_{\sigma=\pm}\delta(\varepsilon_{\boldsymbol{p}'}-\varepsilon_{\boldsymbol{p}}+\sigma\omega_{\boldsymbol{p}-\boldsymbol{p}'}), \quad B_E = \frac{1}{T^2}\int_{\boldsymbol{p}}\boldsymbol{v}^2_{\boldsymbol{p}}\varepsilon^2_{\boldsymbol{p}}\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}} = v_F^2 T c_{\rm el}(T). \tag{A.29}$$

In the analysis of the coefficient $A_E(T)$ we can introduce the corresponding function $F_E(\varepsilon,\varepsilon',\omega)$:

$$A_E(T) = \frac{v_F^2}{sT}\int d\varepsilon d\varepsilon' d\omega F_E(\varepsilon,\varepsilon',\omega)\frac{\omega f(\varepsilon)f(\varepsilon')}{|e^{-\varepsilon/T}-e^{-\varepsilon'/T}|}\sum_{\sigma=\pm}\delta(\varepsilon'-\varepsilon+\sigma\omega), \tag{A.30}$$

$$F_E(\varepsilon,\varepsilon',\omega) = \frac{D_0}{T^2v_F^2}\int_{\boldsymbol{p}\boldsymbol{p}'}[\boldsymbol{v}_{\boldsymbol{p}'}\varepsilon_{\boldsymbol{p}'} - \boldsymbol{v}_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}]^2\delta(\omega-\omega_{\boldsymbol{p}-\boldsymbol{p}'})\delta(\varepsilon-\varepsilon_{\boldsymbol{p}})\delta(\varepsilon'-\varepsilon_{\boldsymbol{p}'}). \tag{A.31}$$

Next we notice that

$$[\boldsymbol{v}_{\boldsymbol{p}'}\varepsilon_{\boldsymbol{p}'} - \boldsymbol{v}_{\boldsymbol{p}}\varepsilon_{\boldsymbol{p}}]^2 \approx v_F^2(\varepsilon-\varepsilon')^2 - 2v_F^2\varepsilon\varepsilon'(1-\cos\theta_{\boldsymbol{p}\boldsymbol{p}'}), \tag{A.32}$$

where the second term contains the usual transport scattering cross-section factor $(1-\cos\theta_{\boldsymbol{p}\boldsymbol{p}'})$, however unlike in the case of conductivity, here it gives only a subleading correction for the energy relaxation, and can be neglected. As a result one finds

$$F_E(\varepsilon,\varepsilon',\omega) \approx \frac{4\lambda_{\rm ep}s}{\omega_D}\left(\frac{\varepsilon-\varepsilon'}{T}\right)^2\left(\frac{\omega}{\omega_D}\right). \tag{A.33}$$

Owing to the energy conserving delta-function in A_E one can replace $(\varepsilon-\varepsilon')^2 \to \omega^2$ in the final integrations. Finally, calculating the energy current from Eq. (2.52)

$$\boldsymbol{j}_\varepsilon = -\int_{\boldsymbol{p}} \boldsymbol{v}_{\boldsymbol{p}}(\boldsymbol{v}_{\boldsymbol{p}}\nabla_{\boldsymbol{r}}T)\left(\frac{\varepsilon_{\boldsymbol{p}}}{T}\right)^2 f_{\boldsymbol{p}}(1-f_{\boldsymbol{p}})g_{\boldsymbol{p}} = -\kappa_{\mathrm{ep}}\nabla T, \tag{A.34}$$

with $g_{\boldsymbol{p}} = B_E/A_E$, we determine that the time scale τ_E, that defines thermal conductivity κ_{ep} in Eq. (2.54), is given by

$$\tau_E^{-1} = \frac{2\lambda_{\mathrm{ep}}}{\omega_D^2 T^3}\int_0^{\omega_D}\frac{\omega^5 d\omega}{\cosh(\omega/T)-1}. \tag{A.35}$$

A.4. Spectral properties of the collision integral and super-diffusion on a Fermi surface

In the context of electron liquids when the electron-electron interaction establishes a hydrodynamic regime it is known that there is a fundamental difference between the relaxation of even and odd modes of the distribution function which is specific to the two-dimensional case. As first shown by Gurzhi and coauthors [61, 62] the ratio of corresponding decay rates is $\gamma_{\mathrm{odd}}/\gamma_{\mathrm{even}} \sim (T/E_F)^2 \ll 1$ and physically comes from the kinematics of head-on collisions. This problem was recently re-analyzed in the work by Ledwith *et al.* [63] where special attention was paid to the dependence of these rates on the angular momentum. It was found that $(\gamma_l)_{\mathrm{even}} \sim (T^2/E_F)\ln l$ whereas $(\gamma_l)_{\mathrm{odd}} \sim (T^4/E_F^3)l^4\ln l$ for $1 < l < l_{\max} \sim \sqrt{E_F/T}$. In the context of graphene with electron-electron Coulomb interaction, it was recently shown that the corresponding rate behaves as $\gamma_l \propto (e^2/v_{\mathrm{F}})^2 T\,|l|$. The non-analytic dependence with respect to the angular mode l gives rise to super-diffusion on the Dirac cone and Lévy-flight behavior in phase space, described by a Fokker-Planck equation in phase space with a fractional Laplacian [64].

It is perhaps surprising, but to the best of our knowledge, a similar analysis has not been carried out for electron-phonon liquids. We are aware of two related studies. In the work by Kabanov and Alexandrov [46] the lowest eigenmode of the electron-phonon collision operator corresponding to the energy relaxation was found. This result was obtained by a Fourier transform of the linearized Boltzmann equation that thus can be reduced to an auxiliary problem to an effective Schrödinger equation in the Pöschl-Teller potential. In the work by Gurevich and Laikhtman [54] energy and momentum transport in fluids was analyzed in the regime dominated by phonon-phonon collisions. It was shown that at low enough temperatures the relaxation is primarily governed by near-collinear scattering between acoustic phonons. Globally, however, the relaxation is hierarchical. These collisions first thermalize unidirectional modes on fast scale leading to angle-dependent temperature, which is followed by a slower relaxation process of angular diffusion on a 2D sphere in 3D momentum space. Below we present general results for the electron-phonon collisions applicable for any angular harmonic of non-equilibrium distributions and carry out the analysis for the 3D case where we reveal the super-diffusive character of the relaxation.

We aim to solve the linearized Boltzmann equation

$$\left(\frac{\partial}{\partial t}+\boldsymbol{v}\nabla_{\boldsymbol{r}}\right)\left(-T\frac{\partial f}{\partial\varepsilon_{\boldsymbol{p}}}\right)\psi_{\boldsymbol{p}}(\boldsymbol{r},t) = \delta\,\mathrm{St}_{\mathrm{ep}}\{\psi\} + S_{\boldsymbol{p}} \tag{A.36}$$

with the source term $S_{\boldsymbol{p}}$, by expanding the non-equilibrium distribution function into angular momentum eigenmodes of the spherical harmonics

$$\psi_{\boldsymbol{p}} = \sum_{lm} Y_{lm}(\theta_{\boldsymbol{p}},\varphi_{\boldsymbol{p}})\phi_{lm}(\varepsilon_{\boldsymbol{p}},\boldsymbol{r},t). \tag{A.37}$$

In the limit of degenerate fermions we can ignore the $|\boldsymbol{p}|$ dependence of ϕ_{lm}. Then we multiply the Boltzmann equation with the mode expansion by $Y^*_{l'm'}(\theta_{\boldsymbol{p}},\varphi_{\boldsymbol{p}})$ and integrate over momenta with the usual prescription $\int_{\boldsymbol{p}} \to \frac{\nu}{4\pi}\int d\varepsilon_{\boldsymbol{p}}\int d\Omega_{\boldsymbol{p}}$ where the solid angle measure is $d\Omega_{\boldsymbol{p}} = \sin\theta_{\boldsymbol{p}}d\theta_{\boldsymbol{p}}d\varphi_{\boldsymbol{p}}$. Then it follows after the spacetime Fourier transform

$$(-i\omega+\tau_l^{-1})\phi_{lm}\delta_{ll'}\delta_{m,m'} + iv_F q(a_{lm}\delta_{l',l+1}+b_{lm}\delta_{l',l-1})\phi_{lm}\delta_{m,m'} = S_{lm} \tag{A.38}$$

where

$$\tau_l^{-1} = \int_{\boldsymbol{p}} Y_{lm}(\theta_{\boldsymbol{p}}, \varphi_{\boldsymbol{p}}) \delta \, \mathrm{St}\{Y_{lm}\}, \quad S_{lm} = \int_{\boldsymbol{p}} Y_{lm}(\theta_{\boldsymbol{p}}, \varphi_{\boldsymbol{p}}) S_{\boldsymbol{p}}. \tag{A.39}$$

Here we used that τ_l^{-1} should not depend on m if the system is rotation invariant. The coefficients are $a_{lm} = \sqrt{\frac{(l+1-m)(l+1+m)}{4l(l+2)+3}}$ and $b_{lm} = \sqrt{\frac{(l-m)(l+m)}{4l^2-1}}$. By using the explicit form of the collision integral, the decay rates (inverse relaxation times) for the given angular harmonic can be presented as follows

$$\gamma_l = \tau_l^{-1} = \frac{1}{2(2l+1)} \sum_m \int_{\boldsymbol{p}\boldsymbol{p}'} D(\boldsymbol{p}, \boldsymbol{p}') [Y_{lm}(\theta_{\boldsymbol{p}}, \varphi_{\boldsymbol{p}}) - Y_{lm}(\theta_{\boldsymbol{p}'}, \varphi_{\boldsymbol{p}'})]^2 \sum_{\sigma=\pm} \delta(\varepsilon_{\boldsymbol{p}} - \varepsilon_{\boldsymbol{p}'} + \sigma \omega_{\boldsymbol{p}-\boldsymbol{p}'}). \tag{A.40}$$

The summation over the azimuthal components of the angular momentum can be completed explicitly with the help of the well known formula from the theory of spherical functions

$$\frac{1}{2l+1} \sum_m Y_{lm}(\theta_{\boldsymbol{p}}, \varphi_{\boldsymbol{p}}) Y_{lm}(\theta_{\boldsymbol{p}'}, \varphi_{\boldsymbol{p}'}) = \frac{1}{4\pi} P_l(\cos \theta_{\boldsymbol{p}-\boldsymbol{p}'}), \tag{A.41}$$

where $P_l(x)$ are the Legendre polynomials. This leads to the following result for the matrix elements of the collision operator as function of angular momentum:

$$\gamma_l = \frac{1}{2s} \int d\varepsilon d\varepsilon' d\omega F_l(\varepsilon, \varepsilon', \omega) \frac{\omega f(\varepsilon) f(\varepsilon')}{|e^{-\varepsilon/T} - e^{-\varepsilon'/T}|} \sum_{\sigma=\pm} \delta(\varepsilon' - \varepsilon + \sigma\omega), \tag{A.42}$$

$$F_l(\varepsilon, \varepsilon', \omega) = \frac{D_0}{2\pi} \int_{\boldsymbol{p}\boldsymbol{p}'} [1 - P_l(\cos \theta_{\boldsymbol{p}-\boldsymbol{p}'})] \delta(\varepsilon - \varepsilon_{\boldsymbol{p}}) \delta(\varepsilon' - \varepsilon_{\boldsymbol{p}'}) \delta(\omega - \omega_{\boldsymbol{p}-\boldsymbol{p}'}). \tag{A.43}$$

Adopting the same reasoning as explained in Sec. A.1, we can ignore the $\varepsilon, \varepsilon'$ dependency of F_l for small fermionic energies. The result then simplifies considerably and gives for $\omega < \omega_D$

$$F_l(\omega) = \frac{4\lambda_{\mathrm{ep}} s}{\omega_D} \left(\frac{\omega}{\omega_D}\right) \left(1 - P_l\left(1 - \left(\frac{\omega}{\omega_D}\right)^2\right)\right), \tag{A.44}$$

with the same convention for the electron-phonon coupling constant λ_{ep} as used earlier. This yields

$$\gamma_l = \frac{2\lambda_{\mathrm{ep}} T^3}{\omega_D^2} \int_0^{\frac{\omega_D}{T}} dx \frac{x^3 \left(1 - P_l\left(1 - \left(\frac{T}{\omega_D}\right)^2 x^2\right)\right)}{\cosh(x) - 1}. \tag{A.45}$$

For $l = 1$ and l =2 we recover, of course, the known results for the scattering rates relevant in the resistivity without drag

$$\gamma_1 = \begin{cases} 480\zeta(5) \lambda_{\mathrm{ep}} T^5/\omega_D^4 & \text{if } T \ll \omega_D \\ \lambda_{\mathrm{ep}} T & \text{if } T \gg \omega_D \end{cases}, \tag{A.46}$$

and for the viscosity

$$\gamma_2 = \begin{cases} 3\gamma_1 & \text{if } T \ll \omega_D \\ 2\gamma_1 & \text{if } T \gg \omega_D \end{cases}. \tag{A.47}$$

To analyze the rate for arbitrary l we first use $1 - P_l(1 - y^2) \approx \frac{1}{2} l(l+1) y^2 (1 + \mathcal{O}(l^2 y^2))$. This expansion is sufficient for temperatures $T \ll \omega_D/l$ and yields after a few steps $\gamma_l = \frac{1}{2} l(l+1) \gamma_1$. The situation is more subtle in the regime $\omega_D/l \ll T \ll \omega_D$. To analyze the large-$l$ behavior we split $\gamma_l = \delta\gamma_0 - \delta\gamma_l$ where

$$\delta\gamma_l = \frac{2\lambda_{\mathrm{ep}} T^3}{\omega_D^2} \int_0^{\frac{\omega_D}{T}} dx \frac{x^3 P_l\left(1 - \left(\frac{T}{\omega_D}\right)^2 x^2\right)}{\cosh(x) - 1}. \tag{A.48}$$

Notice that $\delta\gamma_l < \delta\gamma_0$ for $l \geq 1$. Next, we employ the identity

$$\sum_{l=0}^{\infty} P_l\left(1-x^2\right) t^l = \frac{1}{\sqrt{(t-1)^2 + 2tx^2}} \tag{A.49}$$

and obtain for the generating function

$$\delta\gamma\left(t\right) = \sum_{l=0}^{\infty} \delta\gamma_l t^l = \frac{2\lambda_{\mathrm{ep}} T^3}{\omega_D^2} \int_0^{\frac{\omega_D}{T}} dx \frac{x^3}{\sqrt{(t-1)^2 + 2t\left(\frac{T}{\omega_D}\right)^2 x^2 \left(\cosh\left(x\right)-1\right)}}. \tag{A.50}$$

The behavior of $\delta\gamma\left(t \to 1 - 0^+\right)$ determines the large-l asymptotics of $\delta\gamma_l$. The generating function has a well defined limit as $t \to 1$ with leading corrections that are linear in $1-t$. This implies that $\delta\gamma_l$ cannot decay slower than l^{-2}. Hence in the regime $T \ll \omega_D/l$ follows that

$$\gamma_{l\to\infty} = \delta\gamma_0 = \begin{cases} 24\zeta\left(3\right)\lambda_{\mathrm{ep}} T^3/\omega_D^2 & \text{if } T \ll \omega_D \\ 2\lambda_{\mathrm{ep}} T & \text{if } T \gg \omega_D \end{cases} \tag{A.51}$$

This analysis reveals that the behavior at low temperatures and for a sufficiently small angular momentum modes l can be captured via diffusion processes on the Fermi surface. However, at any finite T, high angular modes with angular momentum $l \gg \omega_D/T$ undergo super-diffusion. These results are further discussed in section 2.4 in the main text.

References

[1] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, J. Schmalian, Annalen der Physik **529**, 1700043 (2017).

[2] A. Lucas and K. C. Fong, Journal of Physics: Condensed Matter **30**, 053001 (2018).

[3] L. W. Molenkamp and M. J. M. de Jong, Phys. Rev. B **49**, 5038 (1994).

[4] M. J. M. de Jong and L. W. Molenkamp, Phys. Rev. B **51**, 13389 (1995).

[5] H. Predel, H. Buhmann, L. W. Molenkamp, R. N. Gurzhi, A. N. Kalinenko, A. I. Kopeliovich, and A. V. Yanovsky, Phys. Rev. B **62**, 2057 (2000).

[6] X. P. A. Gao, G. S. Boebinger, A. P. Mills, Jr., A. P. Ramirez, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **94**, 086402 (2005).

[7] A. Gupta, J. J. Heremans, S. Fallahi, G. C. Gardner, M. J. Manfra, preprint arXiv:2002.12806 [cond-mat.mes-hall].

[8] D. A. Bandurin, I. Torre, R. Krishna Kumar, M. Ben Shalom, A. Tomadin, A. Principi, G. H. Auton, E. Khestanova, K. S. Novoselov, I. V. Grigorieva, L. A. Ponomarenko, A. K. Geim, M. Polini, Science **351**, 1055 (2016).

[9] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, K. Chung Fong, Science **351**, 1058 (2016).

[10] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim Phys. Rev. Lett. **116**, 136802 (2016).

[11] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, A. K. Geim, Nature Physics **13**, 1182 (2017).

[12] Y. Nam, D.-K. Ki, Da. Soler-Delgado, and A. F. Morpurgo, Nat. Phys. **13**, 1207 (2017).

[13] D. A. Bandurin, A. V. Shytov, L. Levitov, R. Krishna Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, G. Falkovich, Nature Communications **9**, 4533 (2018).

[14] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, D. A. Bandurin, Science **364**, 162 (2019).

[15] C. Tan, D. Y. H. Ho, Lei Wang, J. I. A. Li, I. Yudhistira, D. A. Rhodes, T. Taniguchi, K. Watanabe, K. Shepard, P. L. McEuen, C. R. Dean, S. Adam, J. Hone, preprint arXiv:1908.10921 [cond-mat.mes-hall].

[16] Ramzy Daou, Raymond Frésard, Sylvie Hébert, and Antoine Maignan, Phys Rev. B **91**, 041113(R) (2015).

[17] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Science **351**, 1061 (2016).

[18] N. Nandi, T. Scaffidi, P. Kushwaha, S. Khim, M. E. Barber, V. Sunko, F. Mazzola, P. D. C. King, H. Rosner, P. J. W. Moll, M. König, J. E. Moore, S. Hartnoll, A. P. Mackenzie, npj Quantum Materials **3**, 66 (2018).

[19] C. W. Hicks, A. S. Gibbs, A. P. Mackenzie, H. Takatsu, Y. Maeno, and E. A. Yelland Phys. Rev. Lett. **109**, 116401 (2012).

[20] E. Mun, H. Ko, G. J. Miller, G. D. Samolyuk, S. L. Budko, and P. C. Canfield Phys. Rev. B **85**, 035135 (2012).

[21] C. Fu, T. Scaffidi, J. Waissman, Y. Sun, R. Saha, S. J. Watzman, A. K. Srivastava, G. Li,W. Schnelle, P.Werner, M. E. Kamminga, S. Sachdev, S. S. P. Parkin, S. A. Hartnoll, C. Felser, and J. Gooth, preprint arXiv:1802.09468 [cond-mat.mtrl-sci].

[22] J. Gooth, F. Menges, C. Shekhar, V. Suth, N. Kumar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[23] A. Jaoui, B. Fauqué, C. W. Rischau, A. Subedi, C. Fu, J. Gooth, N. Kumar, V. Süß, D. L. Maslov, C. Felser, and K. Behnia, npj Quantum Materials **3**, 64 (2018).

[24] Alexandre Jaoui, Benoit Fauqué, Kamran Behnia, preprint arXiv:2006.01861 [cond-mat.str-el].

[25] M. P. Jura, M. Grobis, M. A. Topinka, L. N. Pfeiffer, K. W. West, and D. Goldhaber-Gordon Phys. Rev. B **82**, 155328 (2010).

[26] B. A Braem, F. M D Pellegrino, A. Principi, M. Röösli, C. Gold, S. Hennel, J. V Koski, M. Berl, W. Dietsche, W. Wegscheider, M. Polini, T. Ihn, K. Ensslin, Phys. Rev. B **98**, 241304 (2018).

[27] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Stern, T. Scaffidi, A. K. Geim, S. Ilani, Nature **576**, 75 (2019).

[28] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, P. Kim, A. Yacoby, R. L. Walsworth, preprint arXiv:1905.10791 [cond-mat.mes-hall].

[29] A. Jenkins, S. Baumann, H. Zhou, S. A. Meynell, D. Yang, K. Watanabe, T. Taniguchi, A. Lucas, A. F. Young, A. C. Bleszynski Jayich, preprint arXiv:2002.05065 [cond-mat.mes-hall].

[30] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics*, Course of Theoretical Physics – Vol. 10 (Elsevier India: 1st edition 2010).

[31] A. A. Abrikosov, *Fundamentals of the Theory of Metals*, (Dover Publications; Reprint edition 2017).

[32] V. F. Gantmakher and Y. B. Levinson, *Carrier scattering in metals and semiconductors*, Modern Problems in Condensed Matter Science Vol. 19, General Editors V. M. Agranovich and A. A. Maradudin (North-Holland, 1987).

[33] A.V. Andreev, Steven A. Kivelson, and B. Spivak, Phys. Rev. Lett. **106**, 256804 (2011).

[34] R. A. Davison, K. Schalm, and J. Zaanen Phys. Rev. B **89**, 245116 (2014).

[35] A. Levchenko, H.-Y. Xie, A. V. Andreev, Phys. Rev. B **95**, 121301 (2017).

[36] S. A. Hartnoll, P. K. Kovtun, M. Müller, and S. Sachdev, Phys. Rev. B **76**, 144502 (2007).

[37] M. S. Foster and I. L. Aleiner, Phys. Rev. B **79**, 085415 (2009).

[38] A. Lucas, J. Crossno, K. C. Fong, P. Kim, and S. Sachdev, Phys. Rev. B **93**, 075426 (2016).

[39] A. A. Patel, R. A. Davison, and A. Levchenko, Phys. Rev. B **96**, 205417 (2017).

[40] S. Li, A. Levchenko, A. V. Andreev, preprint arXiv:2004.13726 [cond-mat.mes-hall].

[41] L. Fritz, J. Schmalian, M. Müller, S. Sachdev, Phys. Rev. B **78**, 08505 (2008).

[42] M. Müller, J. Schmalian, L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[43] B. Spivak, S. V. Kravchenko, S. A. Kivelson, and X. P. A. Gao, Rev. Mod. Phys. **82**, 1743 (2010).

[44] P. B. Allen, Phys. Rev. Lett. **59**, 1460 (1987).

[45] R. H. M. Groeneveld, R. Sprik, and Ad. Lagendijk, Phys. Rev. B **51**, 11433 (1995).

[46] V. V. Kabanov and A. S. Alexandrov, Phys. Rev. B **78**, 174514 (2008).

[47] A. Sergeev, M. Yu. Reizer, and V. Mitin, Phys. Rev. Lett. **94**, 136602 (2005).

[48] J. Coulter, R. Sundaraman, and P. Narang, Phys. Rev. B **98**, 115130 (2018).

[49] R. E. Peierls, Ann. Phys. (5) **12**, 154 (1931).

[50] L. E. Gurevich, Zh. Eksp. Teor. Fiz. **16**, 193 (1946).

[51] R. N. Gurzhi, Sov. Phys. Usp. **11**, 255 (1968).

[52] H. Nielsen and B. I. Shklovskii, Sov. Phys. JETP **29**, 386 (1969).

[53] L. E. Gurevich and B. I. Shklovskii, Sov. Phys. Solid State **8**, 2434 (1967).

[54] V. L. Gurevich, B. D. Laikhtman, Sov. Phys. JETP **50**, 941-979 (1979).

[55] Yu. G. Gurevich and O. L. Mashkevich, Physics Reports **181**, 327 (1989).

[56] M. S. Steinberg, Phys. Rev. **109**, 1486 (1958).

[57] R. N. Gurzhi and A. I. Kopeliovich, Sov. Phys. JETP **34**, 1345 (1972).

[58] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian Phys. Rev. Lett. **120**, 196801 (2018).

[59] C. Q. Cook and A. Lucas, Phys. Rev. B **99**, 235148 (2019).

[60] R. N. Gurzhi, Zh. Eksp. Teor. Fiz. **44**, 771 (1963).

[61] R. N. Gurzhi, A. I. Kopeliovich, and S. B. Rutkevich, Adv. Phys. **36**, 221 (1987).

[62] R. N. Gurzhi, A. N. Kalinenko, and A. I. Kopeliovich, Phys. Rev. Lett. **74**, 3872 (1995).

[63] P. J Ledwith, H. Guo, L. Levitov, Ann. Phys. **411**, 167913 (2019).

[64] E. I. Kiselev, J. Schmalian, Phys. Rev. Lett **123**, 195302 (2019).

[65] C. Cercignani, *The Boltzmann Equation and Its Applications*, Springer (1988).

[66] A. A. Abrikosov and I. M. Khalatnikov, Rep. Prog. Phys. **22**, 329 (1959).

[67] Y.Liao and V. Galitski, Phys. Rev. B. **101**, 195106 (2020).

[68] W. Chen, A. V. Andreev, and A. Levchenko, Phys. Rev. B **91**, 245405 (2015).

[69] Joseph Callaway, Phys. Rev. **113**, 1046 (1958).

[70] R. N. Gurzhi, Sov. Phys. JETP **19**, 490 (1964).

[71] H. Beck, P. F. Meier, A. Thellung, Phys. Stat. Sol. (a) **24**, 11 (1974).

[72] Sangyeop Lee, Xun Li, preprint arXiv:1903.05731 [cond-mat.mes-hall].

[73] Jiecheng Zhang, Erik D. Kountz, Kamran Behnia, Aharon Kapitulnik, Proc. Natl. Acad. Sci. USA **116**, 19869 (2019).

[74] Connie H. Mousatov, Sean A. Hartnoll, preprint arXiv:1908.04792 [cond-mat.mtrl-sci cond-mat.stat-mech].

[75] Michele Simoncelli, Nicola Marzari, and Andrea Cepellotti, Phys. Rev. X **10**, 011019 (2020).

[76] Tobias Holder, Phys. Rev. B **100**, 235121 (2019).

[77] Songci Li and Dmitrii L. Maslov, Phys. Rev. B **98**, 245134 (2018).

[78] E. I. Kiselev and J. Schmalian, Phys. Rev. B **99**, 035430 (2019).
[79] R. Moessner, N. Morales-Durán, P. Surówka, and P. Witkowski Phys. Rev. B **100**, 155115 (2019).
[80] P. S. Alekseev, Phys. Rev. Lett. **117**, 166601 (2016).
[81] T. Scaffidi, N. Nandi, B. Schmidt, A. P. Mackenzie, and J. E. Moore Phys. Rev. Lett. **118**, 226601 (2017).
[82] L. V. Delacrétaz and A. Gromov Phys. Rev. Lett. **119**, 226602 (2017).
[83] F. M. D. Pellegrino, I. Torre, and M. Polini, Phys. Rev. B **96**, 195401 (2017).
[84] I. Matthaiakakis, D. Rodríguez Fernández, C. Tutschku, E. M. Hankiewicz, J. Erdmenger, R. Meyer, Phys. Rev. B **101**, 045423 (2020).
[85] A. Tomadin, G. Vignale, and M. Polini Phys. Rev. Lett. **113**, 235901 (2014).
[86] T. Holder, R. Queiroz, A. Stern, Phys. Rev. Lett. **123**, 106801 (2019).
[87] M. Shavit, A. Shytov, G. Falkovich, Phys. Rev. Lett. **123**, 026801 (2019).
[88] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*: Volume 6 Course of Theoretical Physics Series (Butterworth-Heinemann, 2 edition, 1987).
[89] Kin Chung Fong, preprint arXiv:1711.04005 [cond-mat.mes-hall].
[90] T. Holder, R. Queiroz, T. Scaffidi, N.Silberstein, A. Rozen, J. A. Sulpizio, L. Ella, S. Ilani, A. Stern, Phys. Rev. B **100**, 245305 (2019).
[91] A. V. Kavokin, B. L. Altshuler, S. G. Sharapov, P. S. Grigoryev, and A. A. Varlamov, PNAS **117**, 2846 (2020).
[92] U. Briskot, M. Schütt I. V. Gornyi, M. Titov, B. N. Narozhny, A. D. Mirlin, Phys. Rev. B **92**, 115426 (2015).
[93] Z. Sun, D. N. Basov, M. M. Fogler, PNAS **115**, 3285 (2018).
[94] P. S. Alekseev, Phys. Rev. B **98**, 165440 (2018).
[95] M. Semenyakin and G. Falkovich, Phys. Rev. B **97**, 085127 (2018).
[96] R. Moessner, P. Surówka, and P. Witkowski, Phys. Rev. B **97**, 161112(R) (2018).
[97] D. Forcella, J. Zaanen, D. Valentinis, and D. van der Marel Phys. Rev. B **90**, 035143 (2014).
[98] R. Toshio, K. Takasan, N. Kawakami, preprint arXiv:1905.11698 [cond-mat.mes-hall].
[99] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, Ha.Zhang, K. Watanabe, T. Taniguchi, F. Wang, Science **364**, 150 (2019).
[100] J. D. Crossno, *Electronic Thermal Conductance of Graphene via Electrical Noise*, (PhD Thesis, Harvard, 2017).
[101] Ya. M. Blanter and M. Büttiker, Phys. Rep. **336**, 1–166 (2000).
[102] A. H. Steinbach, J. M. Martinis, and M. H. Devoret, Phys. Rev. Lett. **76**, 3806 (1996).
[103] C. W. J. Beenakker and M. Büttiker, Phys. Rev. B **46**, 1889 (1992).
[104] K. E. Nagaev, Phys. Lett. A **169**, 103 (1992).
[105] Sh. Kogan, *Electronic Noise and Fluctuations in Solids*, (Cambridge University Press 1996).
[106] V. I. Kozub and A. M. Rudin, Phys. Rev. B **52**, 7853 (1995)
[107] M. J. M. de Jong, C. W. J. Beenakker, Physica A **230**, 219 (1996).
[108] D. A. Bagrets, Phys. Rev. Lett. **93**, 236803 (2004).
[109] D. B. Gutman, A. D. Mirlin, Yuval Gefen, Phys. Rev. B **71**, 085118 (2005).
[110] V. L. Gurevich and A. M. Rudin, Phys. Rev. B **53**, 10078 (1996).
[111] K. E. Nagaev, Phys. Rev. B **52**, 4740 (1995).
[112] Y. Naveh, D. V. Averin, and K. K. Likharev, Phys. Rev. B **58**, 15371 (1998).
[113] Y. Naveh, "XVIII Rencontres de Moriond: Quantum Physics at Mesoscopic Scale", Edited by D. C. Glattli and M. Sanquer (Edition Frontiers, France); preprint arXiv:cond-mat/9910256 [cond-mat.mes-hall].
[114] D. Belitz, Phys. Rev. B **36**, 2531 (1987).
[115] J. M. Burgers, Advances in applied mechanics **1**, 171 (1948).
[116] G. B. Whitham, *Linear and nonlinear waves*, Vol. 42. John Wiley and Sons (2011).
[117] P. T. Dumitrescu, Phys. Rev. B **92**, 121102 (2015).
[118] J. M. Link and I. F. Herbut, Phys. Rev. B **101**, 125128 (2020).

arXiv:2006.13989v3 [cond-mat.mes-hall] 18 Jun 2021

Corbino FETs in magnetic field: highly tunable photodetectors

Bailey Winstanley,[1, *] Henning Schomerus,[2] and Alessandro Principi[1, †]
[1] *Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK*
[2] *Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK*

We study gated field effect transistors (FETs) with an eccentric Corbino-disk geometry, such that the drain spans its circumference while the off-center inner ring acts as a source. An AC THz potential difference is applied between source and gate while a static source-drain voltage, rectified by the nonlinearities of FET electrons, is measured. When a magnetic field is applied perpendicular to the device, a strong resonance appears at the cyclotron frequency. The strength of the resonance can be tuned by changing the eccentricity of the disk. We show that there is an optimum value of the eccentricity that maximizes the responsivity of the FET.

I. INTRODUCTION

Electromagnetic radiation is one of the prime tools to investigate matter and its properties. This is made possible by the existence of efficient and compact sources and detectors in the whole spectrum, with the crucial exception of the low-THz range (between 0.1 and 30 THz). This fact, commonly referred to as the "terahertz gap", has slowed down technological developments in, e.g., nondestructive imaging, biosensing, and spectroscopy of materials [1–3]. In modern optoelectronics there is a deep need for efficient and tunable photodetectors that operate in this range [1–6]. Dyakonov and Shur, in 1996, predicted that a field effect transistor (FET), or any gated two-dimensional (2D) electron liquid, could be used to generate and detect THz radiation [7–9].

The device in their seminal work consists of a square semiconductor quantum-well cavity, hosting a 2D electron gas, connected to a source and a drain and in close proximity to a metal gate. When a THz AC source-gate voltage is applied, typically from incoming THz radiation impinging on an antenna, asymmetric boundary conditions and intrinsic nonlinearities of the electron fluid produce a rectified DC source-drain voltage. Resonances are observed in the rectified (photo)voltage at frequencies that allow plasmons (collective long-wavelength charge density fluctuations [10]) to undergo constructive interference. This phenomenon has been experimentally verified in semiconductor quantum wells at room [11–13] and low temperatures [14] and in graphene-based FETs [15–19].

Recently, it has been shown that the responsivity of Dyakonov-Shur THz detectors can be greatly enhanced by shaping them as Corbino disks [20]. In such geometry, the electric field becomes singular at the inner contact ring (the source), and the field enhancement results in a strong nonlinear rectification at the outer ring (the drain). Motivated by such findings, here we study similar photodetectors in a uniform magnetic field perpendicular to the electron liquid, previously performed in other geometries and shown to enhance photodetection [21 and 22]. Under this condition, the spectrum of plasmon modes, labelled by their "winding number" η, *i.e.* the number of oscillations of the electric field in the angular direction, is recontructed. Notably, the plasmon spectrum splits into two parts, revealing both bulk and edge modes. Edge magnetoplasmons have frequencies below the cyclotron frequency for values of η that are not too large. Bulk-plasmons' frequencies are instead "pushed" above the cyclotron frequency.

As shown in what follows, the energy of magnetoplasmons depends on the sign of η, with edge modes appearing only at positive winding numbers (for magnetic fields along the direction orthogonal to the disk). Furthermore, depending on device parameters and at odds with Corbino disks characterized by symmetric boundary conditions [23–25], the dispersion of bulk modes can exhibit a nearly-flat band close to the cyclotron frequency. When the radii of the source and drain rings are comparable, modes characterized by different winding numbers appear to have all very similar frequencies. Because of this feature, we would expect the response of the system to be greatly enhanced when the frequency of the external field is close to the cyclotron one, if we would be able to excite plasmon modes with different winding numbers at once. Since the cyclotron frequency can be tuned with the external magnetic field, the Corbino photodetector could be capable of selectively detecting frequencies deep in the THz gap with a high responsivity. Unfortunately, in the Corbino geometry this would require a careful fine-tuning of the potential profile at the source (inner) ring, which is highly unlikely to be realized experimentally with a simple circular contact connected to an antenna. The circular symmetry of the Corbino disk indeed forbids the mixing of modes of different winding numbers, and therefore a homogeneous potential at the source would only excite non-winding plasmons with $\eta = 0$.

To overcome this limitation, we study an "eccentric" Corbino geometry, whereby the inner source ring is off-centered and made closer to the outer edge on one side of the disk. By breaking the circular symmetry, the eccentric geometry enables the excitation of modes characterized by different winding numbers with a simple uniform source potential. The photoresponse is greatly enhanced at frequencies near the cyclotron one when the source is in close proximity of the drain. This requirement is reminiscent of the condition needed to obtain a plasmon flat band in concentric Corbino geometries. Therefore, in eccentric geometries, the photoresponse enhancement is controlled not only by the size of the inner ring, but also by its closeness to the drain. We find that, for any pair of source and drain radii, there is an *optimal* value of the eccentricity that maximizes the photoresponse.

In Sect. II we present the model of the electron cavity as a hydrodynamic fluid in the presence of a uniform perpendicular magnetic field. In Sect. III we apply said theory to model a Corbino disk. In Sect. IV we study an eccentric Corbino disk. In Sect. V we report the summary of our findings and our main conclusions. We note that the description we use holds for a variety of different systems [7–9, 18, and 26], and therefore our predictions have a broad range of applicability.

II. THE MODEL OF THE CAVITY

We consider a general FET, where the active component is a 2D electron liquid placed in close proximity to a metal gate. The geometry used in this paper is that of a Corbino disk with source and drain electrodes attached to the inner and outer edges, respectively. It should be noted, however, that the following applies to general 2D geometries. A radiation field oscillating at frequency ω is applied between the source and the gate, typically via an antenna, while the drain is left fluctuating, *i.e.* no current flows through it. We will study rectification of the oscillating field due to the intrinsic hydrodynamic nonlinearities of the electron liquid [18, 27–32] (we discuss below the applicability of such model). A rectified DC source-drain potential difference, proportional to the power of the incident radiation, is therefore measured between source and drain at zero applied bias.

Since we focus on the long-wavelength low-frequency dynamics of the electron liquid, we model it by means of hydrodynamic equations [18 and 33]. These govern the relationship between the density, current and electric field within the device. We stress that equations formally equivalent to hydrodynamic ones can be derived by inverting the nonlinear relation between current and electric field of the electron fluid [18], with no reference to typical scattering times [34] (*i.e.* the relations hold true also for non-interacting electrons). Therefore hydrodynamic equations should be seen here as an efficient way to incorporate nonlinearities in the long-wavelength description of the electron liquid. The first of these relations is the continuity equation, $\partial_t \rho(\boldsymbol{r},t)+\nabla\cdot[\rho(\boldsymbol{r},t)\boldsymbol{v}(\boldsymbol{r},t)]=0$, which connects the periodic accumulation of charge density due to the oscillating radiation field, $\rho(\boldsymbol{r},t)$, to the flow velocity, $\boldsymbol{v}(\boldsymbol{r},t)$. Since electrons are charged, $\rho(\boldsymbol{r},t)$ induces a nonlocal Hartree-like electric potential according to [18]

$$U(\boldsymbol{r},t)=\int d\boldsymbol{r}' V(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}',t) \ , \tag{1}$$

which in turn acts as the restoring force that sustains charge oscillations in a feedback loop. In Eq. (1), $V(\boldsymbol{r}-\boldsymbol{r}')$ is the Coulomb interaction between two charges at positions $\boldsymbol{r}$ and $\boldsymbol{r}'$. The nearby gate, which we assume to be a perfect conductor, has an important effect: mirror charges screen the tail of the Coulomb interaction and make it effectively short-ranged. In view of this fact, and to simplify the following derivation, we will employ the so-called "local-gate approximation" [15 and 18]. The latter consists in assuming a local relation between the self-induced field and charge density,

$$U(\boldsymbol{r},t)=\rho(\boldsymbol{r},t)/C \ , \tag{2}$$

in lieu of the nonlocal one of Eq. (1). This approximation has been shown [24 and 35] to well reproduce results obtained with Eq. (1) when the gate is explicitly accounted for. In the specific case under consideration, it allows for the emergence of edge magnetoplasmons in both semi-infinite planes and hollow disks. Using the local-gate relation between electric potential and charge density, the continuity equation becomes

$$\partial_t U(\boldsymbol{r},t)=-\nabla\cdot[U(\boldsymbol{r},t)\boldsymbol{v}(\boldsymbol{r},t)] \ . \tag{3}$$

The equation relating the flow velocity to the self-induced field is assumed to have the following Euler-like form [18]

$$\begin{aligned}\frac{e}{m}\nabla U(\boldsymbol{r},t)&=\partial_t\boldsymbol{v}(\boldsymbol{r},t)+\frac{1}{\tau}\boldsymbol{v}(\boldsymbol{r},t)+\omega_c\hat{\boldsymbol{z}}\times\boldsymbol{v}(\boldsymbol{r},t)\\&+\big[\boldsymbol{v}(\boldsymbol{r},t)\cdot\nabla\big]\boldsymbol{v}(\boldsymbol{r},t)+\boldsymbol{v}(\boldsymbol{r},t)\times\big[\nabla\times\boldsymbol{v}(\boldsymbol{r},t)\big] \ .\end{aligned} \tag{4}$$

In these equations, $-e$ is the electron charge, m their effective mass and τ the average time between two successive momentum-non-conserving collisions with impurities or phonons. Finally, $\omega_c=eB/m$ is the cyclotron frequency and B is the magnetic field applied orthogonal to the 2D electron liquid. The term $\boldsymbol{v}(\boldsymbol{r},t)\times\big[\nabla\times\boldsymbol{v}(\boldsymbol{r},t)\big]$, known as the Lamb vector, represents a nonlinear Lorentz force due to the vortical movement of the electron fluid itself [36], and can be combined with the term $\big[\boldsymbol{v}(\boldsymbol{r},t)\cdot\nabla\big]\boldsymbol{v}(\boldsymbol{r},t)$ into the single term $\nabla v^2(\boldsymbol{r},t)/2$. We solve the problem posed by the hydrodynamic equations (3)-(4) in conjunction with the usual Dyakonov-Shur boundary conditions

$$\begin{aligned}&U(\boldsymbol{r},t)|_{\text{source}}=U_{\text{ext}}(\boldsymbol{r})\cos(\omega t) \ ,\\&\hat{\boldsymbol{n}}\cdot\boldsymbol{v}(\boldsymbol{r},t)|_{\text{drain}}=0 \ ,\end{aligned} \tag{5}$$

corresponding to an oscillating gate-source potential from the antenna output and an open-circuit drain. Here $\hat{\boldsymbol{n}}$ is the unit vector normal to the drain surface.

To solve the problem above, we resort to a perturbative treatment of the system of nonlinear equations. We assume U_{ext} to be a small parameter and calculate the rectified nonlinear response as a perturbation to the potential. We then expand

$$\begin{aligned}&U(\boldsymbol{r},t)=U_0+U_1(\boldsymbol{r},t)+U_2(\boldsymbol{r},t)+\mathcal{O}(U_{\text{ext}}^3),\\&\boldsymbol{v}(\boldsymbol{r},t)=\boldsymbol{v}_1(\boldsymbol{r},t)+\boldsymbol{v}_2(\boldsymbol{r},t)+\mathcal{O}(U_{\text{ext}}^3).\end{aligned} \tag{6}$$

Here $U_0 < 0$ is the equilibrium gate potential (which fixes the charge density in the FET according to $\rho_0 = CU_0$), and the equilibrium velocity, $\boldsymbol{v}_0$, is zero by definition. $U_1(\boldsymbol{r}, t)$ and $\boldsymbol{v}_1(\boldsymbol{r}, t)$, and $U_2(\boldsymbol{r}, t)$ and $\boldsymbol{v}_2(\boldsymbol{r}, t)$ are the linear (order $U_{\rm ext}$) and nonlinear (order $U_{\rm ext}^2$) contributions to the potential and velocity, respectively. Note that, although small, $U_2(\boldsymbol{r}, t)$ is responsible for the only nontrivial DC rectified potential, which can be detected by measuring an averaged source-to-drain voltage drop [7–9, 15, and 18].

Plugging the expansions in Eq. (6) into the set of equations (2)-(5), we collect terms of order $U_{\rm ext}$ and $U_{\rm ext}^2$ into two systems of differential equations, which are linear in $U_1(\boldsymbol{r}, t)$ and $\boldsymbol{v}_1(\boldsymbol{r}, t)$, and $U_2(\boldsymbol{r}, t)$ and $\boldsymbol{v}_2(\boldsymbol{r}, t)$, respectively. The former yields the linear response of the system which oscillates at the same frequency as the external source-gate perturbation potential, *i.e.* $U_1(\boldsymbol{r}, t) = U_1(\boldsymbol{r})e^{-i\omega t} + U_1^*(\boldsymbol{r})e^{i\omega t}$ and $\boldsymbol{v}_1(\boldsymbol{r}, t) = \boldsymbol{v}_1(\boldsymbol{r})e^{-i\omega t} + \boldsymbol{v}_1^*(\boldsymbol{r})e^{i\omega t}$. Conversely, the system of equations for $U_2(\boldsymbol{r}, t)$ and $\boldsymbol{v}_2(\boldsymbol{r}, t)$ yields solutions oscillating at $\pm 2\omega$ and a rectified (time-independent) one. To focus on the latter part of the potential $U_2(\boldsymbol{r}, t)$, we average equations over time by integrating over a period of oscillation, $T = 2\pi/\omega$. In this way, the time-dependent parts of $U_2(\boldsymbol{r}, t)$ and $\boldsymbol{v}_2(\boldsymbol{r}, t)$ vanish.

The details of the derivation are given in App. A. The linear systems of equations for $U_1(\boldsymbol{r}, t)$ and $\boldsymbol{v}_1(\boldsymbol{r}, t)$, and $U_2(\boldsymbol{r}, t)$ and $\boldsymbol{v}_2(\boldsymbol{r}, t)$ read

$$\begin{cases} \left[\omega_c^2 - \omega^2 f_\omega^2\right] U_1(\boldsymbol{r}) - s^2 f_\omega \nabla^2 U_1(\boldsymbol{r}) = 0 \\ U_1(\boldsymbol{r})\Big|_{\rm source} = U_{\rm ext}(\boldsymbol{r}) \\ \hat{\boldsymbol{n}} \cdot \left[i\omega f_\omega \nabla U_1(\boldsymbol{r}) + \omega_c \hat{\boldsymbol{z}} \times \nabla U_1(\boldsymbol{r})\right]\Big|_{\rm drain} = 0 \end{cases} \tag{7}$$

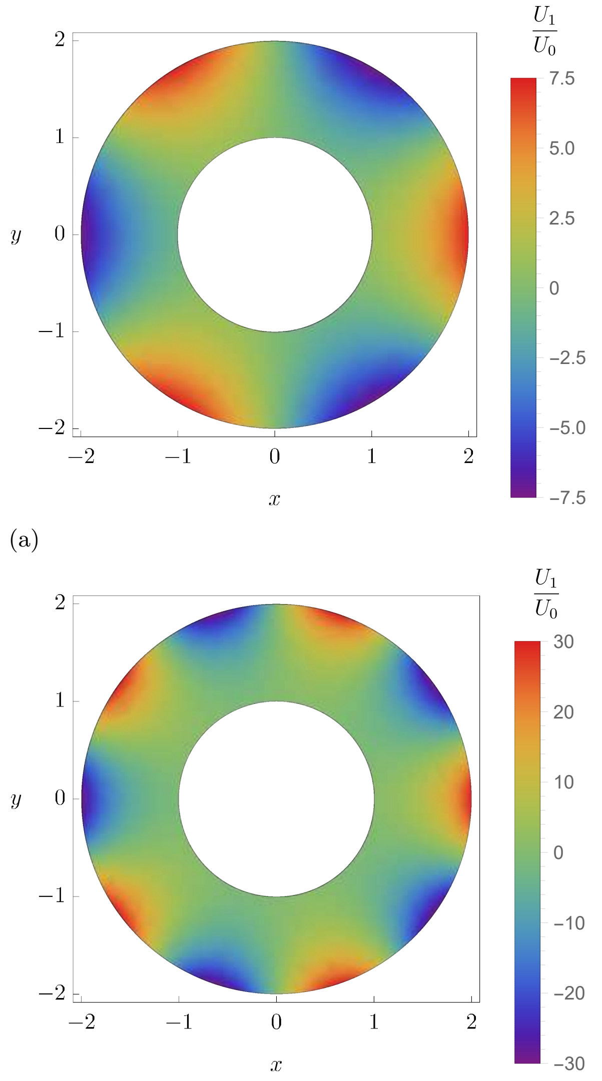
where $s = \sqrt{-eU_0/m}$ is the plasma wave velocity, $f_\omega = 1 + i/(\omega\tau)$, and

$$\begin{cases} \dfrac{1 + (\tau\omega_c)^2}{U_0 \tau} \nabla \cdot \left[U_1^*(\boldsymbol{r})\boldsymbol{v}_1(\boldsymbol{r}) + U_1(\boldsymbol{r})\boldsymbol{v}_1^*(\boldsymbol{r})\right] = \nabla^2 \phi(\boldsymbol{r}) \\ \phi(\boldsymbol{r}) - \boldsymbol{v}_1^*(\boldsymbol{r}) \cdot \boldsymbol{v}_1(\boldsymbol{r})\Big|_{\rm source} = 0 \\ \hat{\boldsymbol{n}} \cdot \left[\omega_c \hat{\boldsymbol{z}} \times \nabla\phi(\boldsymbol{r}) - \dfrac{1}{\tau}\nabla\phi(\boldsymbol{r})\right]\Big|_{\rm drain} = 0 \end{cases} . \tag{8}$$

Here, $\phi(\boldsymbol{r}) = \boldsymbol{v}_1^*(\boldsymbol{r}) \cdot \boldsymbol{v}_1(\boldsymbol{r}) - eU_2(\boldsymbol{r})/m$ and

$$\boldsymbol{v}_1(\boldsymbol{r}) = \frac{s^2}{U_0} \frac{i\omega f_\omega \nabla U_1(\boldsymbol{r}) + \omega_c \hat{\boldsymbol{z}} \times \nabla U_1(\boldsymbol{r})}{\omega_c^2 - \omega^2 f_\omega^2} . \tag{9}$$

The Poisson problem in Eqs. (8) admits a unique solution for $\phi(\boldsymbol{r})$ and therefore for $U_2(\boldsymbol{r}) = m/e\big[\boldsymbol{v}_1^*(\boldsymbol{r}) \cdot \boldsymbol{v}_1(\boldsymbol{r}) - \phi(\boldsymbol{r})\big]$. In the absence of a magnetic field, the photoresponse of the system will exhibit resonances at given frequencies dependant on the geometry of the system. The lowest of these frequencies is denoted as ω_B and determined numerically for any given disk geometry for later use (see the following sections).



(a)

(b)

FIG. 1. Density plots of the real part of the linear potential, $U_1(\boldsymbol{r})$ for $\omega = \omega_c = 5\ s/r_0$. Panel (a) is evaluated at $\eta = 3$. We can observe three complete oscillations of the potential around the circumference. Panel (b) is evaluated at $\eta = 5$. Similarly, here we observe five complete oscillations. These are evaluated at $\tau = 5\ r_0/s$, where $r_1 = 2\ r_0$.

III. CONCENTRIC CORBINO DISK

We first solve Eqs. (7)-(8) for a concentric Corbino disk akin to the one studied in Ref.[20], whose inner (source) and outer (drain) radii are r_0 and r_1, respectively. In this geometry, we can readily solve Eqs. (7) analytically and determine the full spectrum of magnetoplasmon modes, owing to the inherent rotational symmetry of the system. Such symmetry enables the separation of radial and angular variables within the solution. We note that our study differs from that of Ref.[20] in two respects. Firstly, we consider the role of the magnetic field in modifying the spectrum of magnetoplasmons. Secondly, we consider the impact of source-to-gate voltages having a finite (integer) winding number η. We therefore impose that $U_1(\boldsymbol{r})$ is equal to $\bar{U}_{\rm ext}\cos(\eta\theta)$ at the source, where $\bar{U}_{\rm ext}$ is the magnitude of the external potential and θ is the angle between $\boldsymbol{r}$ and the $\hat{\boldsymbol{x}}$-axis. Hence the linear solution will have winding numbers $\pm\eta$. Defining $k^2 = (\omega_c^2 - \omega^2 f_\omega^2)\mathrm{sign}(\omega_c^2 - \omega^2)/(s^2 f_\omega)$, the solutions of the system of linear differential equations (7) takes the form $U_1(\boldsymbol{r}) = U_1^{(\eta)}(r,\theta) + U_1^{(-\eta)}(r,\theta)$, where

$$U_1^{(\eta)}(r,\theta) = \begin{cases} [A_\eta I_\eta(kr) + B_\eta K_\eta(kr)]e^{i\eta\theta} \ , & \text{if } \omega^2 < \omega_c^2 \\ [C_\eta J_\eta(kr) + D_\eta Y_\eta(kr)]e^{i\eta\theta} \ , & \text{if } \omega^2 > \omega_c^2 \end{cases} \tag{10}$$

Here, $J_\eta(x) = J_{-\eta}(x)$ $[I_\eta(x) = I_{-\eta}(x)]$ and $Y_\eta(x) = Y_{-\eta}(x)$ $[K_\eta(x) = K_{-\eta}(x)]$ are (modified) Bessel functions of the first and second kind, respectively. The coefficients A_η, B_η, C_η and D_η are determined by applying the boundary conditions. After some lengthy but straightforward algebra we find, for $|\omega| < |\omega_c|$,

$$U_1^{(\eta)}(r,\theta) = \frac{\bar{U}_{ext}}{2}\left[\frac{I_\eta(kr)}{I_\eta(kr_0)} - \frac{I'_\eta(kr_1) - \gamma_\eta I_\eta(kr_1)}{\mathcal{D}_\eta(\omega) I_\eta(kr_0)} \right. \\ \left. \times \left(\frac{K_\eta(kr)}{K_\eta(kr_0)} - \frac{I_\eta(kr)}{I_\eta(kr_0)}\right)\right] e^{i\theta\eta}, \tag{11}$$

where $\gamma_\eta = \omega_c\eta/(\omega f_\omega k r_1)$, $I'_\eta(x) = dI_\eta(x)/dx$, $K'_\eta(x) = dK_\eta(x)/dx$, and

$$\mathcal{D}_\eta(\omega) = \frac{K'_\eta(kr_1) - \gamma_\eta K_\eta(kr_1)}{K_\eta(kr_0)} - \frac{I'_\eta(kr_1) - \gamma_\eta I_\eta(kr_1)}{I_\eta(kr_0)}. \tag{12}$$

For $|\omega| > |\omega_c|$, $U_1^{(\eta)}(r,\theta)$ has the same form of Eqs. (11)-(12), with $J_n(kr)$ and $Y_n(kr)$ in lieu of $I_n(kr)$ and $K_n(kr)$, respectively. In Fig. 1 we plot the real part of the linear potential $U_1^{(\eta)}(\boldsymbol{r})$. Counting oscillations at the outer perimeter of the disk (the drain), it can be seen that the two edge plasmons produced by manual injection at the source have $\eta = 3$ [panel (a)] and $\eta = 5$ [panel (b)], respectively. In this figure we scale the electrical potential with U_0, lengths with the source radius r_0 and times with r_0/s.

Bulk and edge magnetoplasmons can be identified as the zeros of $\mathcal{D}_\eta(\omega)$ and its counterpart for $|\omega| > |\omega_c|$. For $\omega_c > 0$, the frequencies of magnetoplasmon modes as a function of the winding number η are shown in Fig. 2 (a). In this figure potential, lengths and times are given in the same units of Fig. 1. For convenience, frequencies are scaled with the first resonant frequency at zero magnetic field, ω_B. There, ungapped edge modes are seen to wind in the $+\hat{\boldsymbol{\theta}}$ direction (as they only exist for positive η) and are localised at the outer edge of the disk. Winding in the opposite direction cannot occur as plasmons would be bound to the inner edge, which is however held at a fixed potential.

We also observe that bulk modes exhibit a variable degree of asymmetry: in general, the frequencies are higher for magnetoplasmons characterized by negative winding numbers. The asymmetry can be traced back to γ_η defined after Eq. (11), the only parameter that depends on the sign of η. Physically, this asymmetry arises from the relative alignment between the Lorentz force induced by the magnetic field, acting on the plasmons' constituent electrons, and the plasmons' electric field. The splitting in frequency of bulk modes can be observed in Fig. 2 (b) where upper branches refer to negative winding numbers. In passing, we note that analogous splittings of frequencies of bulk modes have previously been observed in conventional disk geometries [23 and 24] as well as Corbino disks with symmetric boundary conditions [25]. In contrast, the present case, characterized by asymmetric boundary conditions, admits an additional nearly-flat band of normal modes. In fact, while the lowest branch of negative-η bulk modes displays an approximately linear dispersion, positive-η modes oscillate at around the cyclotron frequency. The latter plasmon nearly-flat band has no counterpart in conventional disks [23 and 24] or Corbino disks under symmetric boundary conditions [25]. We stress that the nearly-flat band becomes a clear feature of the spectrum only when the inner and outer radii of the Corbino disk are comparable. When this is not realized, it becomes unstable against the introduction of a small damping $1/\tau$, and the conventional disk solution is recovered [23 and 24].

The flat plasmon band at $\omega \simeq \omega_c$ and $\eta > 0$ in Fig. 2 (a) has an important consequence for the nonlinear responsivity of the Corbino disk. For every external source-to-gate potential $\bar{U}_{\rm ext}\cos(\eta\theta)$, we expect the nonlinear rectified potential $U_2(\boldsymbol{r})$ to exhibit a resonance at $\omega \simeq \omega_c$. In fact, $U_1(r_0,\theta)$ can be decomposed into the sum of two counter-winding potentials, characterized by winding numbers $\pm\eta$, one of which (depending on the direction of the magnetic field and the sign of ω_c) can excite a magnetoplasmon mode at the cyclotron frequency. In turn, such mode produces a rectified voltage $U_2(\boldsymbol{r})$ at the outer rim of the Corbino disk. We note that such voltage, thanks to the interference between oppositely-winding magnetoplasmons, not only is time-independent but it also contains a non-winding component characterized by $\eta = 0$ that does not vanish when integrated over the drain.

In Fig. 3(a) we show $U_2(\boldsymbol{r})$, obtained by numerically solving Eq. (8), integrated over the outer rim of the Corbino disk (*i.e.* the drain) for the first few values of η and as a function of ω. We clearly recognize a resonance at $\omega \simeq \omega_c$ for all values of η. In Fig. 3(b), we show how the maximum of such resonance scales with η.

Such result has an attractive implication. If we would be able to excite at once magnetoplasmons of frequency $\omega \simeq \omega_c$ in a broad range of winding numbers, the resulting resonance would grow to become particularly strong, therefore greatly enhancing the responsivity of the device. Furthermore, its position could be tuned by changing the external magnetic field, and it could be made to span the THz range practically at will. Unfortunately, the current geometry does not allow to easily achieve such result: to excite magnetoplasmons with different winding numbers it is necessary to carefully engineer the potential applied at the source. This requires superimposing various harmonics characterized by different values of η, a fact that is at present experimentally challenging.

For this reason, we will now move to study the experimentally more relevant case of an *eccentric* Corbino disk. In fact, while in the Corbino disk circular symmetry leads to the decoupling of various modes, the lack of symmetry of the eccentric disc allows their mixing. In turn, this enables the use of more realistic source potentials (*i.e.* uniform along the inner ring) to access the strong resonance at $\omega \simeq \omega_c$, as we proceed to show.

IV. ECCENTRIC CORBINO DISK

The eccentric Corbino FET geometry is shown schematically in Fig. 4. In this geometry, the inner source ring is shrunk and placed off-centre. The non-linear hydrodynamic problem, with the asymmetric boundary conditions of Eq. (5), can be solved numerically as described in Sect. II. First, Eqs. (7) are solved for the linear potential. Then, by using Eq. (9), Eqs. (8) are solved for $\phi(\boldsymbol{r})$. From the latter, we can then calculate the non-linear potential $U_2(\boldsymbol{r})$. Since both Eqs. (7) and (8) are Poisson problems, they admit unique solutions for a given set of boundary conditions. We define the eccentricity as $\xi = d/r_1$ where d is the distance of the centre of the source from the centre of the disk, and r_1 is the outer radius of the disk. To aid comparison with the previous section, we keep the drain radius identical to that of the concentric Corbino disk, and we therefore scale lengths [times] with $r_1/2$ $[r_1/(2s)]$. Similarly, as in the previous section, frequencies will be scaled by ω_B, the lowest resonance frequency at zero magnetic field determined numerically for any given geometry.

We plot the non-linear potential $U_2(\boldsymbol{r})$, integrated along the drain, as a function of the AC driving frequency for various magnetic field strengths in Fig. 5 (a). For each curve, resonances at $\omega < \omega_c$ correspond to edge modes, while those at $\omega \geq \omega_c$ can be due to both bulk or edge ones. Now that the source has been placed off-centre and close to the drain, we can see that edge plasmons with differing winding numbers, and hence different frequencies, can propagate. As an example, for $\omega_c = 7\omega_B$, we can see three edge modes below the cyclotron frequency (of frequencies $\omega/\omega_B \approx 1.5,\ 3.2,\ 4.8$) and one mode above it (at $\omega/\omega_B \approx 8.4$). As expected from the discussion in the previous section, for all field strengths the first bulk mode, fixed around the cyclotron frequency, results in the largest resonance peak.

It should be noted that although Fig. 5 (a) is obtained by setting the eccentricity $\xi = 0.95$, this is not the optimum value that maximizes the photoresponse. In fact Fig. 5 (b) shows that for a source of radius $r_0 = 0.05\, r_1/2$, with $\omega_c = 7\omega_B$ as used in panel (a), the optimum eccentricity is $\xi \approx 0.8$. Fig. 5 (b) further shows that the optimum eccentricity is inversely proportional to the source radius, r_0. It can be further shown that it increases with the drain radius, r_1, and cyclotron frequency, ω_c. As such the geometry of such a device must be tailored to the expected frequency of incoming light.

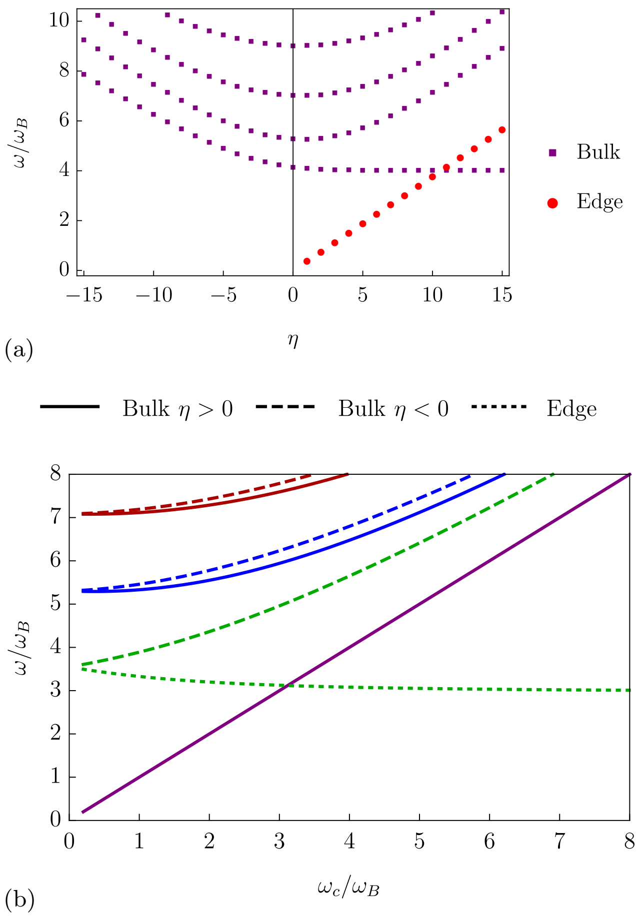


FIG. 2. Panel (a) The resonant frequencies of the *linear* potential, $U_1^{(\eta)}(r,\theta)$, obtained from Eq. (11) and plotted against the winding number η defined before Eq. (10). Bulk magnetoplasmon modes are represented by purple squares, while edge magnetoplasmons are represented by red circles. Panel (b) The resonant frequencies of the *linear* potential, $U_1^{(\eta)}(\boldsymbol{r})$, against cyclotron frequency, at fixed $|\eta| = 8$. Solid (dashed) lines refer to plasmons propagating in the counterclockwise (clockwise) direction. The dotted line denotes the edge state. The purple solid line corresponds to the mode oscillating at the cyclotron frequency. Units are the same as in Fig. 1. Both figures are obtained in the limit $\tau \to \infty$.

We now wish to briefly comment on the feasibility of our device. We consider an FET based on doped bilayer graphene at relatively small (*i.e.* non-quantizing) magnetic fields, with dimensions on the order of a few micrometers: similar devices have been recently realised and shown to be significantly tunable via the application of gate voltage [26]. Given the lowest bulk plas-

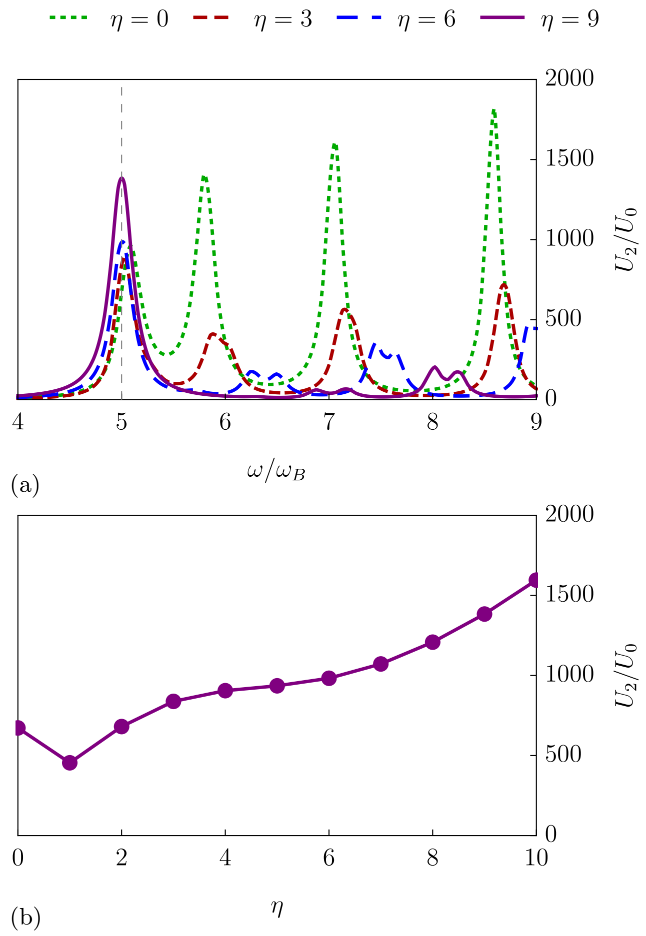


FIG. 3. Panel (a) The *non-linear* potential at the drain, obtained by numerically solving Eqs. (8), plotted against frequency, at $\omega_c = 5\omega_B$. Different curves correspond to different winding numbers, η, of the source potential. We note that the first bulk mode remains pinned at $\omega = \omega_c$ and only increases in intensity with $|\eta|$, while all other modes slowly shift towards higher frequencies. The splitting of higher-order bulk modes becomes more and more evident at larger η: peaks split in two, as seen for e.g. $\eta = 6$. Units are the same as in Figs. 1 and 2. For all curves we have set the collision time $\tau = 5\ r_0/s$, and the outer radius, $r_1 = 2\ r_0$. Panel (b) The value of the nonlinear potential at $\omega = \omega_c$ as a function of η. The dip at $\eta = 1$ is due to the fact that, for small values of η, the peak is slightly shifted to the right.

mon frequency of such devices [17, 26, 37, and 38], $\omega_B = 300$ GHz (this is typically dependent on system size and for graphene can be changed via the gate voltage), and an effective electron mass [26], $m \approx 0.036 m_e$, where m_e is the free electron mass, we can estimate the lower limit for the magnetic field. The lowest observable edge plasmon frequency is always similar to the lowest bulk plasmon frequency provided the source radius is small, thus, by equating the lowest bulk plasmon frequency with the cyclotron frequency, $\omega_c = eB/m^*$, our estimate for the minimum magnetic field becomes $B_{\rm min} \approx 0.06$ T. This magnitude is easily achievable in experiments. In passing, we mention that alternatives to applying an external magnetic field do exist [33 and 39].

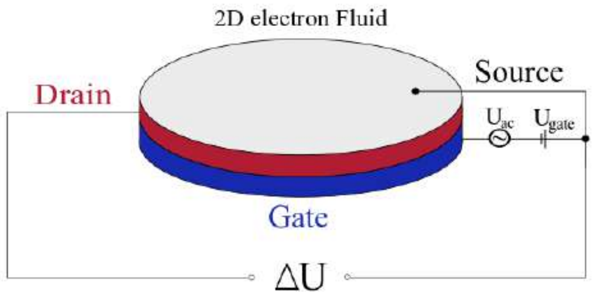


FIG. 4. Schematic of the eccentric Corbino disk FET studied in this paper. The perimeter of the device acts as the drain while a finite small source is connected to the top of the cavity. $U_{\rm gate}$ is the back gate DC bias voltage, which in our case is constant in time and used to fix the charge density. The FET rectifies the AC source-gate voltage, $U_{\rm ac}(t)$, into the DC source-drain voltage ΔU.

V. CONCLUSIONS

In this paper we have studied Corbino-disk-shaped photodetectors with sources at the inner ring which oscillate at the frequency of the incoming radiation with respect to metallic back-gates. The design is similar to that of conventional Dyakonov-Shur devices, in that a rectified potential is measured at the outer rim of the disk, which acts as a fluctuating drain. By applying a magnetic field in the direction perpendicular to the cavity, the rectification of long-wavelength radiation occurs from the constructive interference of not only bulk plasmons, but also edge magnetoplasmons.

In this geometry, plasmons can circulate along the entirety of the disk's perimeter nearly unimpeded [40]. Plasmons in this configuration are categorised by their winding number, *i.e.* the number of complete oscillations of the electron density that occur over a full revolution around the disk. In the first part of the paper, we studied the response of a conventional Corbino-disk photodetector with the source-ring located at the centre of the disk. Said geometry admits an analytic solution. In this configuration individual plasmon modes can be manually injected by selecting the winding number of the external source-to-gate potential. It is important to note that, as shown in Sect. III, all modes, and in particular ones at $\omega \simeq \omega_c$ (which exist only in the presence of asymmetric boundary conditions and up to large winding numbers, as long as inner and outer radii are comparable and damping is small), produce nonlinear rectified potentials that are also uniform along the edge. Therefore, all contributions at $\omega \simeq \omega_c$ can in principle be summed up, with a careful choice of the source-to-gate external potential, and re-

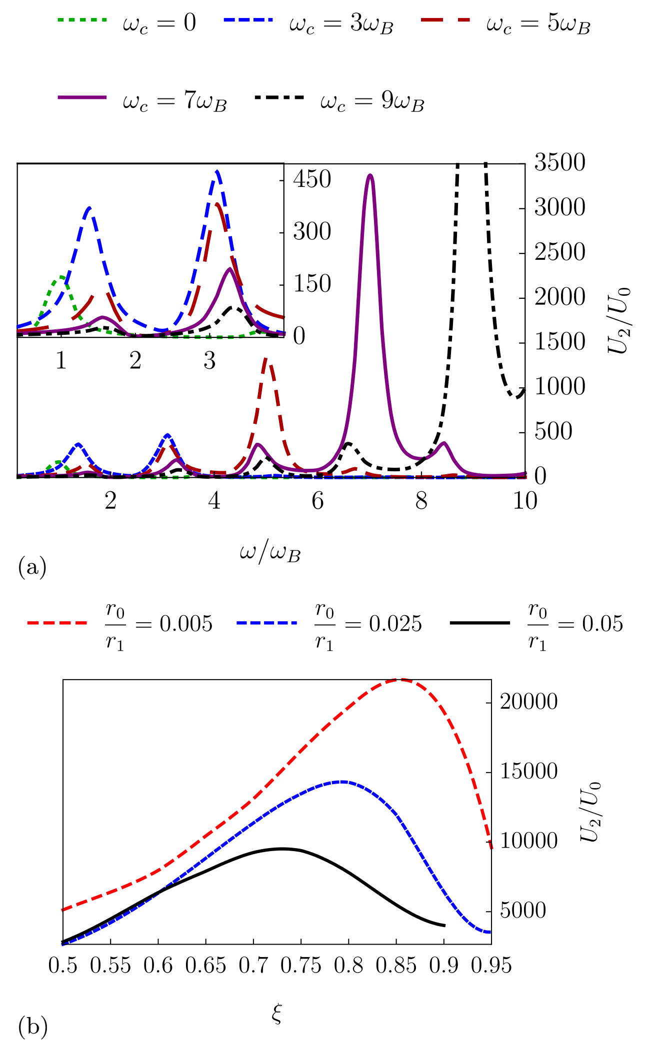


FIG. 5. Panel (a) The nonlinear potential, $U_2(\boldsymbol{r})$ integrated along the drain, obtained by numerically solving the set of Eqs. (7)-(8), plotted as a function of the frequency of the incoming radiation. Different curves correspond to different values of the magnetic field, *i.e.* of the cyclotron frequency ω_c. For all curves we have set the collision time $\tau = 10\, r_1/2s$, and the inner radius $r_0 = 0.05\, r_1/2$, and thus the eccentricity is $\xi = d/r_1 = 0.95$. Additionally $U_{\rm ext} = U_0$. The inset shows a magnification of the graph for $0 < \omega/\omega_B < 4$. Here it is evident that the peak at $\omega \approx 3\omega_B$ shifts to the right as ω_c increases. Panel (b) The strength of the peak of the nonlinear potential at the cyclotron frequency, for $\omega_c = 7\omega_B$, plotted against eccentricity for different source radii. The other parameters are the same of Panel (a).

sult in a large resonance at the cyclotron frequency that greatly enhances the responsivity of the device. Since its frequency depends on the magnetic field, exploiting such strong resonance can lead to the realization of efficient and tunable THz photodetectors. Unfortunately, this programme is difficult to be achieved in practice.

Instead, through breaking the circular symmetry of the system by placing the source off-center and closer to the edge of the disk, magnetoplasmons with various winding numbers can be excited with source-to-gate voltages easily achievable experimentally (*i.e.* uniform along the source perimeter). By tuning the degree of eccentricity of the system, we are able to excite various magnetoplasmons at once. Therefore, we are able to enhance the photodetector responsivity at the frequency range corresponding to the cyclotron one. The best protocol for photodetection clearly depends on one's aims. When searching for the frequency of incoming radiation, it is best to fix the luminosity of the radiation, where possible, and scan over a presumed range of frequencies by changing the magnetic field strength. When measuring the luminosity of incoming radiation it is best to adjust the cyclotron frequency to match the incoming radiation's frequency to achieve a high gain.

ACKNOWLEDGMENTS

A.P. acknowledges support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS) and of the Leverhulme Trust under the grant RPG-2019-363.

Appendix A: Derivation of Eqs. (7)-(8)

Plugging the expansions in Eq. (6) into the set of equations (2)-(5), we collect terms of order $U_{\rm ext}$ and $U_{\rm ext}^2$ into two systems of *linear* differential equations, *i.e.*

$$\begin{cases} \partial_t U_1(\boldsymbol{r},t) = -U_0 \nabla\cdot \boldsymbol{v}_1(\boldsymbol{r},t) \\ \dfrac{e}{m}\nabla U_1(\boldsymbol{r},t) = \partial_t \boldsymbol{v}_1(\boldsymbol{r},t) + \dfrac{1}{\tau}\boldsymbol{v}_1(\boldsymbol{r},t) + \omega_c \hat{\boldsymbol{z}}\times \boldsymbol{v}_1(\boldsymbol{r},t) \\ U_1(\boldsymbol{r},t)|_{\rm source} = U_{\rm ext}\cos(\omega t) \\ \hat{\boldsymbol{n}}\cdot \boldsymbol{v}_1(\boldsymbol{r},t)|_{\rm drain} = 0 \end{cases}, \tag{A1}$$

and

$$\begin{cases} \partial_t U_2(\boldsymbol{r},t) = -\nabla\cdot\left[U_0 \boldsymbol{v}_2(\boldsymbol{r},t) + U_1(\boldsymbol{r},t)\boldsymbol{v}_1(\boldsymbol{r},t)\right] \\ -\nabla\phi(\boldsymbol{r},t) = \partial_t \boldsymbol{v}_2(\boldsymbol{r},t) + \dfrac{1}{\tau}\boldsymbol{v}_2(\boldsymbol{r},t) + \omega_c \hat{\boldsymbol{z}}\times \boldsymbol{v}_2(\boldsymbol{r},t) \\ U_2(\boldsymbol{r},t)|_{\rm source} = 0, \\ \hat{\boldsymbol{n}}\cdot \boldsymbol{v}_2(\boldsymbol{r},t)|_{\rm drain} = 0. \end{cases}, \tag{A2}$$

respectively. Here we defined $\phi(\boldsymbol{r},t) = v_1^2(\boldsymbol{r},t)/2 - eU_2(\boldsymbol{r},t)/m$. Eqs. (A1) form a closed set of linear differential equations that can be solved exactly. Their result is then substituted into Eqs. (A2), which are themselves linear in $U_2(\boldsymbol{r},t)$ and $\boldsymbol{v}_2(\boldsymbol{r},t)$ and whose solution yields the rectified potential. The second order set of equations (A2) can be simplified further by noting that we are looking for a time-independent potential, therefore by integrating over a period of oscillation, $T = 2\pi/\omega$, the time-dependent parts of $U_2(\boldsymbol{r},t)$ and $\boldsymbol{v}_2(\boldsymbol{r},t)$ will vanish. For a generic function of time $A(t)$, we define its time-average as

$$\langle A(t)\rangle = \frac{1}{T}\int_0^T A(t)dt \ . \tag{A3}$$

After time averaging, Eq. (A2) becomes

$$\begin{cases} \nabla\cdot\left[U_0\boldsymbol{v}_2(\boldsymbol{r}) + \langle U_1(\boldsymbol{r},t)\boldsymbol{v}_1(\boldsymbol{r},t)\rangle\right] = 0 \\ \dfrac{1}{\tau}\boldsymbol{v}_2(\boldsymbol{r}) + \omega_c\hat{\boldsymbol{z}}\times\boldsymbol{v}_2(\boldsymbol{r}) = -\nabla\phi(\boldsymbol{r}) \\ U_2(\boldsymbol{r})|_{\rm source} = 0 \\ \hat{\boldsymbol{n}}\cdot\boldsymbol{v}_2(\boldsymbol{r})|_{\rm drain} = 0 \end{cases} , \tag{A4}$$

where now $\phi(\boldsymbol{r},t) = \langle v_1^2(\boldsymbol{r},t)\rangle/2 - eU_2(\boldsymbol{r})/m$, and $U_2(\boldsymbol{r})$ and $\boldsymbol{v}_2(\boldsymbol{r})$ denote the time-independent components of $U_2(\boldsymbol{r},t)$ and $\boldsymbol{v}_2(\boldsymbol{r},t)$, respectively.

We will now further simplify Eqs. (A1). We first obtain two equations by applying the operator $\partial_t + 1/\tau$ and the cross product with $\hat{\boldsymbol{z}}$ to the second of Eqs. (A1). We then combine the two equations we obtained, and get

$$\left[\left(\partial_t + \frac{1}{\tau}\right)^2 + \omega_c^2\right]\boldsymbol{v}_1(\boldsymbol{r},t) = \frac{e}{m}\Bigg[\left(\partial_t + \frac{1}{\tau}\right)\nabla U_1(\boldsymbol{r},t) - \omega_c\hat{\boldsymbol{z}}\times\nabla U_1(\boldsymbol{r},t)\Bigg]. \tag{A5}$$

The new set of equations is solved by using the *Ansatz* (see also the main text, Sect. II)

$$\begin{aligned} U_1(\boldsymbol{r},t) &= U_1(\boldsymbol{r})e^{-i\omega t} + U_1^*(\boldsymbol{r})e^{i\omega t}, \\ \boldsymbol{v}_1(\boldsymbol{r},t) &= \boldsymbol{v}_1(\boldsymbol{r})e^{-i\omega t} + \boldsymbol{v}_1^*(\boldsymbol{r})e^{i\omega t}, \end{aligned} \tag{A6}$$

from which we obtain the following set of time-independent linear equations:

$$-i\omega U_1(\boldsymbol{r}) + U_0\nabla\cdot\boldsymbol{v}_1(\boldsymbol{r}) = 0 \ , \tag{A7}$$

and

$$[\omega^2 f_\omega^2 - \omega_c^2]\boldsymbol{v}_1(\boldsymbol{r}) = \frac{e}{m}\left[i\omega f_\omega\nabla U_1(\boldsymbol{r}) + \omega_c\hat{\boldsymbol{z}}\times\nabla U_1(\boldsymbol{r})\right] \ , \tag{A8}$$

subject to the boundary conditions

$$\begin{aligned} U_1(\boldsymbol{r})|_{\rm source} &= \frac{U_{\rm ext}}{2}, \\ \hat{\boldsymbol{n}}\cdot\boldsymbol{v}_1(\boldsymbol{r})|_{\rm drain} &= 0. \end{aligned} \tag{A9}$$

In these equations we introduced $f_\omega = 1 + i/(\omega\tau)$. In addition to Eqs. (A7)-(A9), we have a set of equation for the quantities $U_1^*(\boldsymbol{r})$ and $\boldsymbol{v}_1^*(\boldsymbol{r})$. These are obtained from Eqs. (A7)-(A9) by taking their complex conjugates. Substituting Eq. (A8) into (A7), results in the following closed set of equations for $U_1(\boldsymbol{r})$:

$$\begin{cases} \left[\omega_c^2 - \omega^2 f_\omega^2\right]U_1(\boldsymbol{r}) - s^2 f_\omega\nabla^2 U_1(\boldsymbol{r}) = 0 \\ U_1(\boldsymbol{r})\Big|_{\rm source} = \dfrac{U_{\rm ext}}{2} \\ \hat{\boldsymbol{n}}\cdot\left[i\omega f_\omega\nabla U_1(\boldsymbol{r}) + \omega_c\hat{\boldsymbol{z}}\times\nabla U_1(\boldsymbol{r})\right]\Big|_{\rm drain} = 0 \end{cases} \tag{A10}$$

Here we define the plasma wave velocity, $s = \sqrt{-eU_0/m}$, where U_0, the equilibrium potential, is negative for an electron fluid. The first of Eqs. (A10) defines a Poisson problem which, once boundary conditions are specified as in the second and third of (A10), admits a unique solution. Such solution is determined analytically for the case of a concentric Corbino-disk geometry in Sect. III and numerically for an eccentric disk in Sect. IV.

Once the set of Eqs. (A10) is solved and $U_1(\boldsymbol{r})$ has been determined, the velocity is given by

$$\boldsymbol{v}_1(\boldsymbol{r}) = \frac{s^2}{U_0}\frac{i\omega f_\omega\nabla U_1(\boldsymbol{r}) + \omega_c\hat{\boldsymbol{z}}\times\nabla U_1(\boldsymbol{r})}{\omega_c^2 - \omega^2 f_\omega^2} \ . \tag{A11}$$

It is then possible to approach the problem posed by the set of Eqs. (A4) in a similar fashion. Plugging the definitions in Eqs. (A6) in there, we find

$$\begin{cases} \nabla\cdot\left[U_0\boldsymbol{v}_2(\boldsymbol{r}) + U_1^*(\boldsymbol{r})\boldsymbol{v}_1(\boldsymbol{r}) + U_1(\boldsymbol{r})\boldsymbol{v}_1^*(\boldsymbol{r})\right] = 0 \\ \dfrac{1}{\tau}\boldsymbol{v}_2(\boldsymbol{r}) + \omega_c\hat{\boldsymbol{z}}\times\boldsymbol{v}_2(\boldsymbol{r}) = -\nabla\phi(\boldsymbol{r}) \\ U_2(\boldsymbol{r})|_{\rm source} = 0 \\ \boldsymbol{v}_2(\boldsymbol{r})|_{\rm drain} = 0 \end{cases} \tag{A12}$$

where, explicitly, $\phi(\boldsymbol{r}) = \boldsymbol{v}_1^*(\boldsymbol{r})\cdot\boldsymbol{v}_1(\boldsymbol{r}) - eU_2(\boldsymbol{r})/m$. To further simplify Eq. (A12) and reduce it to a Poisson problem, we first obtain two equations by taking the divergence and applying the operator $\hat{\boldsymbol{z}}\times\nabla$ to the second of its equations. We get

$$\frac{1}{\tau}\nabla\cdot\boldsymbol{v}_2(\boldsymbol{r}) - \omega_c\hat{\boldsymbol{z}}\cdot\nabla\times\boldsymbol{v}_2(\boldsymbol{r}) = \nabla^2\phi(\boldsymbol{r}) \ , \tag{A13}$$

and

$$\frac{1}{\tau}\hat{\boldsymbol{z}}\cdot\nabla\times\boldsymbol{v}_2(\boldsymbol{r}) + \omega_c\nabla\cdot\boldsymbol{v}_2(\boldsymbol{r}) = 0 \ . \tag{A14}$$

Combining such equations with the first of Eqs. (A12) gives

$$\frac{1 + (\tau\omega_c)^2}{U_0\tau}\nabla\cdot\left[U_1^*(\boldsymbol{r})\boldsymbol{v}_1(\boldsymbol{r}) + U_1(\boldsymbol{r})\boldsymbol{v}_1^*(\boldsymbol{r})\right] = \nabla^2\phi(\boldsymbol{r}) \ . \tag{A15}$$

Eq. (A15) has the form of a Poisson equation for $\phi(\boldsymbol{r})$. Given appropriate boundary conditions, the latter can be solved and yield a unique solution for $\phi(\boldsymbol{r})$ and therefore for $U_2(\boldsymbol{r}) = m/e\big[\boldsymbol{v}_1^*(\boldsymbol{r})\cdot\boldsymbol{v}_1(\boldsymbol{r}) - \phi(\boldsymbol{r})\big]$. To determine the boundary conditions for $\phi(\boldsymbol{r})$, we first take the cross product of the second of Eqs. (A12) with $\hat{\boldsymbol{z}}$, which yields

$$\frac{1}{\tau}\hat{\boldsymbol{z}}\times\boldsymbol{v}_2(\boldsymbol{r}) - \omega_c\boldsymbol{v}_2(\boldsymbol{r}) = -\hat{\boldsymbol{z}}\times\nabla\phi(\boldsymbol{r}) \ . \tag{A16}$$

Substituting this back into the second of Eqs. (A12) we get

$$\left[1 + (\omega_c\tau)^2\right]\boldsymbol{v}_2(\boldsymbol{r}) = \omega_c\tau^2\hat{\boldsymbol{z}}\times\nabla\phi(\boldsymbol{r}) - \tau\nabla\phi(\boldsymbol{r}) \ . \tag{A17}$$

This leads us to the following solvable set of differential equations in $\phi(\boldsymbol{r})$:

$$\begin{cases} \dfrac{1+(\tau\omega_c)^2}{U_0\tau}\nabla\cdot\left[U_1^*(\boldsymbol{r})\boldsymbol{v}_1(\boldsymbol{r})+U_1(\boldsymbol{r})\boldsymbol{v}_1^*(\boldsymbol{r})\right]=\nabla^2\phi(\boldsymbol{r}) \\ \phi(\boldsymbol{r})-\boldsymbol{v}_1^*(\boldsymbol{r})\cdot\boldsymbol{v}_1(\boldsymbol{r})\Big|_{\text{source}}=0 \\ \omega_c\hat{\boldsymbol{z}}\times\nabla\phi(\boldsymbol{r})-\dfrac{1}{\tau}\nabla\phi(\boldsymbol{r})\Big|_{\text{drain}}=0 \end{cases} . \tag{A18}$$

---

[*] bailey.winstanley@manchester.ac.uk

[†] alessandro.principi@manchester.ac.uk

[1] R. Kleiner, Science **318**, 1254 (2007).

[2] P. Alonso-González, A. Nikitin, Y. Gao, A. Woessner, M. Lundeberg, A. Principi, N. Forcellini, W. Yan, S. Vélez, A. Huber, K. Watanabe, T. Taniguchi, F. Casanova, L. Hueso, M. Polini, J. Hone, F. Koppens, and R. Hillenbrand, Nature nanotechnology **12**, 31 (2016).

[3] X. Deng, L. Li, M. Enomoto, and Y. Kawano, Scientific Reports **9**, 3498 (2019).

[4] A. Grigorenko, M. Polini, and K. Novoselov, Nature Photon **6**, 749–758 (2012).

[5] F. Bonaccorso, Z. Sun, T. Hasan, and A. C. Ferrari, Nature Photonics **4**, 611–622 (2010).

[6] M. Tonouchi, Nature photonics **1**, 97 (2007).

[7] M. Dyakonov and M. Shur, IEEE Transactions on Electron Devices **43**, 380 (1996).

[8] M. I. Dyakonov and M. S. Shur, Applied Physics Letters **67**, 1137 (1995).

[9] M. I. Dyakonov and M. S. Shur, Phys. Rev. B **51**, 14341 (1995).

[10] G. Giuliani and G. Vignale, Quantum Theory of the Electron Liquid (Cambridge University Press, 2005).

[11] W. Knap, J. Lusakowski, T. Parenty, S. Bollaert, A. Cappy, V. V. Popov, and M. S. Shur, Applied Physics Letters **84**, 2331 (2004).

[12] N. Dyakonova, A. El Fatimy, J. Łusakowski, W. Knap, M. I. Dyakonov, M.-A. Poisson, E. Morvan, S. Bollaert, A. Shchepetov, Y. Roelens, C. Gaquiere, D. Theron, and A. Cappy, Applied Physics Letters **88**, 141906 (2006).

[13] S. Kang, P. J. Burke, L. N. Pfeiffer, and K. W. West, Applied Physics Letters **89**, 213512 (2006).

[14] W. Knap, M. Dyakonov, D. Coquillat, F. Teppe, N. Dyakanova, J. Lusakowski, K. Karpierz, M. Sakowicz, G. Valusis, D. Seliuta, I. Kasalynas, A. E. Fatimy, Y. M. Meziani, and T. Otsuji, J Infrared Milli Terahz Waves **30**, 1319 (2009).

[15] A. Tomadin and M. Polini, Phys. Rev. B **88**, 205426 (2013).

[16] H. Yan, Z. Li, X. Li, W. Zhu, P. Avouris, and F. Xia, Nano Letters **12**, 3766 (2012).

[17] L. Vicarelli, M. S. Vitiello, D. Coquillat, A. Lombardo, A. C. Ferrari, W. Knap, M. Polini, V. Pellegrini, and A. Tredicucci, Nature Materials **11**, 865–871 (2012).

[18] A. Principi, D. Bandurin, H. Rostami, and M. Polini, Phys. Rev. B **99**, 075410 (2019).

[19] F. Koppens, T. Mueller, P. Avouris, A. Ferrari, M. Vitiello, and M. Polini, Nature nanotechnology **9**, 14 (2014).

[20] M. Khavronin, A. Petrov, A. Kazantsev, E. Nikulin, and D. Bandurin, Phys. Rev. Applied **13**, 064072 (2020).

[21] S. Boubanga-Tombet, M. Sakowicz, D. Coquillat, F. Teppe, W. Knap, M. I. Dyakonov, K. Karpierz, J. Łusakowski, and M. Grynberg, Applied Physics Letters **95**, 072106 (2009), https://doi.org/10.1063/1.3207886.

[22] M. Białek, A. M. Witowski, M. Orlita, M. Potemski, M. Czapkiewicz, J. Wróbel, V. Umansky, M. Grynberg, and J. Łusakowski, Journal of Applied Physics **115**, 214503 (2014), https://doi.org/10.1063/1.4881777.

[23] D. C. Glattli, E. Y. Andrei, G. Deville, J. Poitrenaud, and F. I. B. Williams, Phys. Rev. Lett. **54**, 1710 (1985).

[24] A. L. Fetter, Phys. Rev. B **33**, 5221 (1986).

[25] F. A. Reboredo and C. R. Proetto, Phys. Rev. B **53**, 12617 (1996).

[26] D. Bandurin, D. Svintsov, I. Gayduchenko, S. Xu, A. Principi, M. Moskotin, I. Tretyakov, D. Yagodkin, S. Zhukov, T. Taniguchi, K. Watanabe, I. Grigorieva, M. Polini, G. Goltsman, A. Geim, and G. Fedorov, Nature Communications **9** (2018), 10.1038/s41467-018-07848-w.

[27] Z. Sun, D. N. Basov, and M. M. Fogler, Proceedings of the National Academy of Sciences **115**, 3285 (2018), https://www.pnas.org/content/115/13/3285.full.pdf.

[28] H. Rostami, M. I. Katsnelson, and M. Polini, Phys. Rev. B **95**, 035416 (2017).

[29] S. A. Mikhailov, Phys. Rev. B **95**, 085432 (2017).

[30] S. A. Mikhailov, Phys. Rev. B **93**, 085403 (2016).

[31] J. L. Cheng, N. Vermeulen, and J. Sipe, Scientific Reports **7**, 43843 (2017).

[32] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Phys. Rev. B **91**, 235320 (2015).

[33] A. Principi, M. I. Katsnelson, and G. Vignale, Phys. Rev. Lett. **117**, 196803 (2016).

[34] A. Principi, G. Vignale, M. Carrega, and M. Polini, Phys. Rev. B **88**, 121405(R) (2013).

[35] A. L. Fetter, Phys. Rev. B **32**, 7676 (1985).

[36] H. Marmanis, Physics of Fluids **10**, 1428 (1998).

[37] D. Spirito, D. Coquillat, S. L. De Bonis, A. Lombardo, M. Bruna, A. C. Ferrari, V. Pellegrini, A. Tredicucci, W. Knap, and M. S. Vitiello, Applied Physics Letters **104**, 061111 (2014).

[38] H. Qin, J. Sun, S. Liang, X. Li, X. Yang, Z. He, C. Yu, and Z. Feng, Carbon **116** (2017), 10.1016/j.carbon.2017.02.037.

[39] S. O. Potashin, V. Y. Kachorovskii, and M. S. Shur, "Hydrodynamic inverse faraday effect in two dimensional electron liquid," (2020), arXiv:2001.08015 [cond-mat.mes-

hall].

[40] R. C. Ashoori, H. L. Stormer, L. N. Pfeiffer, K. W. Baldwin, and K. West, Phys. Rev. B **45**, 3894 (1992).

Second harmonic generation as a minimal model of turbulence

N. Vladimirova[1,2], M. Shavit[2], S. Belan[3,4], and G. Falkovich[2]
[1] *Brown University, Providence, RI 02912, USA*
[2] *Weizmann Institute of Science, Rehovot 76100 Israel*
[3] *Landau Institute for Theoretical Physics, 142432 Chernogolovka, Russia and*
[4] *National Research University Higher School of Economics, 101000 Moscow, Russia*
(Dated: March 30, 2021)

When two resonantly interacting modes are in contact with a thermostat, their statistics is exactly Gaussian and the modes are statistically independent despite strong interaction. Considering noise-driven system, we show that when one mode is pumped and another dissipates, the statistics (of such cascades) is never close to Gaussian no matter the interaction/noise relation. One finds substantial phase correlation in the limit of strong interaction (weak noise). Surprisingly, for both cascades, the mutual information between modes increases and entropy further decreases when interaction strength decreases. We use the model to elucidate the fundamental problem of far-from equilibrium physics: where the information (entropy deficit) is encoded and how singular measures form. For an instability-driven system (a laser), even a small added noise leads to large fluctuations of the relative phase near the stability threshold, while far from it we show that the conversion into the second harmonic is weakly affected by noise.

I. INTRODUCTION

Second harmonic generation is the simplest fundamental process of nonlinear wave physics, which is also in the center of numerous practical applications in laser physics and beyond. The dynamics of the process has been studied exhaustively [1, 2], which cannot be said about statistics, despite the fact that understanding the influence of noise on the energy conversion is of paramount practical importance, recently enhanced by the use of metamaterials [3]. Here we address this problem by studying theoretically a two-mode resonant system driven by a combination of pumping and random noise. Our motivation is two-fold. Apart from the classical conversion problem, we find this system ideally suited for elucidating the fundamental problems of non-equilibrium physics. When one mode is stochastically forced and another is dissipated, that presents a minimal model of turbulence cascade. The freedom to force either mode allows us to elucidate the basic differences between direct and inverse cascades. Apart from energy, we shall be interested in the entropy of such far-from equilibrium state, which is expected to be much lower than in thermal equilibrium with the same energy.

A remarkable property, common for all systems of resonantly interacting waves and shared with hydrodynamic systems [4], is that the canonical thermal equilibrium has exactly Gaussian statistics, and the modes fluctuate independently, regardless of the interaction strength. Here we describe how deviations from equilibrium diminish entropy and build correlations between the two modes. Far from equilibrium the joint two-mode statistics is never close to Gaussian, even when the marginal distribution of every mode is close to Gaussian. On the one hand, the entropy decrease means that the statistical distribution is getting more non-uniform, which poses the question: Can it lead all the way to singularity like the celebrated Sinai-Ruelle-Bowen measures in dynamical systems [5, 6]? We show that this is indeed so: the measure in the phase space is getting singular in the double limit of strong non-equilibrium and weak interaction. On the other hand, since entropy is missing information, any entropy decrease poses another question: Where all this extra information about non-equilibrium is encoded? First, we find out how the entropies of the three marginal distributions, of each mode amplitude and their phase difference, go down as the system deviates from equilibrium. Second, we find out which part of the entropy decrease is due to inter-mode correlation. This is properly measured by the mutual information (rather than by the pair correlation function, suitable for Gaussian statistics only).

The process of the second harmonic generation is described by the following model Hamiltonian (assuming perfect resonance)

$$\mathcal{H}_0 = \omega|a_1|^2 + 2\omega|a_2|^2 + V a_1^{*2} a_2 + V^* a_1^2 a_2^* \,. \qquad (1)$$

Here a_1 and a_2 are the complex amplitudes of two nonlinearly coupled modes having frequencies ω and 2ω, respectively, and V is the interaction constant (considered real positive without loss of generality). The two coupled complex equations govern dynamics: $\dot{a}_k = -i\partial\mathcal{H}_0/\partial a_k^*$, $k = 1, 2$. We eliminate the linear terms in these equations by introducing the envelopes

$$b_1 = a_1 e^{i\omega t}, b_2 = a_2 e^{2i\omega t}. \qquad (2)$$

That results in a strongly interacting system with a cubic Hamiltonian $\mathcal{H} = V b_1^{*2} b_2 + V^* b_1^2 b_2^*$.

Due to the symmetry $b_1 \to b_1 e^{i\phi}$, $b_2 \to b_2 e^{2i\phi}$ the system $\dot{b}_k = -i\partial\mathcal{H}/\partial b_k^*$ has an extra integral of motion $N = |b_1|^2 + 2|b_2|^2$ and is completely integrable; the phase portrait is presented in Appendix A 1. Let us add dissipation and stochastic pumping:

$$\dot{b}_1 = -2iV^* b_1^* b_2 - \gamma_1 b_1 + \xi_1(t), \qquad (3)$$

$$\dot{b}_2 = -iV b_1^2 - \gamma_2 b_2 + \xi_2(t). \qquad (4)$$

Here γ_1 and γ_2 are the damping coefficients, and ξ_1 and ξ_2 are independent Gaussian random forces with zero mean $\langle\xi_i(t)\rangle = 0$ and the variance $\langle\xi_i(t_1)\xi_j^*(t_2)\rangle = P_i\delta_{ij}\delta(t_2 - t_1)$.

We mainly focus on the properties of the statistically steady solutions of the system (3,4) in the case when one mode is forced, while the other is damped. Since the modes enter the Hamiltonian in a non-symmetric way, there are two possibilities: one either can pump the first (lower frequency) mode and dump the second (higher frequency) mode or vice versa. The former scenarios qualitatively corresponds to the direct energy cascade, while the second is reminiscent to the inverse cascade.

We wish to understand how much information is needed in order to build a turbulent state and how much one learns about one mode by observing another. For that we will use the metrics from information theory: entropy and mutual information. The answer to the first question is given by the decrease in entropies

$$\begin{aligned} S_{12} &= -\int db_1 db_1^* db_2 db_2^*\, \rho(b_1, b_2) \ln \rho(b_1, b_2)\,, \\ S_1 &= -\int db_1 db_1^*\, \rho(b_1) \ln \rho(b_1)\,, \\ S_2 &= -\int db_2 db_2^*\, \rho(b_2) \ln \rho(b_2)\,, \end{aligned}$$

where ρ is either full or marginal probability distribution. The answer to the second question is given by the mutual information between the modes:

$$I_{12} = S_1 + S_2 - S_{12}\ . \tag{5}$$

Fig.1 demonstrates the growth of the mutual information versus the degree of non-equilibrium (an analog of the Reynolds number defined below, see (8)).

As one of the simplest model of energy transfer, the system of two coupled oscillators has received considerable attention in the literature [7–15]. In particular, in the mathematical literature, one finds an analysis of a two-mode system with a quadratic Hamiltonian $\mathcal{H} = Ta_1^{*2}a_2^2$ with the purpose to get insight into the energy transfer in wave turbulence [14, 15]. What distinguishes our model is that it directly corresponds to physical reality and allows experimental validation. In addition, an asymmetry between the modes allows us to compare direct and inverse cascades, which turn out quite different. Another distinction is that we add entropic and informational consideration to the energetic analysis.

The paper is organized from the point of view of entropy: we start from the maximal-entropy equilibrium and investigate near-equilibrium states in Section II. We then move to study the noise-driven direct and inverse cascades in Section III. We define a dimensionless parameter akin to the Reynolds number and show how entropy decreases as the Reynolds number increases. We begin with the limit of small Reynolds number in III A and continue to large Reynolds number in III B, where entropy decreases all the way down as the measure becomes singular in the limit of infinite Reynolds number. In Section IV we consider an instability-driven first harmonic and study the influence of noise on the conversion process, this can serve as a simple model for a laser generating second harmonic. Conclusion V briefly lists our main results.

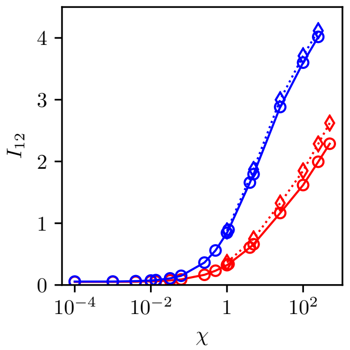


FIG. 1. Mutual information versus Reynolds number for direct and inverse cascades (red and blue lines respectively). Three dimensional distribution are computed with bin size $\Delta\theta = 2\pi/32$ and $\Delta\rho_{1,2}^2/n_{1,2} = 1$ (circles) and 0.5 (diamonds).

II. NEAR THERMAL EQUILIBRIUM

Adopting the language of stochastic thermodynamics, one can call the ratios $P_1/\gamma_1 \equiv T_1$ and $2P_2/\gamma_2 \equiv T_2$ effective temperatures experienced by two modes b_1 and b_2 which are governed by the Langevin equations (3) and (4). If $\Delta T = T_1 - T_2 = 0$, then it is straightforward to find from the Fokker-Planck equation or entropy maximum the steady-state joint probability distribution:

$$\mathcal{P}_0 = \frac{1}{Z}\exp\left(-\frac{2|b_1|^2 + 4|b_2|^2}{T}\right) = \frac{1}{Z}\exp\left(-\frac{2N}{T}\right). \tag{6}$$

Despite strong interaction, this distribution is exactly Gaussian and the modes are statistically independent. The later means that the mean energy flux and the mutual information between modes are zero. Thermal equilibrium corresponds to the equipartition of the quadratic invariant: $\langle|b_1|^2\rangle \equiv n_1 = 2n_2 \equiv 2\langle|b_2|^2\rangle$.

What can we say about the system's statistics when modes are subject to different effective temperatures? Let us introduce the dimensionless measure of non-equilibrium

$$\sigma = \frac{\Delta T}{T}, \quad \text{where} \quad T = \frac{P_1 + 2P_2}{2(\gamma_1 + \gamma_2)}. \tag{7}$$

Another dimensionless parameter quantifies interaction strength relative to the dissipation:

$$\chi = \frac{(\gamma_1 + \gamma_2)^3}{(P_1 + 2P_2)|V|^2}\ . \tag{8}$$

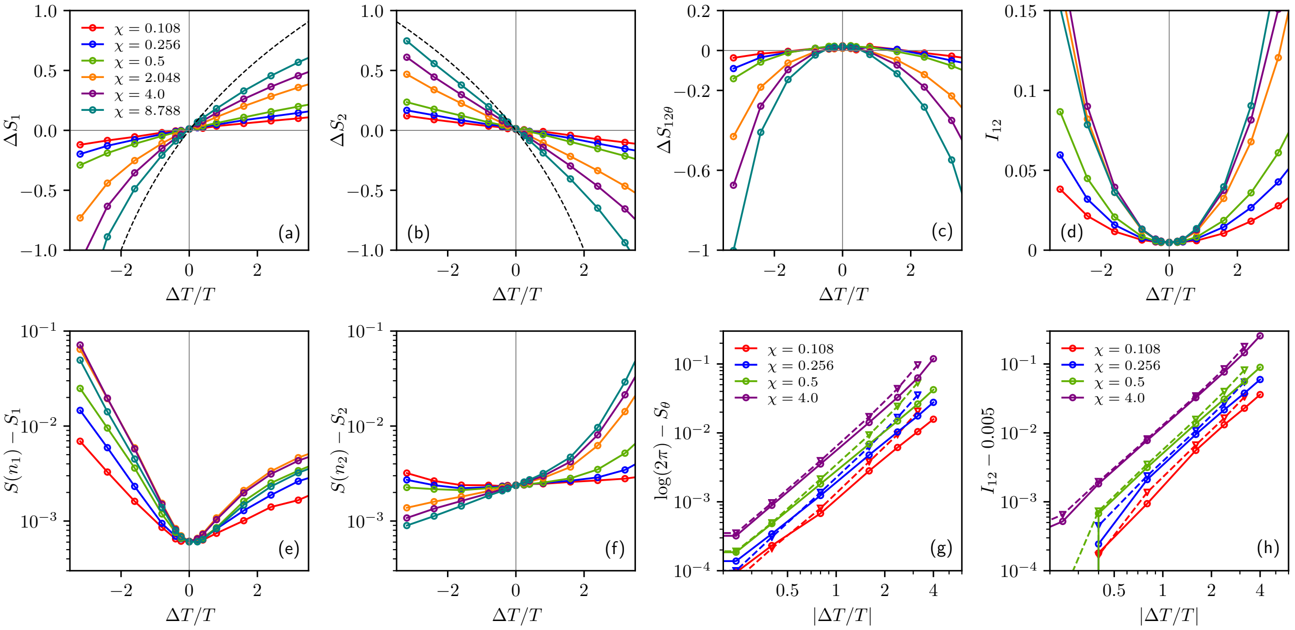


FIG. 2. Top row: deviation of entropies from equilibrium ($\Delta T = 0$) and mutual information for $\gamma_1 = \gamma_2$ and different strength of interaction marked by color. Equilibrium is Gaussian with $S_{\mathrm{eq}1} = 1/\ln(2)$, $S_{\mathrm{eq}2} = 1/\ln(2) - 1$, and $S_{\mathrm{eq}\theta} = \log_2(2\pi)$. Dashed lines show the entropy deviation of a single mode respectively at T_1 and T_2 for marginal distributions, $\Delta S_1 = \log_2(1 + \Delta T/4T)$ and $\Delta S_2 = \log_2(1 - \Delta T/4T)$. The entropies of individual modes are affected by the change of variance of (nearly Gaussian) distribution and by deviation from the Gaussian shape. For close-to-equilibrium cases presented here the first effect is dominant. Bottom row: panels (e) and (f) show deviation of entropies from the entropies of Gaussian distributions with the same variance; panel (g) illustrates the quadratic dependence of the entropy of phase distribution, panel (h) illustrates the quadratic dependence for mutual information (solid lines for $\Delta T > 0$; dashed lines for $\Delta T < 0$). The offset 0.005 is the effect of finite ensemble and bin size; the offset decreases as the size of data set and resolution improve. 3D distributions are computed with bin sizes $2\pi/32$ for the phase and $0.1T$ for $\rho_{1,2}$.

Denote $\rho_{1,2} = |b_{1,2}|$ and $\theta = \arg(b_1^2 b_2^*)$. From (3) and (4), the steady-state equations on the second moments read:

$$-4V\langle\rho_1^2\rho_2\sin\theta\rangle - 2\gamma_1\langle\rho_1^2\rangle + P_1 = 0, \tag{9}$$

$$2V\langle\rho_1^2\rho_2\sin\theta\rangle - 2\gamma_2\langle\rho_2^2\rangle + P_2 = 0, \tag{10}$$

The time derivative of the real part of the third moment is given by: $d\langle\mathcal{H}\rangle/dt = -(2\gamma_1+\gamma_2)\langle\mathcal{H}\rangle$, since $\langle\xi_1 b_1 b_2^*\rangle = \langle\xi_2^* b_1^2\rangle = 0$. Therefore, in any steady state, either in thermal equilibrium or out of it, one has

$$\langle\mathcal{H}\rangle = 2V\langle\rho_1^2\rho_2\cos\theta\rangle = 0. \tag{11}$$

Equations (9)-(11) are valid for any values of σ and χ.

At $\sigma \neq 0$, the probability density $\mathcal{P}(b_1, b_2)$ is non-Gaussian in non-equilibrium, yet it is close to Gaussian when $|\sigma| \ll 1$ for all values of χ. The simplest to treat is the limit of small interaction, $\chi \gg 1$. In this case, the first correction to Eq. (6) is determined by the energy flux between modes, which is small and proportional to the temperature difference:

$$\ln\mathcal{P}(b_1,b_2) \approx -\frac{2|b_1|^2}{T_1} - \frac{4|b_2|^2}{T_2} - \tag{12}$$

$$-\frac{4\Delta T}{P_1T_2+P_2T_1}\,\mathrm{Im}\,[V^* b_1^{*2} b_2] + O(\chi^{-2}) \tag{13}$$

Smallness of interaction multiplies the parameter of non-equilibrium $\Delta T/T$ in the right hand side of Eq. (12), so that this result is valid even when $\Delta T/T$ is not small. That means that, as long as both temperatures remain finite and interaction is weak, even far from equilibrium the relative entropy is small:

$$D(\mathcal{P}|\mathcal{P}_0) = \int db_1 db_1^* db_2 db_2^* \mathcal{P}\ln(\mathcal{P}/\mathcal{P}_0) \propto \chi^{-1} \ll 1\ ,$$

as well as the mutual information.

In the opposite limit $\chi \ll 1$ or $V \to \infty$, the non-Gaussian correction is again proportional to the product of the degree of non-equilibrium and the small parameter χ. In terms of $x = |b_1|^2$ and $y = 2|b_2|^2$ we obtain:

$$\ln\mathcal{P} \approx -\frac{x+y}{T} + \frac{\chi^{1/2}\Delta T}{T} f(x,y,\theta)\ ,$$

where the correction satisfies the equation

$$\frac{2xy}{\sqrt{yT}}\left[\sin\theta\left(\frac{\partial}{\partial y}-\frac{\partial}{\partial x}\right)+\frac{x-2y}{2xy}\cos\theta\frac{\partial}{\partial\theta}\right]f=x-y\,.$$

The correction is odd in phase difference, $f(-\theta)=-f(\theta)$, and scales linearly with amplitudes, so that it is substantial at small amplitudes. In the limits, $f\to g(\theta)\sqrt{y/2T}$, where $g=\sin\theta$ at $y\ll x$, and $g=\int d\theta/\cos\theta$ at $y\gg x$.

It makes sense to compare entropies at the same mean quadratic energy N. To see how entropy goes down on the way to turbulence we shall subtract the total entropy from its maximal equilibrium value, which quantifies the amount of information one needs to create a turbulent state: $\Delta S(N,n_2/n_1)=S_0-S_{12}$.

Numerics support quadratic decrease of $S_{12}(\Delta/T)$ and increase of $I_{12}(\Delta/T)$ up to $\Delta T\simeq 4T$, see Figure 2.

When $\Delta T/T$ exceeds one, the functions are not even which demonstrates the statistical difference between upward and downward energy conversion. We see stronger deviations from Gaussianity for negative $\Delta T<0$, which corresponds to the downward energy flow and to an inverse cascade at $\Delta T/T\to-\infty$. The physical difference is that the first mode pumps the second one as an additive force, while the second mode pumps the first one as a multiplicative instability. Therefore, it seems natural that the entropy is generally lower and the mutual information higher for an inverse transfer. The analysis of the separate distributions of two amplitudes and the relative phase shows that the entropy of the driven mode (say, $S_1(\Delta/T)$ for a direct transfer) grows with ΔT slower than the entropy of the dissipated mode and $S_\theta(\Delta/T)$ decrease, see panels (g) and (h) in Figure 2.

III. TURBULENT CASCADES

Now let us have an energy cascade in our model: pumping one mode and dissipating another. When energy flows from lower frequency mode to higher, i.e $0=\xi_2=\gamma_1$ in Eqs. (3) and (4), the cascade is called direct, and inverse when $0=\xi_1=\gamma_2$. In these cases, the only dimensionless parameter is $\chi=\gamma^3/(P|V|^2)$, where P is the intensity of noise acting upon the driving mode, and γ denotes the damping coefficient of the dissipating mode. As we shall see below, χ to some extent plays the role of the Reynolds number of hydrodynamics in a sense that it determines how low is the entropy and how much the occupation numbers deviate from the equipartition $n_1=2n_2$, even though the system is not close to thermal equilibrium for however small or large χ.

Balance of the quadratic invariant, N, means that the dissipating mode keeps the magnitude of order of its equilibrium value: $n_2=P/4\gamma$ for the direct cascade and $n_1=P/\gamma$ for the inverse cascade. How much the mode which is pumped exceeds the equipartition value is determined by the value of χ, as described below. Note that this parameter can be interpreted as the squared ratio of the dissipation rate γ and the nonlinear transfer rate $Vn\simeq V\sqrt{P/\gamma}$.

When χ is small, the interaction between modes is strong and the energy transfer is fast, so that the occupation numbers are expected to be close to equipartition, yet the statistics is not expected to be close to separable Gaussian form given by (6). Even though the noise is weak, it is white, that is a singular perturbation destroying integrability everywhere in the phase space [16]; we shall see below how non-trivial the probability distribution is already in this limit.

One may naively expect that in the opposite limit of large χ, when the noise is strong and interaction is weak, the correlation between modes would be weak too. We shall show below that the opposite is true far from equilibrium: the necessity to carry the flux makes the modes strongly correlated precisely because of a strong noise and weak interaction. It is in this limit we find the lowest entropy and the maximal mutual information between modes, as well as appearance of singular measure in phase space.

A. Small Reynolds number: strong-interaction-weak-noise limit

1. Inverse cascade

In an inverse cascade, energy goes from high to low frequency, so we set $\gamma_2=0$ and $P_1=0$ in (3,4):

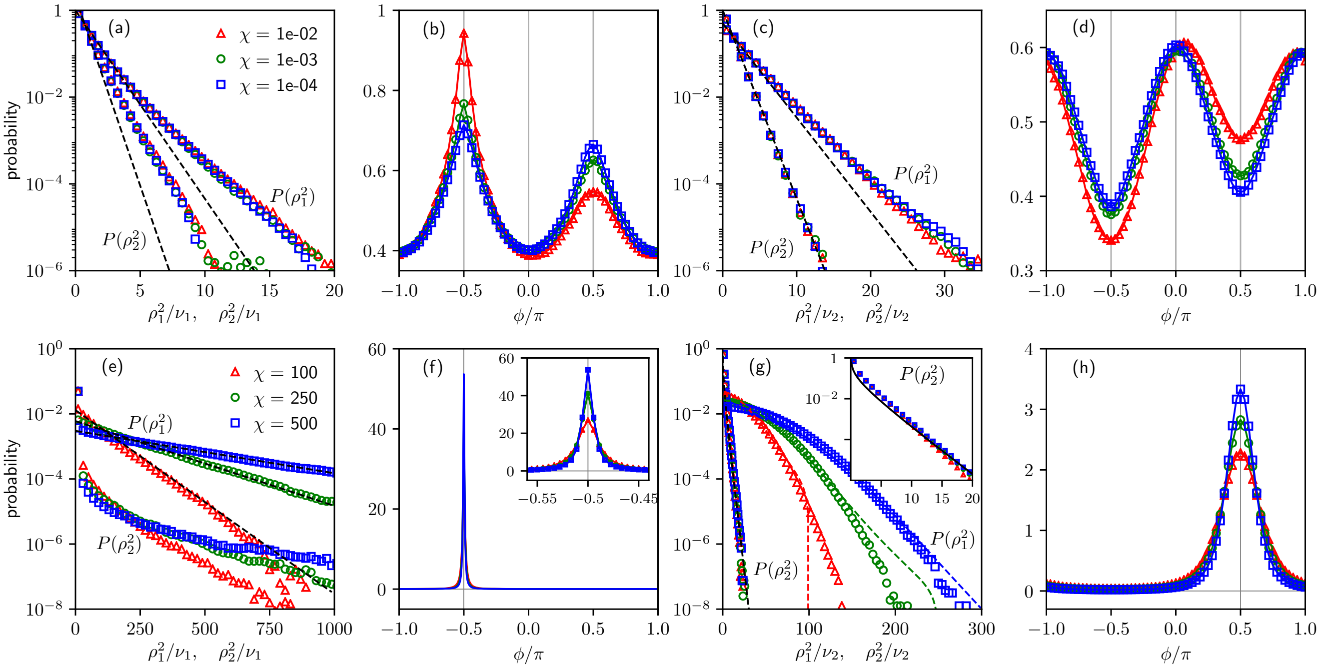
$$\dot b_1=-2iV^*b_1^*b_2-\gamma b_1, \tag{14}$$

$$\dot b_2=-iVb_1^2+\xi(t). \tag{15}$$

In the steady state, the energy input rate P must be equal to the dissipation rate γn_1 and to the energy flux from the second mode to the first given by the imaginary part of the third cumulant: $2V\langle\rho_1^2\rho_2\sin\theta\rangle=-P$. Then, from the energy balance we obtain $n_1=P/\gamma$ and $2V\langle\rho_1^2\rho_2\sin\theta\rangle=-P$, so that $\langle\rho_1^2\rho_2\rangle\geq P/2V$. Also, from the condition $\frac{d}{dt}\langle\ln|b_1|^2\rangle=0$ we find $2V\langle\rho_2\sin\theta\rangle=-\gamma$ and, therefore $\langle\rho_2\rangle\geq\gamma/2V$.

From Eqs. (14,15) it is straightforward to see that when $\chi\ll 1$ the steady-state probability distribution $\mathcal{P}(b_1,b_2)$ cannot be close to the equilibrium Gaussian (6) with the temperature $T=2P/\gamma$ and the equipartition $P/\gamma=n_1=2n_2$. Indeed, the stationarity of $\langle\mathcal{H}^2\rangle=4|V|^2\langle\rho_1^4\rho_2^2\cos^2\theta\rangle$ requires $\langle\mathcal{H}^2\rangle/\langle\rho_1^4\rangle=\frac{|V|^2P}{2\gamma}$, but this contradicts the Gaussian ratio which gives $\langle\mathcal{H}^2\rangle/\langle\rho_1^4\rangle=\frac{|V|^2P}{\gamma}$.

Thus, small value of χ does not mean that the system is near equipartition. In agreement with this conclusion, results of numerical modeling presented at two lower left panels of Figure 3 show that at neither of marginal distributions of the mode amplitudes is close to Gaussian and that the phase distribution does not become uniform as χ tends to zero. This is also reflected in nonzero



.

FIG. 3. Probabilities of the occupation numbers and the phase for: (a, b) inverse cascade, small χ; (c, d) direct cascade, small χ; (e, f) inverse cascade, large χ; (g, h) direct cascade, large χ. Each dataset contains 20M datapoints, at $\Delta t = 0.01$ for inverse cascade and large χ and $\Delta t = 0.1$ for all other cases. For the inverse cascade, $\chi = \gamma_1^3/(2P_2V^2)$ and $\nu_1 = P_2/\gamma_1$. We use $\gamma_{1,2} = 0.01$ for small χ and $\gamma_{1,2} = 1$ for large χ. In all cases, $V = 1$. Broken lines in (g) correspond to the approximation (31,36).

value $I_{12}(+0)$ of mutual information between modes in this limit, see the right panel of Figure 1.

2. *Direct cascade*

Direct cascade corresponds to the choice $\gamma_1 = 0$ and $P_2 = 0$ in (3,4):

$$\dot{b}_1 = -2iV^* b_1^* b_2 + \xi(t), \tag{16}$$

$$\dot{b}_2 = -iV b_1^2 - \gamma b_2. \tag{17}$$

Now energy goes from low to high frequency. Again, in the steady-state regime, the energy input rate $P/2$ must be equal to the dissipation rate $2\gamma n_2$ and to the energy flux from the first mode to the second given by the imaginary part of the third cumulant: $2V\langle\rho_1^2\rho_2 \sin\theta\rangle = P/2$, so $n_2 = P/4\gamma$. From $\frac{d}{dt}\langle \ln|b_2|^2\rangle = 0$ we find $\langle \rho_1^2 \sin\theta/\rho_2\rangle = \gamma/V$. Therefore $\langle\rho_1^2\rho_2\rangle \geq P/4V$ and $\langle\rho_1^2/\rho_2\rangle \geq \gamma/V$.

When $\chi \to 0$, the dimensionless flux $\langle\rho_1^2\rho_2\sin\theta\rangle/n_1 n_2^{1/2} = \chi^{1/2}$ is small, which may suggest that phase-space distribution is close to the Gaussian equilibrium (6) with $T = P/\gamma$ and that the phase distribution is close to uniform. Furthermore, as opposed to the case of inverse cascade discussed above, the equality obtained from the stationarity of $\langle\mathcal{H}^2\rangle$:

$$\langle\mathcal{H}^2\rangle = 4|V|^2\langle\rho_1^4\rho_2^2\cos^2\theta\rangle = \frac{2|V|^2P}{\gamma}\langle\rho_1^2\rho_2^2\rangle\ , \tag{18}$$

is achieved by the Gaussian distribution with $n_1 = 2n_2 = P/2\gamma$. However, numerical data, as can be seen from two upper left panels of Figure 3, shows that even though the marginal distributions of amplitudes are close to Gaussian with equipartition, $n_1 \approx 2n_2$, the phase distribution is far from flat and deviation from equilibrium is substantial. The mutual information between modes as a function of χ exhibits a non-zero value of $I_{12}(+0)$ (see Figure 1) which is also a clear footprint of non-equilibrium.

B. Large Reynolds number: weak interaction, strong noise limit

1. Inverse cascade

The pair of complex equations (14,15) can be rewritten as three real ones since the overall phase drops out:

$$\dot{\rho}_1 = -2|V|\rho_1\rho_2 \sin\theta - \gamma\rho_1, \tag{19}$$

$$\dot{\rho}_2 = |V|\rho_1^2 \sin\theta + \frac{P}{4\rho_2} + \frac{\zeta(t)}{\sqrt{2}}, \tag{20}$$

$$\dot{\theta} = |V|\frac{\rho_1^2 - 4\rho_2^2}{\rho_2}\cos\theta + \frac{\zeta(t)}{\sqrt{2}\rho_2}, \tag{21}$$

where $\zeta(t)$ is the real white noise with zero mean $\langle\zeta(t)\rangle = 0$ and the pair correlation function $\langle\zeta(t_1)\zeta(t_2)\rangle = P\delta(t_1 - t_2)$.

When $\chi \gg 1$, Eqs. (19)-(21), can be further simplified by assuming that relative phase is locked on $\theta = -\pi/2$ most the time. Then, one gets the following closed equations for the amplitudes dynamics

$$\dot{\rho}_1 = 2|V|\rho_1\rho_2 - \gamma\rho_1, \tag{22}$$

$$\dot{\rho}_2 = -|V|\rho_1^2 + \frac{P}{4\rho_2} + \frac{\zeta(t)}{\sqrt{2}}. \tag{23}$$

A hypothesis that the modes are statistically independent in this limit is shown incorrect in the Appendix A 4. This result is in sharp contrast with the model described in [14], where authors found the factorized joint probability density $\mathcal{P}(\rho_1, \rho_2)$ of mode amplitudes in the limit when their analogue of the parameter χ is large.

While constructing the probability densities for inverse cascade at $\chi \to \infty$ turns out to be a tricky task, it is straightforward to describe general features of stochastic dynamics dictated by Eqs. (22) and (23). Namely, these pair of nonlinearly coupled equations suggest the following cyclical evolution: ρ_1 stays close to zero most of the time while ρ_2 undergoes diffusion in a repulsive logarithmic potential; when ρ_2 sufficiently outgrows the threshold level $\gamma/2|V|$, ρ_1 shoots up and quickly diminishes ρ_2; after that ρ_1 also resets to the near-zero level and the stochastic dynamics of ρ_2 starts from scratch. The mode dynamics during the intermittent burst events can be described by simplified equations

$$\dot{\rho}_1 = 2|V|\rho_1\rho_2 - \gamma\rho_1, \tag{24}$$

$$\dot{\rho}_2 = -|V|\rho_1^2. \tag{25}$$

Compared with Eqs. (22) and (23), we neglected the terms associated with noise. Equations (24) and (25) are exactly solvable yielding

$$\rho_1^2(t) = r_1^2 - 2\left(\rho_2(t) - \frac{\gamma}{2V}\right)^2 + 2\left(r_2 - \frac{\gamma}{2|V|}\right)^2, \tag{26}$$

$$\rho_2(t) = \frac{\gamma}{2V} + \sqrt{\left(r_2 - \frac{\gamma}{2V}\right)^2 + \frac{r_1^2}{2}}\tanh\left[\frac{1}{2}\ln\frac{\sqrt{\left(r_2 - \frac{\gamma}{2V}\right)^2 + \frac{r_1^2}{2}} + r_2 - \frac{\gamma}{2V}}{\sqrt{\left(r_2 - \frac{\gamma}{2V}\right)^2 + \frac{r_1^2}{2}} - r_2 + \frac{\gamma}{2V}} - 2Vt\sqrt{\left(r_2 - \frac{\gamma}{2V}\right)^2 + \frac{r_1^2}{2}}\right], \tag{27}$$

where $r_1 = \rho_1(0)$, $r_2 = \rho_2(0)$ are the initial conditions. Estimating $r_2 \sim \gamma/V$ and $r_1 \ll r_2$, we see from Eq. (27) that the duration of such burst event is $\sim \gamma^{-1}$, which is much smaller than the typical inter-events period $\sim \gamma^2/PV^2$. As follows from Eq. (26), the amplitude of the dissipated mode grows from the initial value $r_1 \ll r_2$ to the maximum value $\rho_{1\max} = \sqrt{r_1^2 + 2(r_2 - \frac{\gamma}{2V})^2} \approx \sqrt{2}(r_2 - \frac{\gamma}{2V})$ (attaining it at the moment when $\rho_2(t) = \frac{\gamma}{2V}$) and finally returns to the starting level $\rho_1(\infty) = r_1$. Such bursts are likely responsible for pulses running in shell models, which are chains of interacting triplets.

Numerical simulations confirm the intermittent nature of system dynamics described above. Namely, panel (f) of Figure 3 reveals that the relative phase is indeed locked at $-\pi/2$. Figure 4 illustrates that analytical prediction based on the assumption of phase locking (see Eqs. (26) and (27)) are in excellent agreement with numerical data extracted from simulations of Eqs. (14) and (15). As visible in panel (e) of Figure 3, the tails of the amplitudes probability distributions strongly depend on χ; the fits of $\mathcal{P}(\rho_1)$ and $\mathcal{P}(\rho_2)$ by the broken lines in Figure 3 are empirical. As for the mutual information, from Figure 1 we see that $I_{12} \propto \ln\chi$ for inverse cascade with $\chi \gg 1$.

2. Direct cascade

We conclude the treatment of our turbulent cascades with the case of a direct cascade in the limit of large Reynolds number, $\chi \to \infty$. It is in this limit we find a window to the way a singular measure is formed far away from equilibrium. We indeed find that in this limit the full probability distribution is singular and thus corresponds to the lowest entropy state $S_{12} \to -\infty$.

In addressing the weak-interaction-strong-noise limit $\gamma \to \infty$ ($\chi \to \infty$), it is convenient to express b_2 from (17) as an integral, which in the leading order shows that the

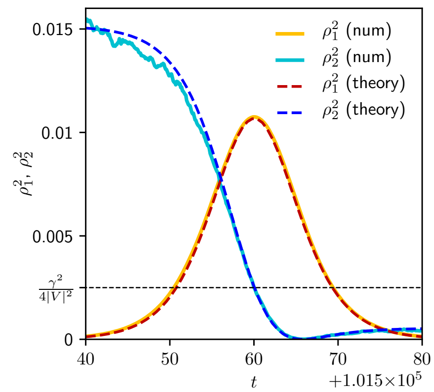


FIG. 4. An individual realization of the modes trajectories during one of the events. The continues lines are obtained from numerical simulations of Eqs. (14) and (15, and the dashed lines represent theoretical fit based on Eqs. (26) and (27)).

second amplitude is enslaved to the first one:

$$b_2(t) = -iV \int_{-\infty}^{t} b_1^2(t') e^{\gamma(t'-t)} dt' \approx -\frac{iVb_1^2}{\gamma}. \tag{28}$$

Substituting this relation into (16), we get a closed equation on the stochastic dynamics of b_1

$$\dot{b}_1 = -\frac{2V^2}{\gamma} b_1^* b_1^2 + \xi(t), \tag{29}$$

from which one finds the following expressions for the marginal probability distributions

$$\mathcal{P}(b_1) = Z_1^{-1} \exp\left(-\frac{2V^2}{\gamma P}|b_1|^4\right), \tag{30}$$

$$\mathcal{P}(b_2) = Z_2^{-1} \exp\left(-\frac{2\gamma}{P}|b_2|^2\right), \tag{31}$$

which are valid at $\rho_1, \rho_2 \ll \gamma/|V|$. Thus, the whole probability density in the four-dimensional phase space is singular at $\chi \to \infty$, sitting on a three-dimensional manifold

$$\mathcal{P}(b_1, b_2) = \frac{2}{\sqrt{\pi P\gamma}} \exp\left(-\frac{4V^2|b_1|^4}{P\gamma}\right) \delta(b_2 - \frac{iV}{\gamma} b_1^2), \tag{32}$$

so that the total entropy $S_{12} \to -\infty$. Note also that Eq. (32) yields large ratio of the typical mode amplitudes: $\rho_1^2/\rho_2^2 \simeq \sqrt{\gamma^3/P|V|^2} = \sqrt{\chi} \gg 1$.

Since the distribution over the overall phase is flat, one can integrate it out and conclude that the distribution in the the three-dimensional space of variables ρ_1, ρ_2, θ concentrates on the curve $\rho_2 \propto \rho_1^2$. Interestingly, with increasing χ the joint probability distribution $\mathcal{P}(b_1, b_2)$ is getting sharper than Gaussian along this curve. This is different from the model discussed in [14] where the driving mode is nearly Gaussian in this limit, the relative amplitudes of both modes fluctuate, and only the relative phase is fixed, that is the joint probability density is only singular with respect to the phase difference θ.

Considering large but finite χ, instead of delta-function in Eq. (32) one obtains the distribution with a finite width which is the variance of the difference $b_2 - Vb_1^2/\gamma$. To estimate this width we further expand Eq. (28)

$$b_2(t) \approx -iV \int_{-\infty}^{t} \left[b_1^2(t) + (t'-t)\frac{db_1^2(t)}{dt}\right] e^{\gamma(t'-t)} dt' \tag{33}$$

$$= -\frac{iVb_1^2(t)}{\gamma} + \frac{2iVb_1(t)}{\gamma^2}\frac{db_1(t)}{dt}. \tag{34}$$

From Eq. (33) we get

$$\langle |b_2 + iVb_1^2/\gamma|^2 \rangle = \frac{4|V|^2}{\gamma^4} \langle |b_1 \dot{b}_1|^2 \rangle = \frac{4|V|^2 P^2}{\gamma^4}\ . \tag{35}$$

Dividing this result by $\langle |b_2|^2 \rangle = P/\gamma$, one obtains that the relative squared width behaves as $1/\chi$ (as expected, it tends to zero when $\chi \to \infty$). The non-zero width at finite values of χ entails the finite entropy of the distribution $\mathcal{P}(b_1, b_2)$: $S_{12} \simeq -\ln \chi$. This analytical prediction is supported by Figure 1. For comparison, the naive Gaussian ansatz yields $S_1 + S_2 \simeq \ln n_1 n_2 \propto \ln \chi^{-1/2}$, since $n_2 = P/4\gamma$ and $n_1 \simeq \sqrt{P\gamma}/|V|$.

Now let us plug Eq. (28) into Eq.(17), then solving the corresponding steady Fokker-Planck equation (see A 3) one arrives at the next order correction for the marginal probability distribution for $\rho_1 \ll \gamma/|V|$:

$$\mathcal{P}(\rho_1) \propto \rho_1 (1 - \frac{4|V|^2}{\gamma^2}\rho_1^2) \exp(-\frac{2|V|^2}{\gamma P}\rho_1^4 + \frac{16|V|^4}{3\gamma^3 P}\rho_1^6)\ . \tag{36}$$

which is more accurate than Eq. (30). Unfortunately, extracting similar correction for the probability density $\mathcal{P}(\rho_2)$ as well as the further high order corrections to $\mathcal{P}(\rho_1)$ is more challenging.

As can be seen from panel (g) of Figure 3, Eqs. (31) and (36) allow us to fit the numerical data quite accurately in the range $\rho_1, \rho_2 \ll \gamma/|V|$. Expectedly, the agreement between numeric and analytic results improves with the growth of χ.

IV. LASER GENERATION

Let us now pump the first harmonic by an instability, for instance, in a laser. Consider first pure dynamics, setting $\xi_1 = \xi_2 = 0$ in (3,4) and changing sign in front of γ_1, which now describes gain for an optical signal. Then the resulting evolution satisfies three closed equations:

$$\frac{d\rho_1^2}{dt} = -2F + 2\gamma_1\rho_1^2, \tag{37}$$

$$\frac{d\rho_2^2}{dt} = F - 2\gamma_2\rho_2^2, \tag{38}$$

$$\frac{dF}{dt} = (2\gamma_1 - \gamma_2)F + 2|V|^2\big(\rho_1^4 - 4\rho_1^2\rho_2^2\big), \tag{39}$$

where $F = 2V\rho_1^2\rho_2\sin\theta$ is the instantaneous flux.

Apart from the trivial unstable fixed point $\rho_1 = \rho_2 = 0$, Eqs. (37)-(39) have the stationary point $\bar{\rho}_1^2 = \gamma_1\gamma_2/2|V|^2$, $\bar{\rho}_2^2 = \gamma_1^2/4|V|^2$, $\bar{\theta} = \pi/2$ (and, thus, $\bar{F} = \gamma_2\gamma_1^2/2|V|^2$). This means that in the the degenerate case $\gamma_2 = 2\gamma_1$, the system possesses the steady state $\rho_1 = 2\rho_2$ existing for any θ. This marginal stability turns into an instability of the steady state at $\gamma_2 < 2\gamma_1$ and into stability at $\gamma_2 > 2\gamma_1$. In what follows, we consider $\gamma_2 > 2\gamma_1$. Note that in practice both γ_1 and γ_2 often depend on the amplitudes, for instance, due to gain saturation or/and nonlinear damping. However, our main focus here is on the noise impact on the steady state, so we will treat γ_1, γ_2 taken near this state as constants.

Let us now add a random pumping and study its influence on the efficiency and statistics of conversion. The modes amplitudes, ρ_1 and ρ_2 and the relative phase θ are governed by the following equations

$$\dot{\rho}_1 = -2V\rho_1\rho_2\sin\theta + \gamma_1\rho_1 + \frac{P}{4\rho_1} + \frac{\zeta_1(t)}{\sqrt{2}}, \tag{40}$$

$$\dot{\rho}_2 = V\rho_1^2\sin\theta - \gamma_2\rho_2\,, \tag{41}$$

$$\dot{\theta} = \frac{\rho_1^2 - 4\rho_2^2}{\rho_2}V\cos\theta + \frac{\sqrt{2}\zeta_2(t)}{\rho_1}. \tag{42}$$

Here ζ_1 and ζ_2 are two independent real white noises with zero mean values $\langle\zeta_i(t)\rangle = 0$, and the pair correlator $\langle\zeta_i(t_1)\zeta_j(t_2)\rangle = P\delta_{ij}\delta(t_1 - t_2)$. In the limit of weak noise, $P|V|^2 \ll \gamma_1^2(\gamma_2 - 2\gamma_1)$, one can apply a linear approximation near the fixed point. More specifically, we substitute decomposition $\rho_1(t) = \bar{\rho}_1 + u(t), \rho_2(t) = \bar{\rho}_2 + v(t), \theta(t) = \bar{\theta} + \phi(t)$ into Eqs. (41)-(42) and keep only the first order terms with respect to u, v and ϕ. This procedure yields

$$\dot{u} = -2V\bar{\rho}_1 v + \frac{P}{4\bar{\rho}_1} + \frac{\zeta_1(t)}{\sqrt{2}}, \tag{43}$$

$$\dot{v} = 2V\bar{\rho}_1 u - \gamma_2 v, \tag{44}$$

$$\dot{\phi} = (2\gamma_1 - \gamma_2)\phi + \frac{\sqrt{2}\zeta_2(t)}{\bar{\rho}_1}. \tag{45}$$

From Eqs. (43)-(45) we immediately find the variances

$$\langle u^2\rangle - \langle u\rangle^2 = \frac{(2\gamma_1 + \gamma_2)\,P}{8\gamma_1\gamma_2}\,, \tag{46}$$

$$\langle v^2\rangle - \langle v\rangle^2 = \frac{P}{4\gamma_2}\,, \tag{47}$$

$$\langle \phi^2\rangle - \langle \phi\rangle^2 = \frac{2V^2P}{\gamma_1\gamma_2\,(\gamma_2 - 2\gamma_1)}. \tag{48}$$

We see that level of fluctuations in relative phase grows when one approaches the stability threshold. Note also that far from the threshold, i.e. at $2\gamma_1 \ll \gamma_2$, the fluctuations of the second harmonic are suppressed: $\langle v^2\rangle/\langle u^2\rangle \approx \gamma_1/\gamma_2 \ll 1$. In this case, the noise of the first harmonic only weakly influences the conversion into the second one. However, the conversion is least effective in this limit: $\bar{\rho}_2^2/\bar{\rho}_1^2 = \gamma_1/2\gamma_2 \ll 1$.

V. CONCLUSION

Our most important finding is the explicit formula (32) for the singular measure of a direct cascade in the limit of strong noise and weak interaction. We believe that this is a meaningful advance in non-equilibrium statistics, as it opens a window to the study of the formation of singular measures in systems driven far away from equilibrium. We have described also the approach to this limit and have shown that the total entropy decays and the inter-mode mutual information grows logarithmically with the Reynolds number. In the inverse cascade case in this limit, the phase is locked on $-\pi/2$ and the system exhibits an intermittent dynamics of bursts, which we were able to describe analytically. Such bursts are perhaps responsible for pulses running in shell models, which are chains of interacting triplets used in modeling hydrodynamic incompressible turbulence. It is thus may be interesting to apply the methods developed here to the popular shell model $\dot{u}_i = u_{i-1}^2 - u_i u_{i+1}$ [17]. After some elementary transformations, this model can be turned into that with the Hamiltonian $\mathcal{H} = \sum_i V_i(a_i^2 a_{i+1}^* + c.c)$, that is the interacting chain built of our pairs.

The opposite limit of weak noise and strong interaction is a singular one: the probability distribution is not close to a Gaussian distribution determined by N for however small χ, despite occupation numbers being close to equipartition and the marginal one mode distributions close to quadratic. It is expressed, in particular, in the nonzero mutual information $I_{12}(\chi)$ at the limit $\chi \to +0$. Figure 1 combines the mutual information data for both cascades. We see that $I_{12}(\chi) - I_{12}(+0) \propto \chi^2$ at $\chi \ll 1$. We failed to find an analytic solution in this limit either in the direct or inverse cascade, even though it is likely that the probability distribution can be expressed in terms of N and $\mathcal{H}^2$, which are the conserved quantities of the unforced undamped system.

We conclude with suggesting an interesting application of our model to wave turbulence. In a set of $M + 2$

interacting waves, one may consider to model the interaction of a resonant couple with the other M waves by dissipation and random forcing. When $M \gg 1$ we can treat forces from all other modes as a white noise, so that our model (4) applies. In this case, different limits in χ correspond to different situations. If we assume an almost continuous distribution of other modes and estimate from the wave kinetic equation $\gamma \simeq V^2 Mn/\omega$ and $P \simeq \gamma n$ [18], then $\chi = \gamma^3/PV^2 \simeq V^2 Mn/\omega^2 \ll 1$, which is the original parameter of nonlinearity assumed to be small. In this case, we come to the surprising conclusion that a resonant mode within turbulence, when $\Delta T \simeq T$, has a relative entropy of order unity and independent of V. If, however, we have a set of well-isolated resonant interactions, then it makes more sense to assume that the interaction with a given mode is M times smaller so that χ is large (as M or $\sqrt{M}$), then the relative entropy is small. Note that in most cases the number of resonant interactions, is much less than the total number of modes in the system.

The work was supported by the Scientific Excellence Center and Ariane de Rothschild Women Doctoral Program at WIS, grant 662962 of the Simons foundation, grant 075-15-2019-1893 by the Russian Ministry of Science, grant 873028 of the EU Horizon 2020 programme, and grants of ISF, BSF and Minerva. NV was in part supported by NSF grant number DMS-1814619. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF grant number ACI-1548562, allocation DMS-140028.

---

[1] PD Drummond, KJ McNeil, and DF Walls. Nonequilibrium transitions in sub/second harmonic generation. Optica Acta: International Journal of Optics, 27(3):321–335, 1980.

[2] Paul Mandel and Thomas Erneux. Amplitude self-modulation of intracavity second-harmonic generation. Optica Acta: International Journal of Optics, 29(1):7–21, 1982.

[3] Matthias W Klein, Christian Enkrich, Martin Wegener, and Stefan Linden. Second-harmonic generation from magnetic metamaterials. Science, 313(5786):502–504, 2006.

[4] Natalia Vladimirova, Michal Shavit, and Gregory Falkovich. Fibonacci turbulence. arXiv preprint arXiv:2101.10418, 2021.

[5] Jay Robert Dorfman. An introduction to chaos in nonequilibrium statistical mechanics. Number 14. Cambridge University Press, 1999.

[6] Michal Shavit and Gregory Falkovich. Singular measures and information capacity of turbulent cascades. Physical Review Letters, 125(10):104501, 2020.

[7] K Vijay Kumar, Sriram Ramaswamy, and Madan Rao. Active elastic dimers: Self-propulsion and current reversal on a featureless track. Physical Review E, 77(2):020102, 2008.

[8] Federico Bonetto, Joel L Lebowitz, and Jani Lukkarinen. Fourier's law for a harmonic crystal with self-consistent stochastic reservoirs. Journal of statistical physics, 116(1-4):783–813, 2004.

[9] Sergio Ciliberto, Alberto Imparato, Antoine Naert, and Marius Tanase. Heat flux and entropy produced by thermal fluctuations. Physical review letters, 110(18):180601, 2013.

[10] Gianmaria Falasco, Marco Baiesi, Leo Molinaro, Livia Conti, and Fulvio Baldovin. Energy repartition for a harmonic chain with local reservoirs. Physical Review E, 92(2):022129, 2015.

[11] Hyun-Myung Chun and Jae Dong Noh. Hidden entropy production by fast variables. Physical Review E, 91(5):052128, 2015.

[12] Federica Mura, Grzegorz Gradziuk, and Chase P Broedersz. Nonequilibrium scaling behavior in driven soft biological assemblies. Physical review letters, 121(3):038002, 2018.

[13] Junang Li, Jordan M Horowitz, Todd R Gingrich, and Nikta Fakhri. Quantifying dissipation using fluctuating currents. Nature communications, 10(1):1–9, 2019.

[14] Paul A Milewski, Esteban G Tabak, and Eric Vanden-Eijnden. Resonant wave interaction with random forcing and dissipation. Studies in Applied Mathematics, 108(1):123–144, 2002.

[15] RE Lee DeVille, Paul A Milewski, Ricardo J Pignol, Esteban G Tabak, and Eric Vanden-Eijnden. Nonequilibrium statistics of a reduced model for energy transfer in waves. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 60(3):439–461, 2007.

[16] Khanh-Dang Nguyen Thu Lam and Jorge Kurchan. Stochastic perturbation of integrable systems: a window to weakly chaotic systems. Journal of Statistical Physics, 156(4):619–646, 2014.

[17] VN Desnyansky and EA Novikov. Modelling of the cascade processes in a turbulent flows. Prikl. Mat. Mekh, 38(3):507, 1974.

[18] Vladimir E Zakharov, Victor S L'vov, and Gregory Falkovich. Kolmogorov spectra of turbulence I: Wave turbulence. Springer Science & Business Media, 2012.

[19] Natalia Vladimirova, Stanislav Derevyanko, and Gregory Falkovich. Phase transitions in wave turbulence. Physical Review E, 85(1):010101, 2012.

[20] G Falkovich, I Kolokolov, V Lebedev, and A Migdal. Instantons and intermittency. Physical Review E, 54(5):4896, 1996.

[21] Gregory Falkovich and Vladimir Lebedev. Single-point velocity distribution in turbulence. Physical review letters, 79(21):4159, 1997.

[22] AM Obukhov. Some general properties of equations describing the dynamics of the atmosphere(atmosphere hydrodynamic simulation model for cascade energy transfer in turbulent flow, using euler gyro equations). Academy of Sciences, USSR, Izvestiya, Atmospheric and Oceanic Physics, 7:471–475, 1971.

[23] AJ Bray. Random walks in logarithmic and power-law potentials, nonuniversal persistence, and vortex dynamics in the two-dimensional xy model. Physical Review E, 62(1):103, 2000.

[24] Artem Ryabov, Ekaterina Berestneva, and Viktor Holubec. Brownian motion in time-dependent logarithmic potential: Exact results for dynamics and first-passage properties. The Journal of chemical physics, 143(11):114117, 2015.
[25] Martin R Evans and Satya N Majumdar. Diffusion with stochastic resetting. Physical review letters, 106(16):160601, 2011.

Appendix A: Appendix

1. Hamiltonian evolution

Here we briefly sketch some of the elementary properties of the Hamiltonian system defined by Eq. (1). From Eq. (1) we find the system of two coupled complex equations

$$\dot{a}_1 = -i\frac{\partial \mathcal{H}}{\partial a_1^*} = -i\omega a_1 - 2iVa_1^* a_2, \tag{A1}$$

$$\dot{a}_2 = -i\frac{\partial \mathcal{H}}{\partial a_2^*} = -2i\omega a_2 - iV^* a_1^2. \tag{A2}$$

It is easy to see that in addition to $\mathcal{H}$, Eqs. (A1) and (A2) have the second integral of motion,

$$N = \omega|a_1|^2 + 2\omega|a_2|^2, \tag{A3}$$

and, thus, the system is completely integrable. Indeed, two integrals of motion allows one to reduce Eqs. (A1) and (A2) to a single first-order equation, which we write for $x = 2|a_2|^2/N$ and $t \to tVN^{1/2}$

$$\frac{dx}{dt} = \pm 2\sqrt{2x(1-x)^2 - 4K^2/N^3}\ , \tag{A4}$$

where

$$K = \mathcal{H} - N = Va_1^{*2}a_2 + V^* a_1^2 a_2^*, \tag{A5}$$

is also an integral of motion.

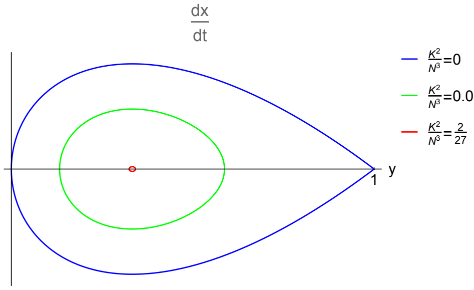


FIG. 5. The phase portrait of the integrable Hamiltonian dynamics plotted for different values of the dimensionless ratio of integrals of motion, K^2/N^3.

Next, using the Euler representation, $a_1 = \rho_1 e^{i\varphi_1}$ and $a_2 = \rho_2 e^{i\varphi_2}$, one obtains from Eqs. (A1) and (A2)

$$\dot{\rho}_1 = -2|V|\rho_1\rho_2 \sin\theta\,, \tag{A6}$$

$$\dot{\rho}_2 = |V|\rho_1^2 \sin\theta\,, \tag{A7}$$

$$\dot{\theta} = \frac{\rho_1^2 - 4\rho_2^2}{\rho_2}\cos\theta, \tag{A8}$$

where $\theta = \arg a_1^2 a_2^* = 2\phi_1 - \phi_2$. From Eqs. (A4) - (A8) one finds that the dynamical system has two fixed points: 1) $x = 1 = 2\rho_2^2/N$, $K = 0$, which means $\rho_1 = 0$ and $\theta = \pm\pi/2$, and 2) $x = 1/3$, $K^2/N^3 = 2/27$, which means $\rho_1^2 = 4\rho_2^2 = 2N/3$ and $\theta = 0$ or $\theta = \pi$. It is easy to show that the first one is unstable, while the second one is stable. The small oscillations near the second point are harmonic with the frequency $4\sqrt{2}/3$. We will see below that this phase portrait explains qualitatively the statistics of system in the presence of dumping and (small) random forcing.

Since $K = 2V\rho_1^2\rho_2\cos\theta$ is a constant, it cannot change sign, an, therefore, we have two separated regions in phase space that correspond to the sign of $\cos\theta$: $[-\pi/2, \pi/2]$ and $[\pi/2, 3\pi/2]$. Both separating planes at $\theta = \pi/2, 3\pi/2$ are critical points of ϕ, on which non-linear interaction is zero.

2. Small-flux limit

This limit can be called alternatively either small-flux limit, because the mean value of the inter-mode energy flux $2\langle\rho_1^2\rho_2\sin\theta\rangle$ is much smaller than its periodic oscillation, or small-noise limit, because most of the time the evolution is unaffected by pumping and damping. The phase portrait described in Section A 1 explains qualitatively the statistics of turbulence in this limit of large wave amplitudes, presented in the upper row of Figure 3. Indeed, the dynamical system has two fixed points: the unstable one $\rho_1 = 0$, $\theta = \pm\pi/2$, and the stable one $\rho_1^2 = 4\rho$, $\theta = 0, \pi$. In this limit, the system spends most of its time close to one of the fixed points, randomly switching between them. In the direct cascade, described in Sect III A 2, random noise acts on the first mode, so that the system spends less time around ρ_1 close to zero, and the probability has minima at $\theta = \pi/2, -\pi/2$. In the direct cascade, the system spends more time oscillating around the second fixed point and the probability has maxima at $\theta = 0, \pi$. On the contrary, dissipation acts on ρ_1 in the inverse cascade, which keeps system longer around the first fixed point, and the probability has maxima at $\theta = \pi/2, -\pi/2$.

3. Large-flux limit for the direct cascade

Here we briefly discuss the derivation of the solution Eq. (36) for the marginal probability distribution for ρ_1 in the main text.

Let us plug Eq. (28) into Eq. (17) to get:

$$\dot{b}_1 = -\frac{2|V|^2}{\gamma}\frac{|b_1|^2 b_1}{1-\frac{4|V|^2}{\gamma^2}|b_1|^2} + \frac{\xi(t)}{1-\frac{4|V|^2}{\gamma^2}|b_1|^2}. \qquad \text{(A9)}$$

The steady-state probability density of the amplitude ρ_1 obeys the Fokker-Planck equation

$$\frac{2|V|^2}{\gamma}\frac{1}{\rho_1}\frac{\partial}{\partial\rho_1}\left[\frac{\rho_1^3}{1-\frac{4|V|^2}{\gamma^2}\rho_1^2}\mathcal{P}(\rho_1)\right] + \qquad \text{(A10)}$$

$$+\frac{P}{4\rho_1}\frac{\partial}{\partial\rho_1}\left[\rho_1\frac{\partial}{\partial\rho_1}\left[\frac{\rho_1^{-1}}{(1-\frac{4|V|^2}{\gamma^2}\rho_1^2)^2}\mathcal{P}(\rho_1)\right]\right] = 0 \qquad \text{(A11)}$$

Solving this equation one arrives at Eq. (36) in the main text.

4. Large-flux limit for the inverse cascade

The stationary Fokker-Planck equation on $\mathcal{P}(\rho_1,\rho_2)$ is as follows

$$\frac{P}{4}\frac{\partial^2\mathcal{P}}{\partial\rho_2^2}+\frac{\partial}{\partial\rho_2}\left[|V|\rho_1^2-\frac{P}{4\rho_2}\right]\mathcal{P}+\frac{\partial}{\partial\rho_1}\left[\gamma\rho_1-2|V|\rho_1\rho_2\right]\mathcal{P}=0\,. \qquad \text{(A12)}$$

From Eq. (23) we obtain

$$\langle\rho_2\rangle = \frac{\gamma}{2|V|}, \qquad \text{(A13)}$$

$$\langle\rho_1^n\rangle = \frac{2|V|}{\gamma}\langle\rho_1^n\rho_2\rangle, \qquad \text{(A14)}$$

$$\langle\rho_1^2\rangle = \frac{P}{4|V|}\langle\frac{1}{\rho_2}\rangle, \qquad \text{(A15)}$$

Multiplying FPE (A12) by ρ_2^2 and integrating over $d\rho_1 d\rho_2$ yields

$$\langle\rho_1^2\rho_2\rangle = \frac{P}{2|V|}, \qquad \text{(A16)}$$

and therefore (due to Eqs. (A14), (A15) and (A16))

$$\langle\rho_1^2\rangle = \frac{P}{\gamma}, \qquad \text{(A17)}$$

$$\langle\frac{1}{\rho_2}\rangle = \frac{4|V|}{\gamma}. \qquad \text{(A18)}$$

Also from Eq. (A12) we find

$$\langle\rho_2^n\rangle = \frac{4|V|}{(n+2)P}\langle\rho_1^2\rho_2^{n+1}\rangle, \quad \langle\rho_1^2\rho_2^2\rangle = \frac{3}{8}\frac{\gamma P}{|V|^2}. \qquad \text{(A19)}$$

It follows from Eqs. (A13), (A14), (A16) and (A17) that

$$\langle\rho_1\rho_2\rangle = \langle\rho_1\rangle\langle\rho_2\rangle, \quad \langle\rho_1^2\rho_2\rangle = \langle\rho_1^2\rangle\langle\rho_2\rangle, \qquad \text{(A20)}$$

which may lead one to hypothesize that in the steady state the random variables ρ_1 and ρ_2 are statistically independent. If such statistical independence was true, then

$$\mathcal{P}(\rho_1,\rho_2) = C\rho_1^{-1+\frac{4P|V|^2}{\gamma^3}} e^{-\frac{2|V|^2\rho_1^2}{\gamma^2}}\rho_2 e^{-\frac{4|V|}{\gamma}\rho_2}. \qquad \text{(A21)}$$

However, direct substitution of Eq. (A21) into the Fokker-Planck equation (A12) shows that this distribution represents the solution only along two lines: $\rho_2 = (\frac{1}{2}+\frac{1}{2\sqrt{2}})\frac{\gamma}{|V|}$ and $\rho_1 = \sqrt{\frac{P}{\gamma}}$. Thus, the hypothesis of statistical independence is not self-consistent.

Singular measures and information capacity of turbulent cascades

Gregory Falkovich and Michal Shavit
Weizmann Institute of Science, Rehovot 76100 Israel

Is there really such a thing as weak turbulence? Here we analyze turbulence of weakly interacting waves using the tools of information theory. It offers a unique perspective for comparing thermal equilibrium and turbulence: the mutual information between modes is shown to be stationary and small in equilibrium but grows linearly with time in weak turbulence. We trace this growth to the concentration of probability on the resonance surfaces, which can go all the way to a singular measure. The surprising conclusion is that no matter how small is the nonlinearity and how close to Gaussian is the statistics of any single amplitude, a stationary phase-space measure is far from Gaussian, as manifested by a large relative entropy. Though it might be upsetting to practitioners of weak turbulence approach, this is a rare piece of good news for turbulence modeling: the resolved scales carry significant information about the unresolved scales. The mutual information between large and small scales is the information capacity of turbulent cascade, setting the limit on the representation of subgrid scales in turbulence modeling.

There are two quite different perspectives to look at the evolution of a statistical system: fluid mechanics and information theory. The first one is the continuum viewpoint, where a Hamiltonian evolution of an ensemble is treated as an incompressible flow in a phase space. Such flows generally mix which leads to a uniform microcanonical equilibrium distribution. On the contrary, to deviate a system from equilibrium, one needs external forces and dissipation that break Hamiltonian conservative nature of evolution and lead to compressible flows in phase space, which generally produce extremely non-uniform measures [1, 2]. The second perspective is the discrete viewpoint of information theory, where the evolution towards equilibrium and entropy saturation are described as the loss of all the information except integrals of motion. On the contrary, to keep a system away from equilibrium, we need to act, producing information and decreasing entropy.

Here we make a step in synthesis of the two approaches, asking: what is the informational manifestation of nonuniform turbulent measures? Such measures are expected to have a low entropy whose limit is set by an interplay between interaction on the one hand and discreteness, coarse-graining or finite resolution on the other hand. We shall look at turbulence from the viewpoint of the mutual information (MI), which measures effective correlations between different degrees of freedom.

To keep a system away from equilibrium, environment extracts entropy thus producing information — where is this information encoded? Here we consider turbulent systems which can be treated perturbatively as long as their statistics is close to Gaussian, such as weak wave turbulence (similar approach can be applied to a passive scalar [3] and other systems). We show that the MI between wave modes is encoded in cumulants (not described by the traditional description in terms of occupation numbers [2]). The information production builds higher and higher correlations which concentrate sharper and sharper on the resonant surfaces, driving the distribution towards a singular measure. When nonlinearity is small, we show that the entropy decay is due to the triple moment concentrating on the three-wave resonance surface, see Figure 1. It is unclear yet how to describe the long-time asymptotic of the entropy decay. When turbulence is driven by a random force, which provides for a phase-space diffusion and smears singularities, the entropy must saturate at a finite value, but the difference with Gaussian random-phase approximation can be large when the Reynolds number is large.

Consider a wave system defined by random complex amplitudes $a_k = |a_k|e^{i\phi}$ which satisfy $i\dot{a}_k = \delta\mathcal{H}/\delta a_k^*$ with the Hamiltonian:

$$\mathcal{H} = \sum_{\mathbf{k}} \omega_k |a_k|^2 + \sum_{\mathbf{kpq}} \frac{1}{2} \int \left(V_{kpq} a_k^* a_p a_q + \text{c.c.}\right) \delta^{\mathbf{k}}_{\mathbf{p+q}} \ .$$

Here c.c means the complex conjugated terms, $\delta^{\mathbf{k}}_{\mathbf{p+q}}$ is the Kronecker delta and we use the shorthand notation $a_k \equiv a(\mathbf{k})$, etc. The medium is assumed scale invariant, that is both ω_k and V are homogeneous functions of degree α and m, respectively. A central problem is to describe the evolution of the phase-space distribution $\rho\left(\{a_k, a_k^*\}\right)$. We assume that the second term in the Hamiltonian is on average much smaller than the first one and that the modes are independently distributed at $t = 0$. Then the occupation numbers $n_k \delta^{\mathbf{k}}_{\mathbf{k}'} := \langle a_k a_k^* \rangle$ satisfy a closed kinetic equation [2, 5–10]:

$$\frac{dn_k}{dt} = \sum_{\mathbf{k}_1\mathbf{k}_2} \mathrm{Im}\,(V_{k12}J_{k12} - 2V^*_{1k2}J^*_{1k2})\delta^{\mathbf{k}_1}_{\mathbf{k}_2+\mathbf{k}} \,, \qquad (1)$$

$$J_{123}(t) = \frac{e^{i\omega^1_{2,3}t} - 1}{\omega^1_{2,3}} V^*_{123}(n_2n_3 - n_1n_2 - n_1n_3) \,. \qquad (2)$$

The brackets $\langle f \rangle$ indicate averaging with respect to ρ, $\langle a_i^* a_j a_k \rangle = J_{ijk}\delta^{\mathbf{i}}_{\mathbf{j+k}}$ is the third moment and $\omega^1_{2,3} \equiv \omega_1 - \omega_2 - \omega_3$. Substituting (2) into (1) gives the collision integral of the kinetic equation, which is a direct analog

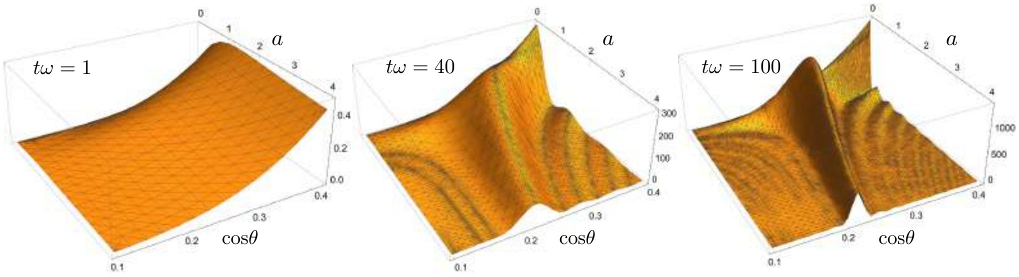


FIG. 1: Temporal build-up of the mutual information between three capillary waves in turbulence. Here $a = k_2/k_3$, $\cos\theta_{23} = \mathbf{k_1} \cdot \mathbf{k_2}/k_1k_2$. For fixed $\mathbf{k_2}$ and $\mathbf{k_1} = \mathbf{k_2} + \mathbf{k_3}$, the resonant surface is the line where the probability and MI peak at $t\omega_2 >> 1$.

of the Boltzmann equation for dilute gases:

$$\frac{dn_k}{dt} \equiv St_k = \sum_{\mathbf{k_1k_2}} \left(U_{k12} - 2U_{1k2}\right) \tag{3}$$
$$U_{ijk} = \pi|V_{ijk}|^2 \delta\left(\omega^i_{jk}\right) \delta^{\mathbf{i}}_{\mathbf{j+k}}(n_jn_k - n_in_k - n_in_j)\,.$$

The nonlinear interaction time $t_{NL}(k) \simeq n_k/St_k$ is assumed large relative to the wave period, so that the nonlinearity parameter $\epsilon_k^2 = 1/\omega_k t_{NL}(k)$ is small. It is known that the Boltzmann kinetic equation is the first term of a regular cluster expansion only at thermal equilibrium, while even weak non-equilibrium leads to singularities as manifested already in the density expansion of kinetic coefficients like viscosity, diffusivity and thermal conductivity [11–16]. The kinetic equation for waves describes spectra of developed wave turbulence, both stationary and non-stationary [2, 6, 8], yet the singularities hidden behind this nice picture haven't been analyzed. Here we open this Pandora box and start such an analysis using the most general approach of information theory. That (and processing real data of experiments) requires a discrete approach, so that we consider the number of modes N finite as well as the nonlinearity parameter ϵ. Subtleties related to taking limits $N \to \infty$ and $\epsilon \to 0$ are subject of the ongoing work [8–10, 17, 18].

The self-consistent weak-turbulence description of the one-mode statistics in terms of the occupation numbers n_k [2, 5–8, 17] guarantees the statistics of any single amplitude stays close to Gaussian, i.e $q(|a_k|)$ is Rayleigh for every single wave mode. That tempts one to approximate the whole distribution using only the set of n_k, assuming that the amplitudes are independent and the phases are random:

$$q\left(\{a_k, a_k^*\}\right) = \prod_k (2\pi)^{-1}\, q\left(|a_k|\right), \tag{4}$$

which implies a Gaussian approximation for $\rho\left(\{a_k, a_k^*\}\right)$. Here we show that ρ is quite different. The difference between distributions can be measured by the relative entropy (Kullback-Leibler divergence), which is the price of non-optimal coding in information theory: $D\left(\rho \mid q\right) = \langle \ln\left(\rho/q\right)\rangle$. Since q is a product, its entropy is a direct sum of the entropies of non interacting modes: $\sum_{\mathbf{k}} \ln\left(e\pi n_k\right) = \sum_{\mathbf{k}} S_k$. The relative entropy then coincides with the multi-mode mutual information $D\left(\rho \mid q\right) = \sum_{\mathbf{k}} S_k - S\left(\rho\right) := I\left(\{a_k, a_k^*\}\right)$. Remind that the mutual information is defined for any subsystems, A and B, via their entropies: $I(A,B) = S(A) + S(B) - S(A,B)$. For example, the mutual information between two parts of the message measures how much of the future part we can predict given the part already received.

Starting with a gaussian ρ at $t = 0$, at the times $1/\omega_k \ll t \ll t_{NL}$, the distribution ρ can be determined by the second and third moments using conditional entropy maximum (see Supplementary for details):

$$\rho = \frac{1}{Z}\exp\left[-\sum_{\mathbf{k}} \alpha_k\, |a_k|^2 + \sum_{\mathbf{kpq}} F_{kpq} a_k^* a_p a_q + \text{c.c}\right]. \tag{5}$$

For ρ to be normalizable, by (A.15) we mean a truncated series in powers of ϵ. Here we consider terms up to second order. Then the parameters α, F of the distribution can be expressed via the moments J and n:

$$F_{123} = J^*_{123}/2n_1n_2n_3 \tag{6}$$

$$\alpha_i^{-1} = n_i - \sum_{\mathbf{k_1k_2}} \frac{\left|J_{i12}\right|^2 + 2\left|J_{12i}\right|^2}{2n_1^2n_2^2}. \tag{7}$$

We saw that continuing concentration of the third moment on the resonant surface had no influence on the kinetic equation (3) since the integral of the imaginary part of the third moment saturates on the short timescale $1/\omega_k$. However, the relative entropy,

$$D(\rho|q) = \sum_{\mathbf{k}_i\mathbf{k}_j} \frac{|J_{i+jij}|^2}{2n_in_jn_{i+j}}, \tag{8}$$

is determined by the squared modulus, which depends dramatically on whether the system is in thermal equilibrium or not. The equilibrium $n_k = T/\omega_k$ is special because the last bracket in (2) is proportional to $\omega^1_{2,3}$, so the third cumulant is regular everywhere in k-space and constant for long times: $J_{123}(t) = -V^*_{123}T^2\omega_1/\omega_2\omega_3$. Therefore, at $t \to \infty$ the relative entropy is small comparing to the total entropy as long as nonlinearity is small:

$$D(\rho|q) = I\{a_k\} = T\sum_{ij}\frac{|V_{i+j,ij}|^2}{\omega_i\omega_j\omega_{i+j}} = \left(\frac{E_{int}}{T}\right)^2 . \tag{9}$$

Away from equilibrium, on the contrary, with time the third cumulant (2) concentrates in a close vicinity of the resonance surface. It leads to a profound difference between statistics of a wave system in equilibrium and in turbulence. The equilibrium probability of a configuration $\{a_1, a_2, a_3\}$ is insensitive to resonances, because it is determined by the instantaneous interaction energy divided by the (uniform) temperature: $exp[-\mathcal{H}/T]$, since $F^*_{123} = J_{123}/2n_1n_2n_3 = -V^*_{123}/2T$ in this case. For turbulence, the interaction energy is additionally weighted by the resonance factor $(n_1^{-1} - n_2^{-1} - n_3^{-1})/(\omega_1 - \omega_2 - \omega_3)$ as the probability is the result of a time averaging. The measure in the phase space is thus regular in equilibrium and tends to singular in turbulence.

The squared cumulants in the relative entropy (A.28) have a secular growth in weak turbulence: $\lim_{t\to\infty}|(e^{\imath\Delta t} - 1)/\Delta|^2 = 2\pi t\delta(\Delta)$. The Liouville theorem requires that this increase of the mutual information and decrease in total entropy is exactly equal to the growth of $S(q)$ due to the change in n_k according to (3):

$$\frac{dS(q)}{dt} = \sum_{\mathbf{k}}\frac{1}{n_k}\frac{dn_k}{dt} = \sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}\frac{1}{2n_1n_2n_3}\frac{d}{dt}|J_{123}|^2 . \tag{10}$$

Hamiltonian evolution by itself does not change the entropy $S(\rho)$, but non-equilibrium state requires pumping and damping by an environment. If its action makes n_k stationary, then the information production is ultimately due to the entropy extraction by the environment: $dD(\rho|q)/dt = -(dS/dt)_{env}$. We see that stationarity of the second moment does not mean stationary distribution. On the contrary, the third moment (and other cumulants) are getting more and more singular, reflecting the total entropy decrease and the growth of the relative entropy between the true distribution and the random-phase Gaussian approximation:

$$D(\rho|q) = t\sum_{\mathbf{kps}} U_{kps}\frac{n_pn_s - n_kn_p - n_kn_s}{n_kn_pn_s} > 0 . \tag{11}$$

Contribution to the relative entropy of every cumulant is proportional to its square. The fourth cumulant is $\propto V^2$, so its contribution is proportional to V^4 and can be neglected in this order. Formula (11) can be written as $D = t\sum_k t^{-1}_{NL}(k)$ and is the first term of the expansion in powers of time, valid at $t < t_{NL}$. The terms with higher powers of time will involve higher cumulants. One can estimate $t^{-1}_{NL}(k) \simeq \sum_j |V_{k+j,kj}|^2 n_j/\omega_j \simeq \omega_k\epsilon_k \propto k^{2m+d-s-\alpha}$. At $t \simeq t_{NL}(k)$, when nonlinearity at the three-wave resonant surfaces $\omega_j + \omega_k = \omega_{j+k}$ is getting of order unity, the triple moment is expected to stabilize. At that stage the entropy change already is not small, but could be comparable to the total entropy. At later time, the total entropy decrease is modified, but does not necessarily stop, contrary to what one may suggest. The reason is that the entropy extraction depends on the environment. We illustrate that for two qualitatively different ways of pumping the system.

Let us first add to the rhs of $\imath\partial a_k/\partial t = \partial\mathcal{H}/\partial a^*_k$ a random force and a damping, $f_k - \gamma_k a_k$, with $\langle f_k(0)f^*_j(t)\rangle = \delta_{kj}P_k\delta(t)$. When force and damping are not in detailed balance, i.e ω_kP_k/γ_k is not a constant, we have entropy production:

$$\left(\frac{dS}{dt}\right)_{\rm env} = 2\sum_{\mathbf{k}} P_k\int\prod_j\frac{da_jda^*_j}{2i\rho}\left|\frac{\partial\rho}{\partial a_k}\right|^2 - 2\sum_{\mathbf{k}}\gamma_k . \tag{12}$$

which depends on the distribution. Averaging in this case is over the force statistics. Let us show that if the steady distribution ρ exists, it must have very sharp gradients, proportional to the Reynolds number, so that the entropy is much smaller than $S(q)$. At the initial perturbative stage, the distribution is given by (A.15) and we can substitute (2,6,7) into (12) and obtain:

$$\sum_{\mathbf{k}}\int\prod_j\frac{da_jda^*_j}{2i}\rho^{-1}\left|\frac{\partial\rho}{\partial a_k}\right|^2 = \sum_{\mathbf{k}}\alpha_k + O(J^4). \tag{13}$$

Here α_k is given (7) where the last two terms are initially small. The pumping then produces much less entropy than the dissipation region absorbs (any non-equilibrium state consumes information, that is exists between a low-entropy source and a high-entropy sink). Indeed, the energy spectral density ω_qn_q is a decreasing function of q in a direct energy cascade, so for any $q > k_{pump}$ we have $\sum_{\mathbf{k}} P_kn_k^{-1} < (\omega_qn_q)^{-1}\sum_{\mathbf{k}}\omega_kP_k$ and

$$\left(\frac{dS}{dt}\right)_{\rm env} < 2(\omega_qn_q)^{-1}\sum_{\mathbf{k}}(\omega_kP_k - \gamma_k\omega_kn_k) = 0$$

follows from the energy balance $\sum_k\omega_k(P_k - \gamma_kn_k) = 0$. For a developed turbulence with a wide inertial interval $k_{max}/k_{min} = k_{damp}/k_{pump} = Re \gg 1$, the spectrum of the direct cascade is $n_k \propto k^{-s}$, and the ratio of the negative damping term to the positive pumping term in (12) can be estimated as $\omega_{pump}n_{pump}/\omega_{damp}n_{damp} \equiv Re^{s-\alpha} \gg 1$. Direct energy cascade requires $s > \alpha$, and indeed the entropy absorption by the small-scale dissipation region by far exceeds the entropy production by the pumping region. However, this is only true at the initial perturbative stage. As time proceeds, the growth of the cumulants and deviation of distribution from Gaussian decreases dD/dt by increasing the pumping contribution. For developed turbulence, the gradients $\partial\rho/\partial a_k$ in the pumping region must increase by a large factor n_{pump}/n_{damp} to reach the steady measure, which is thus very close to singular.

Another way of creating non-equilibrium is by adding to the Hamiltonian equations of motion the terms $\gamma_k a_k$, where positive γ_k corresponds to an instability and negative to dissipation. Averaging in this case is over the ensemble of initial conditions. The entropy rate of change $dS_{\rm env}/dt = 2\sum_{\mathbf{k}}\gamma_k \le 0$ is now independent of the distribution and negative for a steady direct cascade for the same reasons of the energy conservation $\sum_{\mathbf{k}} 2\omega_k\gamma_k n_k = 0$ and $\omega_k n_k$ being larger in the instability region. That means that the entropy decreases non-stop and the measure goes all the way to singular unless coarse-graining saturates the entropy decrease. Profound difference between turbulent measures generated by additive force and instability was probably first noticed in [19].

To verify our other predictions, one needs to obtain numerically and experimentally multi-dimensional probability distributions. The simplest is to start from two modes. The pair correlation function, $\langle a_{\mathbf{k}} a^*_{\mathbf{p}}\rangle = 0$ for $\mathbf{k} \neq \mathbf{p}$ due to translation invariance, but the fourth cumulant is generally nonzero and so must be the mutual information (first introduced in [20] for one-dimensional models of turbulence). In thermal equilibrium and for non-resonant modes in turbulence, steady-state MI must be small for small nonlinearity. MI between two modes is given by $I_{\mathbf{k},\mathbf{p}} = |J_{\mathbf{k},\mathbf{p},\mathbf{k}+\mathbf{p}}|^4/(4n_{\mathbf{k}}n_{\mathbf{p}}n_{\mathbf{k}+\mathbf{p}})^2 + |J_{\mathbf{k},\mathbf{p},\mathbf{k}-\mathbf{p}}|^4/(4n_{\mathbf{k}}n_{\mathbf{p}}n_{\mathbf{k}-\mathbf{p}})^2 \propto \epsilon^4$. The ϵ^2-contribution requires minimum three modes: $I_{\mathbf{k},\mathbf{p},\mathbf{q}} = S(a_k) + S(a_p) + S(a_q) - S(a_k, a_p, a_q) = |J_{\mathbf{k},\mathbf{p},\mathbf{q}}|^2/2n_k n_p n_q$. We expect order-unity cumulants (as seen, for instance, in [21]) and substantially non-Gaussian stationary joint distribution for resonant modes in turbulence. Finding that distribution is a well-posed task for a future work.

Our consideration of the MI growth allows solving the old puzzle: why the direction of the formation of the turbulent spectrum $n_k \propto k^{-m-d}$ is determined by the energetic capacity? When the total energy $\sum_{\mathbf{k}}\omega_k n_k$ diverges at infinity ($m < \alpha$, infinite capacity), the formation proceeds from large to small scales, that is from pumping to dissipation [22]. In the opposite finite-capacity case, $m > \alpha$, formation of the cascade was surprisingly found to start from small and proceeds to large scales, that is opposite to the cascade direction [6, 23, 24]. We note that it is the growth of MI that must determine the direction of the evolution, since it quantifies the build-up of multi-mode correlations necessary for a steady non-equilibrium state. For $n_k \propto k^{-s} = k^{-m-d}$, the growth rate of the three-mode mutual information (learning rate) scales as $dI_{\mathbf{k},\mathbf{p},\mathbf{q}}/dt \propto k^{2m+d-s-\alpha} = k^{m-\alpha}$. One then can characterize the directionality of the information transfer by the sign of $m - \alpha$ — when it is positive, correlations must be established first at small scales and then propagate to larger scales. Since the energetic capacity is also finite for the Kolmogorov spectrum of the incompressible turbulence, tantalizing question is whether it is also formed starting from small scales. Note that we characterized evolution by the growth rate of MI, which is to be distinguished from the transfer entropy [25], which characterizes cause-effect relationships in a steady state. Remark that though the MI between non-interacting Gaussian wave modes is zero, the multi MI between points in physical space, $I(x_1, \ldots, x_N) = N\ln\Big(e\pi N^{-1}\sum_{k=1}^{N} n_k\Big) - \sum_k \ln(e\pi n_k)$, is positive whenever n_k are not all the same.

Let us briefly compare our findings with similar effects near thermal equilibrium. Non-equilibrium anomalies in cumulants were noticed first in the simplest case of the linear response of a dilute gas. Series in powers of density for viscosity and diffusivity contain infinities due to Dorfman-Cohen memory effects in multiple collisions [11–13, 26]. Another profound difference is in the two-particle correlation function in two distinct space points at the same time. Outside of the radius of molecular forces, this correlation is zero in thermal equilibrium and nonzero away from it [15, 16]. Non-equilibrium build-up of long spatio-temporal correlation is a counterpart to our spectral singularities.

Another analogy worth exploring is with many-body localization [27], where phase correlations prevent thermalization and keep the system in a low-entropy state. There is a vast literature devoted to cumulant anomalies away from equilibrium, see e.g. books [1, 2, 12–16] and numerous references there. We believe that complementarity of information theory and singular measures will lead to a unified approach to these anomalies.

To conclude, we reiterate our main results: the probability distribution of weak wave turbulence is very far from Gaussian, the mutual information is substantial for resonant modes.

We thank A. Zamolodchikov, V. Lebedev, L. Levitov, K. Gawedzki and R. Chetrite for useful discussions. The work was supported by the Scientific Excellence Center and Ariane de Rothschild Women Doctoral Program at WIS, grant 662962 of the Simons foundation, grant 075-15-2019-1893 by the Russian Ministry of Science, grants 873028 and 823937 of the EU Horizon 2020 programme, and grants of ISF, BSF and Minerva.

---

[1] Dorfman, J. R.. *An introduction to chaos in nonequilibrium statistical mechanics* (Cambridge Univ Press 1999)

[2] D Evans, G Moriss, *Statistical Mechanics of Non Equilibrium Liquids* (Cambridge Univ Press 2008)

[3] G Falkovich, K. Gawedzki and M Vergassola, *Rev Mod Phys* **73**, 913 (2001)

[4] V. Zakharov, V. Lvov. G. Falkovich, *Kolmogorov Spectra of Turbulence* (Springer 1991)

[5] R Peierls, *Ann. Phys.* **3**, 1055 (1929)

[6] A Newell1 and B Rumpf, *Annu. Rev. Fluid Mech.* **43** 59Ũ78 (2011)

[7] P Jakobsen and A Newell, *J. Stat. Mech.* L10002 (2004)

[8] S. Nazarenko, *Wave turbulence* (Springer 2011)

[9] T Buckmaster, P Germain, Z Hani, J Shatah, arXiv:1907.03667

[10] A Dymov and S Kuksin, arXiv: 1907.05044

[11] J. R. Dorfman and E. G. D. Cohen, *Phys. Lett.* **66**, 124 (1965).

[12] R. Peierls, *Surprises in Theoretical Physics* (Princeton Univ Press 1979).

[13] J. R. Dorfman and E. G. D. Cohen, *Phys. Rev. Lett.* **25**, 1257 (1970).

[14] J. R. Dorfman and E. G. D. Cohen, *J Math Phys* **8**, 282 (1967);

[15] E Lifshits and L Pitaevsky, *Physical Kinetics* (Elsevier, Amsterdam 1981)

[16] Sh. Kogan, *Electronic noise and fluctuations in solids* (Cambridge Univ Press 2010).

[17] Y Choi, Y Lvov and S Nazarenko, *Physica D*, **201**, 121 (2005)

[18] G Eyink and Y-K Shi, *Physica D* **241** 1487 (2012)

[19] N Vladimirova, S Derevyanko, G Falkovich, *Phys Review E* **85**, 010101R (2012)

[20] K. Ikeda and K. Matsumoto, *Phys. Rev. Lett.* 62, 2265 (1989)

[21] Y Pan and DKP Yue, *J Fluid Mech*, **816**, R1 (2017).

[22] G Falkovich and A Shafarenko, *J. Nonl. Sci.* **1**, 457 (1991)

[23] C Connaughton and A Newell, *Phys Rev E* **81**, 036303 (2010)
J. Stat. Mech. Theor. Exp. L10002 (2004)

[24] S. Galtier, S. Nazarenko, A. Newell and A. Pouquet, *J. Plasma Phys.* **63**, 447 (2000).

[25] T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).

[26] R. A. Piccirelli, *J. Math. Phys.* 7, 922 (1966) *Sov. Phys. JETP* **29** 467 (1969)

[27] D Abanin, E Altman, I Bloch, M Serbyn, *Rev Mod Phys*, **91** (2019)

Supplementary Material

The Probability Distribution

Here we outline the derivation of the probability distribution ρ for short times (compared to nonlinear time t_{NL}) and its entropy.

The probability distribution determined by the second and third moments, n_k and J_{kpq} respectively, assuming maximum conditional entropy, is an extremum of the functional

$$Q[\rho] = S(\rho) - \sum_{\mathbf{k}} \alpha_k n_k + \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} F_{kpq} J_{kpq} + \text{c.c} + \lambda \int \prod_j \frac{da_j da_j^*}{2i} \rho, \tag{A.14}$$

where α_k and F_{kpq} are the corresponding Lagrange multipliers and the last term is a normalization condition. The solution to the extremum problem, apart from the normalization term, is given by

$$\rho = Z^{-1} \exp\left[-\sum_{\mathbf{k}} \alpha_k \left|a_k\right|^2 + \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} F_{kpq} a_k^* a_p a_q + \text{c.c} \right]. \tag{A.15}$$

For ρ to be normalizable, by Eq. (A.15) we mean a truncated series in $\epsilon = |J_{kpq}|^2 / n_k n_p n_q \ll 1$, assuming weakness of non-Gaussianity:

$$\rho = Z^{-1} \exp\left[-\sum_{\mathbf{k}} \alpha_k \left|a_k\right|^2 \right] \left(1 + \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} F_{kpq} a_k^* a_p a_q + \text{c.c} + ... \right). \tag{A.16}$$

The first two terms of the entropy of the distribution are then

$$S(\rho) = \sum_{\mathbf{k}} \ln \pi e n_k - \sum_{\mathbf{ij}} \frac{|J_{i+jij}|^2}{2 n_i n_j n_{i+j}} \ , \tag{A.17}$$

which give the relative entropy:

$$D(\rho|q) = \sum_{\mathbf{ij}} \frac{|J_{i+jij}|^2}{2 n_i n_j n_{i+j}} \ . \tag{A.18}$$

where $q = \prod_k \frac{1}{\pi n_k} e^{-\sum_k |a_k|^2/n_k}$ is the Gaussian approximation.

Entropy Production

Here we outline the derivation of (13) in the main text:

$$\sum_{\mathbf{k}} \int \prod_j \frac{da_j da_j^*}{2i} \rho^{-1} \left| \frac{\partial \rho}{\partial a_k} \right|^2 = \sum_{\mathbf{k}} \alpha_k + O(J^4). \tag{A.19}$$

Starting with (A.16), the derivative of ρ with respect to a_k is given by

$$\frac{\partial \rho}{\partial a_k} = \left(-\alpha_k a_k^* + \sum_{ij}^{N} \left(2 F_{ijk} a_i^* a_j + F_{kjl}^* a_j^* a_i^* \right) \right) \rho \tag{A.20}$$

so $\forall k$

$$\begin{aligned}\int\prod_j \frac{da_j da_j^*}{2i}\rho^{-1}\left|\frac{\partial\rho}{\partial a_k}\right|^2 =&\alpha_k^2 n_k - \alpha_k\sum_{ij}^{N}\left(2F_{ijk}^*\left\langle a_i a_j^* a_k^*\right\rangle + F_{kjl}\left\langle a_j a_i a_k^*\right\rangle + 2F_{ijk}\left\langle a_i^* a_j a_k\right\rangle + F_{kji}^*\left\langle a_j^* a_i^* a_k\right\rangle\right)\\ &+\sum_{ijlm}^{N}\left(4F_{lmk}^* F_{ijk}\left\langle a_i^* a_j a_l a_m^*\right\rangle + F_{kji}^* F_{kml}\left\langle a_j^* a_i^* a_m a_l\right\rangle\right) \qquad \text{(A.21)}\\ =&\alpha_k^2 n_k - 2\sum_{ij}^{N}\left(2\left|F_{ijk}\right|^2 + \left|F_{kji}\right|^2\right)n_i n_j + O\left(F^3\right) \qquad \text{(A.22)}\\ =&\alpha_k\left(\alpha_k n_k - \sum_{ij}^{N}\left(\frac{\left|J_{ijk}\right|^2}{n_i n_j n_k} + \frac{\left|J_{kij}\right|^2}{2n_i n_j n_k}\right)\right) + O\left(F^3\right). \qquad \text{(A.23)}\end{aligned}$$

Using the first terms in the expansion for α_k

$$\alpha_k = \frac{1}{n_k}\left(1 + \sum_{ji}\left(\frac{\left|J_{kji}\right|^2}{2n_j n_i n_k} + \frac{\left|J_{jik}\right|^2}{n_j n_i n_k}\right)\right), \qquad \text{(A.24)}$$

yields (A.19).

Information Capacity of Turbulent Cascades

This section includes a complementary and broader discussion to the main text regarding the information capacity of turbulent cascades.

The most important practical problem in modeling multi-mode systems, both in and out of equilibrium, is an inability to resolve all scales. It is then imperative to learn how much information about the whole system one can reliably receive from the scales explicitly accounted for. This can be quantified by the mutual information between small and large scales, which we now compute perturbatively for an interactive wave system in and out of equilibrium. Remind that the relative entropy and the mutual information decrease monotonically upon any partial average.

Let us integrate the N-wave probability distribution (A.15) over the amplitude and phase of the N-th harmonic:

$$\rho\left\{a_1\dots a_{N-1}\right\} = \frac{\pi}{Z_N\alpha_N}\exp\left[-\sum_{k=1}^{N-1}\alpha_k\left|a_k\right|^2 + \sum_{ijk}^{N-1}(F_{ijk}a_i^* a_j a_k + \text{c.c}) + \sum_{ijmn}^{N-1} n_N F_{Nij}F_{Nmn}^* a_i a_j a_m^* a_n^*\right]. \qquad \text{(A.25)}$$

We see that the pdf of the remaining modes depends on $F_{Nji} = J_{Nij}^*/2n_i n_j n_N$, which according to (2) in the main text, $J_{123}(t) = \frac{e^{i\omega_{2,3}^1 t}-1}{\omega_{2,3}^1}V_{123}^*(n_2 n_3 - n_1 n_2 - n_1 n_3)$, can be expressed via $n_i n_j n_N$ for weakly-interacting wave system. However, the entropy of the remaining modes is independent of the eliminated modes in this order:

$$S_{N-1} = \sum_{k}^{N-1}\ln\pi e n_k - \sum_{ijk}^{N-1}\frac{\left|J_{ijk}\right|^2}{2n_i n_j n_k} + \mathrm{O}(\epsilon^4)\ . \qquad \text{(A.26)}$$

As a result, up to ϵ^4-terms, the mutual information between the two sets of waves with the wave numbers below and above some k respectively is determined by the triple correlation between the waves taken from different sets:

$$I(k) = S(1,k) + S(k,N) - S_N = \sum_{j<k<i}\frac{\left|J_{ijl}\right|^2}{2n_i n_j n_l}\ . \qquad \text{(A.27)}$$

It can be called the information capacity of the cascade. If k is the resolution scale then $dI(k,t)/dt$ measures the rate of the information loss. More and more correlations appear at finer and finer scales, all lost below the resolution scale.

Eliminating the N-th harmonic thus decreases the mutual information by removing some contributions from the sum of all positive terms, while the changes in the remaining terms appear only in the next order. The expression for the relative entropy, (9) in the main text,

$$D(\rho|q) = I\{a_k\} = T\sum_{ij} \frac{|V_{i+j,ij}|^2}{\omega_i \omega_j \omega_{i+j}} = (E_{int}/T)^2 \ , \tag{A.28}$$

is uniformly valid for thermal equilibrium in a weakly interacting system of waves. The same is true for the general perturbative expression (A.18) at any given time only as long as the measure did not deviate far from Gaussian.

Scaling of the Relative Entropy and Mutual Information

It is interesting to find out how the relative entropy (multi-mode mutual information) depends on the number N of the degrees of freedom and compare it with the linear dependence of the extensive entropy $S(q)$. For example, MI of a system with the pair interaction (spins, neurons) grows with N as the number of pairs, $N(N-1)/2$, at least for low enough N. Finding the critical N_c when the quadratic part is comparable to the linear one, one can estimate, for instance, the cluster size of strongly correlated neurons in the brain [1]. For a wave system in a fixed box, we assume $N \propto (k_{max}/k_{min})^d$. Let us keep the box size $L = 2\pi/k_{min}$ fixed, so that N changes as k_{max}^d. In computing occupation numbers one must assume that $n_k \propto (k_{max}/k)^s$ so that $S(q) = \sum_k \ln(\pi e n_k)$ is indeed extensive (and not $N \ln N$). In thermal equilibrium, the relative entropy is determined by the modes with $k \simeq k_{max}$, so that one estimates $D \simeq N\epsilon_{k_{max}}^2 \simeq N|V(k_{max})|^2 T/\omega_{max}^3 \propto \epsilon^2 N^{1+(2m+d-3\alpha)/d}$, where we denoted the system-scale nonlinearity parameter: $\epsilon = \epsilon_k (k/k_{min})^{2\alpha+s-2m-d} = \epsilon_k (k/k_{min})^{3\alpha-2m-d}$. The critical number of modes, over which perturbation approach fails and waves cannot be considered weakly correlated, then scales with the nonlinearity as $N_c \simeq \epsilon^{d/(3\alpha-2m-d)}$. On the contrary, when the relative entropy is sub-extensive, the larger the number of modes the less correlated they are effectively. On the Kolmogorov spectrum, $dD/dt \propto k_{max}^{m+d-\alpha} \propto N^{1+(m-\alpha)/d}$, which is super-extensive when $\alpha < m$, that is when the interaction energy grows with k faster than the sum of the energies of the separate waves.

The mutual information (A.27) in thermal equilibrium scales as $I(k) \propto k_{max}^{2m-m_1+d-2\alpha} k^{m_1+d-\alpha}$, when $m_1+d-\alpha > 0$. When $2m - m_1 > 2\alpha$, the mutual information is sub-extensive, which can be thought as an analog of an area law. In turbulence, MI grows according to $dI(k,t)/dt \propto k_{max}^{2m-m_1+d-3\alpha+1} k^{m_1+2\alpha-m-1} \propto N^{1+(2m-m_1-3\alpha+1)/d}$. Note that the spectrum locality requires $m_1 + 2\alpha - m - 1 > 0$ [2].

For reader's convenience, Table 1 presents parameters for different wave systems with resonant three-wave interaction.

Wave system	α	m	m_1
3D acoustic waves	1	3/2	1
2D weakly dispersive waves	1	1	1
Capillary waves on deep water	3/2	9/4	7/2
Capillary waves on shallow water	2	2	0

TABLE I: Parameters for different wave systems with resonant three-wave interaction.

The mutual information is parametrically larger in turbulence than in equilibrium. Indeed as stated in the main text the resolved scales carry significant information about the unresolved scales. However, if we are only interested in large-scale properties, the ignorance of small scales carries much higher price in turbulence than in equilibrium.

There are two ways to mitigate the effect of that ignorance: parameterize the unresolved scales by a small set of variables and/or modify the equations of motion of the scales resolved. How well this task is accomplished is to be measured by the relative entropy between the true distribution and that obtained from solving the restricted set of equations. In particular, Computer modeling poses the question: if there is a way to renormalize the Hamiltonian of the modes explicitly computed so that their statistics is faithfully reproduced. For weakly interacting waves, (A.25) suggests the following renormalization to account for subgrid modes at $k \geq L$:

$$\delta\mathcal{H} = \sum_{ijmn} \sum_{k=L}^{N} T_{ijmn}^{k} a_i^* a_j^* a_m a_n \Delta(k-i-j)\Delta(k-m-n) \, . \tag{A.29}$$

We need to choose T^k_{ijmn} so that the evolution of L modes with the new Hamiltonian approximates the true entropy of (A.25) up to ϵ^4. This is possible in thermal equilibrium by choosing $T^k_{ijmn} = V_{kij}V^*_{kmn}n_k/T = V_{kij}V^*_{kmn}/\omega_k$; essentially the same procedure is integrating out fast degrees of freedom in effective ăquantum field theories. It is straightforward to see that it does not work away from thermal equilibrium. Generally, in turbulence subgrid modes not only renormalize Hamiltonian, but provide also random forces with a non-trivial statistics. More likely, a correct way to account for subgrid modes in turbulence is a direct renormalization of the probability distribution, rather than Hamiltonian; this will be treated elsewhere.

Scattering Processes

It is straightforward to include the four-wave scattering into the entropic treatment of weak wave turbulence. Consider the Hamiltonian

$$\mathcal{H} = \sum_k \omega_{\mathbf{k}} \left|a_{\mathbf{k}}\right|^2 + \frac{1}{2}\sum_{i+j=m+n} T_{ijmn} a^*_i a^*_j a_m a_n \,, \tag{A.30}$$

where $T_{ijmn} = T^*_{mnij}$. We define the renormalized frequency $\tilde{\omega}_k = \omega_k + \sum_i T_{ikik} n_i$, where as before $n_i = \langle |a_i|^2 \rangle$. Denote $\Delta = \tilde{\omega}_1 + \tilde{\omega}_2 - \tilde{\omega}_3 - \tilde{\omega}_4$. Similarly to three-wave ineraction, we obtain in the first order the cumulant $J_{ijkl} = \langle a^*_i a^*_j a_k a_l \rangle - 2n_i n_k \delta_{ik}\delta_{jl} = \langle\langle a^*_i a^*_j a_k a_l \rangle\rangle$:

$$J_{1234} = 2T^*_{1234} n_1 n_2 n_3 n_4 \left(n_4^{-1} + n_3^{-1} - n_2^{-1} - n_1^{-1}\right) \frac{e^{i\Delta t} - 1}{\Delta} \delta\left(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4\right) \,, \tag{A.31}$$

which gives linearly decaying entropy for non-equilibrium spectra. Note that $J_{ijij} = 0$.

The distribution determined by n_i and J_{ijkl} is again given by the conditional entropy maximum under the assumption $|J_{ijkl}|^2/n_i n_j n_k n_l \equiv \varepsilon^2 \ll 1$:

$$\rho\{a_k\} = Z^{-1} \exp\left[-\sum_k \alpha_k |a_k|^2 + \sum_{ijkl} G_{ijkl} a^*_i a^*_j a_k a_l\right] . \tag{A.32}$$

where the overall normalization and second moment are given by

$$Z = \big(1 + 2\sum_{ijkl} \frac{G^2_{ijkl}}{\alpha_i \alpha_j \alpha_k \alpha_l}\big) \prod_{l=1}^{N} \frac{\pi}{\alpha_l}, \tag{A.33}$$

$$n_i = \alpha_i^{-1} + 2\alpha_i^{-2} \sum_{jkl} \frac{G^2_{ijkl} + G^2_{jkil}}{\alpha_j \alpha_k \alpha_l}. \tag{A.34}$$

In this case, $G_{1234} = J^*_{1234}/4n_1 n_2 n_3 n_4$, and the relative entropy (the multi-mode mutual information) is expressed via the sum of the fourth cumulants:

$$D(\rho|q) = \sum_{ijkl} \frac{J^2_{ijkl}}{8 n_i n_j n_k n_l} \,, \tag{A.35}$$

which also grows linearly with time in weak turbulence. Apart from the energy, scattering processes conserve also the number of waves, which makes possible two-cascade state with a direct cascade of energy and inverse cascade of waves. In an inverse cascade, occupation numbers at the sink are larger than at the source, so that $dS_{env}/dt \simeq 2\gamma_k(n_{damp}/n_{pump} - 1) \geq 0$. That means that in a weakly nonlinear regime an inverse cascade absorbs entropy and generates information, that is it cannot exist without a direct cascade, which provides overall entropy production. It is straightforward to show that any stationary weak turbulence producing and absorbing both integrals of motion generates entropy. Examples of wave systems with four-wave scattering are given in Table 2.

Wave system	d	α	m
Surface gravity waves	2	1/2	3
plasmons and manons	3	2	0 or 2
Elastic waves in thin plates	2	2	0

TABLE II: Parameters for different wave systems with four-wave scattering.

One can also consider the general quantum kinetic equation for scattering [2]:

$$\frac{\partial n_k}{\partial t} = \int |T_{k123}|^2 F_{k123} \delta(\mathbf{k}+\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_3)\delta(\omega_k+\omega_1-\omega_2-\omega_3) d\mathbf{k}_1..d\mathbf{k}_4 \,,$$
$$F_{k123} = (n_k+1)(n_1+1)n_2 n_3 - n_k n_1 (n_2+1)(n_3+1)\,. \tag{A.36}$$

In the limit $n_k \gg 1$ it gives the classical scattering kinetic equation (A.31). In the opposite limit $n_k \ll 1$ it gives Boltzmann equation:

$$\frac{\partial n_p}{\partial t} = \int |T_{p123}|^2 (n_p n_1 - n_2 n_3)\delta(\mathbf{p}+\mathbf{p}_1-\mathbf{p}_2-\mathbf{p}_3)\delta(\epsilon_p+\epsilon_1-\epsilon_2-\epsilon_3) d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \,. \tag{A.37}$$

Directions of Future Studies

Let us briefly discuss how one can use the information theory to learn more on different cases of turbulence. Traditionally, turbulence studies have focused on single-point probability distributions and correlations between two or three points. Just like the progress in stochastic thermodynamics and data processing required passing from correlation functions to entropy and mutual information, the future progress of turbulence studies may lie in this direction. The whole multi-point probability functional, of course, contains all the information, but it is too vast. However, the entropy of this distribution must be manageable and may answer questions hitherto unasked. For example, entropy measures the rate with which we acquire extra information upon increasing spatial, temporal or amplitude resolution. In addition, mutual information will let us quantify how much information we can infer from some spatial, temporal or spectral domains about the other, unknown, domains.

Further, note that the mutual information between Fourier modes is interesting even when no waves are possible, as in incompressible turbulence. In this case, the energy flux is cubic in velocity, which suggests studying the three-mode mutual information (going beyond the perturbative approach of [3]). In the other extreme, shock creation in compressible turbulence imposes multi-mode phase correlations, which would be very interesting to characterize by the respective multi-mode mutual information. Maxima of multi-mode mutual information may also reveal a connectivity graph of a system with a finite number of interacting modes.

Much remains to be learnt from the passive scalar turbulence [3]. Consider the scalar field $\theta(\mathbf{r},t)$ passively transported by an incompressible flow $\mathbf{v}(\mathbf{r},t)$ and pumped by $\xi(\mathbf{r},t)$:

$$(\partial_t + v_i \nabla_i)\theta = \xi \,. \tag{A.38}$$

Both velocity and pumping are Gaussian. The passive scalar turbulence is amenable to analytic treatment when velocity is either smooth in space (Batchelor case) or rough in time (Kraichnan case) [3]. In the Batchelor case, the scalar statistics is close to Gaussian, yet the statistics of scalar gradients is not. We think that singularity of the measure must manifest itself in angular singularities of cumulants near collinear configurations. The only known case is a cusp for four-point cumulant [4]. One may hope that information-theoretical treatment of this case might be possible within a perturbative approach. In the Kraichnan case, the scalar statistics is close to Gaussian in the limits of large space dimensionality and spatially rough velocity. These two limits are amenable to analytic treatment, where the cumulants can be computed perturbatively. That would be of much interest to compute the entropy and the mutual information for the scalar field and see how they depend on the space dimensionality and the velocity roughness.

Another promising direction could be a renormalization-group (RG) analysis of the information content of turbulence, in particular, the insight it gives into irreversibility of RG as the best way to learn by forgetting, see [5–7]. As we have seen, in thermal equilibrium the multi-mode MI is proportional to ϵ_k. The mutual information of the wave system decreases monotonically under coarse-graining. However, two other steps of RG, re-scaling and renormalization, may increase MI depending on the k-dependence of ϵ_k. For example, a perturbative analysis for the system where the lowest nonlinearity is four-wave scattering, described in Section , is expected to go along the lines of Wilson epsilon-expansion [8]. In that case, the Gaussian fixed point will be either stable or unstable depending on the scaling $\epsilon_k \propto k^{\varepsilon} = k^{s+\alpha-m-d}$, that is ε plays the role of $4-d$ in our case. The mutual information between modes will then respectively increase/decrease upon RG flow for negative/positive ε. This is apparently because some of the information about eliminated degrees of freedom is stored in the renormalized values of the cumulants. Therefore, MI cannot universally play the role of C-function that guarantees the irreversibility of RG flow. Finding the proper informational characteristics for different RG schemes remains the task for the future [9]. Mention also the use of MI for identifying the relevant degrees of freedom and executing RG steps by a machine-learning algorithm without any prior knowledge about the system [10]. RG analysis may also illuminate the profound differences between the fixed points of RG: equilibrium versus turbulence and direct versus inverse cascade. In turbulence, $I(\mathbf{k},\mathbf{p},\mathbf{q})$ must be vanishingly small away form the resonant surfaces $\mathbf{k}=\mathbf{p}+\mathbf{q}$, $\omega_k=\omega_p+\omega_q$. The perturbative consideration demonstrated the growth on the resonant surfaces, but cannot determine where the growth stops and how the resulting mutual information depends on ϵ_k and the ratios k/k_{min} and k/k_{max}. Finding the ultimate mutual information of the invariant measure of weak turbulence remains the task for future, since the weak turbulence fixed point must be very far from Gaussian. Particularly interesting it is to establish how the total entropy of a turbulence system scales with the number of degrees of freedom to see if there an "area law of turbulence".

[1] Schneidman, E., Berry, M.J., Segev, R. and Bialek, W., 2006. *Nature*, 440(7087), pp.1007-1012.
[2] V. Zakharov, V. Lvov. G. Falkovich, *Kolmogorov Spectra of Turbulence* (Springer 1991)
[3] T M Brown, *J. Phys. A* **15** 2285 (1982)
[4] E. Balkovsky, M. Chertkov, I. Kolokolov, V. Lebedev, arXiv:chao-dyn/9508002, *JETP Lett.* **61**, 1012 (1995).
[5] A B Zamolodchikov *JETP Lett.* **43** 730 (1986)
[6] H Casini and M Huerta, *J Phys A* **40**, 7031 (2007)
[7] Z Komargodski and A Schwimmer, *J. High Energ. Phys.* 2011:99 (2011)
[8] K Wilson and J Kogut, *Physics Reports* **12**, 75-199 (1974)
[9] S M Apenko, Information theory and renormalization group flows, *Physica A* **391**, 67 (2012)
[10] Koch-Janusz, M. and Ringel, Z., 2018. *Nature Physics*, 14(6), pp.578-582.

Fibonacci turbulence

Natalia Vladimirova[1,2], Michal Shavit[1] and Gregory Falkovich[1]
Weizmann Institute of Science, Rehovot 76100 Israel
Brown University, Providence, RI 02912, USA

(Dated: April 20, 2021)

Never is the difference between thermal equilibrium and turbulence so dramatic, as when a quadratic invariant makes the equilibrium statistics exactly Gaussian with independently fluctuating modes. That happens in two very different yet deeply connected classes of systems: incompressible hydrodynamics and resonantly interacting waves. This work presents the first detailed information-theoretic analysis of turbulence in such strongly interacting systems. The analysis involves both energy and entropy and elucidates the fundamental roles of space and time in setting the cascade direction and the changes of the statistics along it. We introduce a beautifully simple yet rich family of discrete models with triplet interactions of neighboring modes and show that it has quadratic conservation laws defined by the Fibonacci numbers. Depending on how the interaction time changes with the mode number, three types of turbulence were found: single direct cascade, double cascade, and the first ever case of a single inverse cascade. We describe quantitatively how deviation from thermal equilibrium all the way to turbulent cascades makes statistics increasingly non-Gaussian and find the self-similar form of the one-mode probability distribution. We reveal where the information (entropy deficit) is encoded and disentangle the communication channels between modes, as quantified by the mutual information in pairs and the interaction information inside triplets.

I. INTRODUCTION

Existence of quadratic invariants and Gaussianity of equilibrium in a strongly interacting system may seem exceptional. Indeed, generic systems have no invariants except Hamiltonian. Strongly interacting systems have non-quadratic Hamiltonians, so that equilibrium Gibbs distribution (the exponent of the Hamiltonian) is generally non-Gaussian. And yet two very distinct wide classes of physical systems have quadratic invariants and Gaussian statistics at thermal equilibrium. The first class is the family of hydrodynamic models, starting from the celebrated hydrodynamic Euler equation and including many equations for geophysical, astrophysical and magnetohydrodynamic flows. The second class, as will be described in this paper, contains systems of resonantly interacting waves. We show that the discretized models of the first class exactly correspond to the second one. We shall consider one particular (arguably the simplest) family of such models and describe far-from equilibrium (turbulent) states of such systems.

One calls turbulence a state of any system, where many degrees of freedom are deviated far from thermal equilibrium. Therefore, studies of turbulence encompass a wide variety of phenomena in nature and industry, from pipe flows to ripples on a paddle. It can be studied from the viewpoint of a mathematician, engineer or a physicist. Here we employ the perspective of statistical physics, which is interested in fundamental principles that determine statistical distributions in turbulence and thermal equilibrium. We shall use both the traditional viewpoint of cascades and the relatively recent viewpoint of information theory, that is we address both energy and entropy of turbulence. So far, statistical physics approach to turbulence was to a large extent devoted to two quite distinct classes: systems of interacting waves like those on the surface of the ocean or a paddle and incompressible vortical flows where no waves are possible. Here we build a bridge between these two classes and show that discrete models of a certain kind can describe both.

On the one hand, the vorticity, $\omega = \nabla \times \mathbf{v}$, of an isentropic flow of incompressible fluid satisfies the Euler equation: $\partial\omega/\partial t = \nabla \times (\mathbf{v} \times \omega)$. Quite similar are two-dimensional hydrodynamic models, where a scalar field a (vorticity, temperature, potential) is linearly related to the stream function ψ of the velocity carrying the field: $\partial a/\partial t = -(\mathbf{v}\cdot\nabla)a$, $\mathbf{v} = (\partial\psi/\partial y, -\partial\psi/\partial x)$, $\psi(\mathbf{r}) = \int d\mathbf{r}'|\mathbf{r}-\mathbf{r}'|^{m-2}a(\mathbf{r}')$. For the 2D Euler equation, $m = 2$. Other cases include surface geostrophic ($m = 1$), rotating shallow fluid or magnetized plasma ($m = -2$), etc. After Fourier transform,

$$\dot a_{\mathbf k} = \sum\nolimits_{\mathbf q}[\mathbf k \times \mathbf q]\, q^{-m} a_{\mathbf q} a_{\mathbf k - \mathbf q}\,. \tag{1}$$

All such equations have quadratic nonlinearity and quadratic invariants. Then it was suggested [1] to model different cases of fluid turbulence by the chains of ODEs having quadratic invariant $g_{ij}u^i u^j$ and these properties:

$$\dot u_i = \Gamma^i_{jl} u_j u_l\,, \quad \Gamma^i_{il} = 0 = g_{ik}\Gamma^k_{jl} + g_{lk}\Gamma^k_{ji} + g_{jk}\Gamma^k_{li}\,. \tag{2}$$

On the other hand, consider resonantly interacting waves with the general Hamiltonian,

$$\mathcal{H}_w = \sum\nolimits_i \omega_i |b_i|^2 + \sum\nolimits_{ijl}\left(V_{l,ij} b_i^* b_j^* b_l + V^*_{l,ij} b_i b_j b_l^*\right)\,, \tag{3}$$

where $V_{l,ij} \neq 0$ only if $\omega_i + \omega_j = \omega_l$. By the gauge transformation, $a_i = b_i \exp(\imath\omega_i t)$, we can turn the equations of motion, $\imath\dot b_i = \partial\mathcal{H}_w/\partial b_i^*$ into a system of the type (1,2):

$$\imath\dot a_i = \sum\nolimits_{jl}\left(V^*_{i,jl} a_j a_l + 2V_{l,ij} a_j^* a_l\right)\,. \tag{4}$$

This means that quadratic and cubic parts of the Hamiltonian are conserved separately. If such a system is brought into contact with thermostat, it is straightforward to show that the statistics is Gaussian: $\ln\mathcal{P}\{a_i\}\propto-\sum_i\omega_i|a_i|^2$.

Our interest in resonances is connected to that in non-equilibrium. Thermal equilibrium does not distinguish between resonant and non-resonant interactions because of the detailed balance: whatever correlations can be built over time between resonantly interacting modes, the reverse process destroying these correlations is equally probable. This is not so away from thermal equilibrium, especially in turbulence.

Neglecting non-resonant and accounting only resonant interactions is the standard approach to weakly interacting systems, even though the weak nonlinearity assumption breaks for resonant modes. Weak turbulence theory gets around this by considering continuous distribution and integrating over resonances to get the kinetic wave equation, which describes nonlinear evolution that is slow compared to linear oscillations with wave frequencies [2–5]. There is a tendency in theoretical statistical physics to restrict consideration to two opposite limits: either treat few modes or infinitely many. That preference is even stronger in the studies of non-equilibrium. And yet not only most of the real-world phenomena fall in between these limits, but, as we show here, one learns some fundamental lessons comparing equilibrium and non-equilibrium states of systems with a finite number of degrees of freedom, where phase coherence can play a prominent role. A similar lesson condensed matter physics taught us by discovering the world of mesoscopic phenomena, where the system size was made smaller than the phase coherence length.

The previous treatment of mode discreteness was focused on the sparseness of resonances for the particular cases when resonant surfaces $\omega_k+\omega_q=\omega_{|\mathbf{k}+\mathbf{q}|}$ did not pass through integer lattice determined by a box [5, 6]. Yet in many cases resonance surfaces lay in the lattice. For example, in a quite generic case of quadratic dispersion relation, $\omega_k\propto k^2$, Pythagorean theorem makes the resonance surface for three-wave interactions just perpendicular to any wavevector, so that in any rectangular box resonantly interacting triads fill the lattice of the box eigen modes.

Class of models (1,2,4) is ideally suited for the comparative analysis of thermal equilibrium and turbulence. We show here that such analysis sheds light on the most fundamental aspects of turbulence, particularly the roles of spatial and temporal scales in determining cascade directions and build-up of intermittency. We consider the particular sub-class of models that allow only neighboring interactions, and find it the most versatile tool to date to study turbulence as an ultimate far-from-equilibrium state. We carry here such detailed study of the known types of direct-only and double cascades with unprecedented numerical resolution. Even more important, our models allow for an inverse-only cascade never encountered before.

II. FIBONACCI TURBULENCE

We consider a sub-class of the models (1,2,4) which is Hamiltonian with a local interaction:

$$\mathcal{H}=\sum\nolimits_i V_i\left(a_i^*a_{i+1}^*a_{i+2}+a_ia_{i+1}a_{i+2}^*\right). \tag{5}$$

The equations of motion $\imath\dot a_i=\partial\mathcal{H}/\partial a_i^*$ are as follows:

$$\imath\dot a_i=V_{i-2}a_{i-1}a_{i-2}+V_{i-1}a_{i-1}^*a_{i+1}+V_ia_{i+1}^*a_{i+2}. \tag{6}$$

This family of models (each characterized by V_i) can have numerous classical and quantum applications, since i can be denoting real-space sites, spectral modes, masses of particles, number of monomers in a polymers, etc. The Hamiltonian describes, in particular, decay and coalescence of waves or quantum particles, breakdown and coagulation of particles or polymerization of polymers, etc, when interactions of comparable entities are dominant. In particular, the model describes the resonant interaction of waves whose frequencies are the Fibonacci numbers $F_i=\{1,1,2,3,5\ldots\}$ defined by the identity $F_i+F_{i+1}=F_{i+2}$ with $F_0=0$. Indeed, such waves are described by the Hamiltonian

$$\mathcal{H}_0=\sum\nolimits_i\left[F_i|a_i|^2+V_i\left(a_i^*a_{i+1}^*a_{i+2}+a_ia_{i+1}a_{i+2}^*\right)\right]. \tag{7}$$

The first term corresponds to the linear terms in the equations of motion, while the second term represents the only possible resonant interactions, since no non-consecutive Fibonacci numbers sum into another Fibonacci number (Zeckendorf theorem). For any real t, the Hamiltonian (7) is invariant under the $U(1)\times U(1)$ transformation $a_i\to a_ie^{\imath F_it}$ due to $F_i+F_{i+1}=F_{i+2}$. The transformation (to the wave envelopes) reduces the equation of motion $\dot a_i=\partial\mathcal{H}_0/\partial a_i^*$ to (6).

If i are spectral parameters, they are usually understood as shell numbers. That means that one can *define* wave numbers as $k=F_i=[\phi^i-(-\phi)^{-i}]/\sqrt{5}$, where $\phi=(1+\sqrt{5})/2$ is the golden mean. It plays here the role of an intershell ratio, since asymptotically at $|i|\gg1$, the wave number depends exponentially on the mode number: $F_i\propto\phi^{|i|}$. The model (6) thus belongs to the class of the so-called shell models [7], that is (2) with neighboring interactions. Coefficients of shell models are chosen to have one or two quadratic integrals of motion. In particular, the Sabra shell model [8, 9] for a particular choice of coefficients (non-surprisingly, connected by the golden ratio) coincides with (6), which is Hamiltonian and has the cubic integral of motion (5).

It is straightforward to show that for arbitrary V_i, the dynamical equations (6) conserve a one-parameter family of quadratic invariants (generalizations of the Manley-

Rowe invariants for three-wave interactions):

$$\mathcal{F}_k = \sum\nolimits_{i=1} F_{i+k-1}|a_i|^2 \,, \tag{8}$$

where k could be of either sign if we define negative Fibonacci numbers: $F_{-j} = (-1)^{j+1}F_j$. All invariants can be obtained as linear combinations of any two of them. For example, the first two integrals are positive, independent, and in involution:

$$\mathcal{F}_1 = \sum\nolimits_{i=1} F_i|a_i|^2 \,, \quad \mathcal{F}_2 = \sum\nolimits_{i=1} F_{i+1}|a_i|^2 \;. \tag{9}$$

In a closed system, the microcanonical equilibrium is $\mathcal{P} = \delta(\mathcal{H} - C)\delta(\mathcal{F}_1 - C_1)\delta(\mathcal{F}_2 - C_2)$. We now add dissipation and white-in-time pumping:

$$\dot{a}_i = -\imath\partial\mathcal{H}/\partial a_i^* + \xi_i - \gamma_i a_i \;. \tag{10}$$

Here $\langle\xi_i a_j^*\rangle = \delta_{ij}P_i/2$. It is straightforward to show, also in a general case (3,4), that such forcing on average does not change the cubic Hamiltonian, since $\langle\xi_i a_{i+1}a_{i+2}^*\rangle = P_i\langle\partial(a_{i+1}a_{i+2}^*)/\partial a_i^*\rangle = 0$ for any i. Denoting $\mathcal{H}_i = 2\mathrm{Re}(V_i a_i^* a_{i-1}a_{i-2})$, we then obtain $\sum_i d\langle\mathcal{H}_i\rangle/dt = -\sum_i(\gamma_i + \gamma_{i-1} + \gamma_{i-2})\langle\mathcal{H}_i\rangle$, which must be zero in a steady state. At least when all sums $\gamma_i + \gamma_{i-1} + \gamma_{i-2}$ are the same, $\sum_i\langle\mathcal{H}_i\rangle = \langle\mathcal{H}\rangle = 0$ (one can probably imagine exotic cases where separate $\langle\mathcal{H}_i\rangle \neq 0$ but we shall not consider them). If pumping and damping are in a detailed balance, so that $\sum_k \alpha_k F_{i+k-1} = \gamma_i/P_i$ for every i, the thermal equilibrium distribution is Gaussian: $\mathcal{P} = \exp(-\sum_k \alpha_k\mathcal{F}_k)$ — it is a steady solution of the Fokker-Planck equation:

$$\begin{aligned}\partial_t\mathcal{P} = \{\mathcal{P},\mathcal{H}\} + \sum\nolimits_i \Big[P_i\partial_{a_i}\partial_{a_i^*} + \gamma_i\big(\partial_{a_i}a_i + \partial_{a_i^*}a_i^*\big)\Big]\mathcal{P}\\ \propto \sum\nolimits_i \big(2\gamma_i - P_i\sum\nolimits_k \alpha_k F_{i+k}\big) = 0\,.\end{aligned}$$

That solution realizes maximum entropy for given values of the invariants. The distribution is exactly Gaussian despite the system being described by a cubic Hamiltonian and thus strongly interacting. The only restriction on the numbers α_k is normalization. In particular, when only $\alpha_1 = 1/2T$ is nonzero, we get the equilibrium equipartition with the occupation numbers $n_i \equiv \langle|a_i|^2\rangle = P_i/2\gamma_i = T/F_i$.

In a turbulent cascade, the fluxes of the quadratic invariants can be expressed via the third cumulant. Gauge invariance and Zeckendorf theorem ensure that the triple cumulants are nonzero only for consecutive modes in the inertial range:

$$J_i \equiv \mathrm{Im}\,\langle a_i^* a_{i-1}a_{i-2}\rangle \,, \tag{11}$$

$$\begin{aligned}F_{i+k-1}\frac{d\langle|a_i|^2\rangle}{dt} &= 2F_{i+k-1}(V_{i-2}J_i - V_{i-1}J_{i+1} - V_iJ_{i+2})\\ &= \Pi_k(i-1) - \Pi_k(i) = -\partial_i\Pi_k(i) \;.\end{aligned} \tag{12}$$

The right hand side is the discrete divergence of the flux

$$\begin{aligned}\Pi_k(m) &\equiv -\sum_{i=1}^{m} F_{i+k-1}\frac{d\langle|a_i|^2\rangle}{dt}\\ &= 2F_{m+k}V_{m-1}J_{m+1} + 2F_{m+k-1}V_mJ_{m+2} \;.\end{aligned} \tag{13}$$

The 3rd order cumulants are zero in equilibrium, but in turbulence they are nonzero to carry the flux. In the inertial interval, the flux must be constant and its divergence zero. For our class of models, we are able to find analytically the form of the 3rd cumulant (the analog of Kolmogorov's 4/5-law for fluid turbulence):

$$J_m = CF_{M-m+1}/V_{m-2} \;, \tag{14}$$

where real constant C and integer M can be of either sign. Let us substitute (14) into (13) and show that all the fluxes are non-zero constants independent of m:

$$\begin{aligned}\Pi_k(m) = 2F_{m+k}V_{m-1}CF_{M-m}/V_{m-1}\\ + 2F_{m+k-1}V_mCF_{M-m-1}/V_m = CF_{M+k-1} \;.\end{aligned} \tag{15}$$

The last equality follows from the Cassini identity: $F_mF_n + F_{m-1}F_{n-1} = F_{m+n-1}$. All the fluxes have the same sign for any k, that is all the integrals $\mathcal{F}_k$ flow in the same direction for such solutions. We shall show in the next section what kind of fine-tuning is needed to get a double cascade when both cascades carry the same integrals. In [8], the (quadric) spectral flux of the (cubic) Hamiltonian was also defined, but pumping does not produce it, so that $\langle\mathcal{H}\rangle = 0$ in a steady turbulent state, as well as in thermal equilibrium.

Every model of our family is completely characterized by specifying the dependence of V_i on i. While thermal equilibrium does not depend on V_i and is universal for the whole family, turbulence depends on V_i, as clear from (14). In what follows, we shall consider the power-law dependence $V_i = F_i^\alpha$, which turns into exponential dependence $V_i \approx \phi^{i\alpha}$ for $i \gg 1$. Therefore, the single real parameter α determines the model. Our choice of particular values for α below will make the connection between wave and hydrodynamical turbulence through the Fibonacci model more explicit.

III. CASCADE DIRECTION

To get an analytic insight into our turbulence, particularly, to understand the flux direction, consider an invariant sub-space of solutions with purely imaginary $a_k = i\rho_k$ for all k:

$$\frac{\partial\rho_i}{\partial t} = V_{i-2}\rho_{i-1}\rho_{i-2} - V_{i-1}\rho_{i-1}\rho_{i+1} - V_i\rho_{i+1}\rho_{i+2} \tag{16}$$

In this case, $\mathcal{H} \equiv 0$. The invariant subspace owes its existence to the invariance of (6) with respect to the symmetry $a \to -a^*$.

Consider the chain running between some integers M

and N, either positive or negative, and assume $V_i/V_{i-1} = \phi^\alpha$. Then for $\rho_i = A\phi^{i\beta}$ and $M+1 < i < N-1$ we obtain:

$$\frac{\partial \rho_i}{\partial t} = A^2 V_{i-2}\phi^{2i\beta}\left(\phi^{-3\beta} - \phi^{\alpha} - \phi^{2\alpha+3\beta}\right) \ . \tag{17}$$

The right hand side of (17) turns into zero for $\beta = -(1+\alpha)/3$, which defines a steady solution $\rho_i = \phi^{-i(1+\alpha)/3}$ (also with the replacement $\phi \to -1/\phi$). This solution can describe either direct or inverse cascade, since the symmetry $\rho \to -\rho$, $t \to -t$ means that one reverses the flux by changing the sign of ρ in this case. Indeed, consider the evolution from the initial state where all amplitudes are zero except the first two ρ_M, ρ_{M+1}. The first term in (16) then will produce ρ_{M+2} of the same sign as $V_M\rho_M\rho_{M+1}$, which makes the flux positive, as it should be for a direct cascade. Alternately, by pumping the last two modes, the last term of (16) produces a negative flux. Which cascade can be realized in reality: direct, inverse or both? Physically it is clear that the sign of the flux must be determined by the only parameter α, that is by how mode interaction depends on the mode number. Indeed, for $\alpha = 1/2$, the scaling of the flux steady solution coincides with that of the thermal equilibrium: $\langle\rho_i\rangle = 0$, $\langle\rho_i\rho_j\rangle = n_i\delta_{ij} = \delta_{ij}T/F_i \propto \phi^{-i}$, for $i \gg 1$. Such state can be excited, for instance, by an imaginary pumping acting on every mode in detailed balance with dissipation. Physical common sense suggests that the cascade must carry the conserved quantity $\sum_i F_i\rho_i^2$ from excess to scarcity [3, 10]. For $\alpha > 1/2$ the steady solution $\rho_i^2 = \phi^{-2(1+\alpha)i/3}$ decays with i faster than the equipartition $\rho_i^2 \propto 1/F_i \propto \phi^{-i}$, so that it must correspond to a direct cascade. By the same token, we must have an inverse cascade for $\alpha < 1/2$. Of course, such consideration is a plausible argument, not a rigorous proof of the cascade sign. Getting a little ahead of ourselves, mention here that we observe a double-cascade turbulence exactly at $\alpha = 1/2$.

In a general complex case, arguing that the cascade changes direction when α crosses $1/2$ is even less straightforward. The flux constancy determines the third moment, which only bounds the product of the second and fourth moments (the claim that it bounds the square root of the products of three second moments made in [11] is incorrect). Yet a plausible argument can be made as follows. The input rate of $\mathcal{F}_k$ is equal to $\Pi = PF_{p+k-1}$ where p is the position of the pumping. The input rate must be equal to the dissipation rate $\Pi = 2\gamma_d F_{d+k-1} n_d$ for any choice of γ_d taken at the dissipation position d. In order for n_d to smoothly match the cascade, one must choose γ_d comparable to the nonlinear interaction time: $\gamma_d \simeq V_d J_d^{1/3} \simeq V_d(\Pi/V_dF_d)^{1/3}$. This gives an order-of-magnitude estimate $n_d \simeq (\Pi/V_dF_d)^{2/3}$. Such reasoning can be applied to every i, which in turn gives the estimate for the spectrum of occupation numbers:

$$n_i \simeq (\Pi/V_iF_i)^{2/3} \, . \tag{18}$$

Since the direction of the flux is toward the occupation numbers that are lower than thermal equilibrium, $n_i \propto F_i^{-1}$, then again we see that the flux changes direction when $V_i \propto F_i^{1/2}$. The dimensionless degree of non-Gaussianity on such a spectrum,

$$\xi \equiv \frac{J_i}{n_i^{3/2}} \simeq \frac{\Pi}{V_iF_in_i^{3/2}} \simeq \frac{PF_p}{V_iF_in_i^{3/2}} \, , \tag{19}$$

must be independent of i. For the spectrum close to equilibrium, $\xi \propto F_i^{3/2}/V_iF_i = F_i^{1/2}/V_i$.

Figures 1 and 2 confirm these predictions. We place the pumping at a single mode, $i = p$, between two dissipation regions on the ends, letting the system to choose the cascade direction. The system (10) with pumping and damping has been evolved numerically using `LSODE` solver [12]. At each step, random Gaussian noise of power P is applied to the pumping-connected mode injecting flux $\Pi_p = PF_p$. Damping with γ_L and γ_R is applied to the two left-most and two right-most modes respectively. For $\alpha = 1/2$ ($V_i = \sqrt{F_i}$), the system is weakly distorted from equilibrium, with a constant flux on each side of the pumping. For $\alpha \neq 1/2$ we find that the invariants are absorbed only on one end of the spectrum. For $\alpha > 1/2$ ($V_i = F_i$), we have a thermal equilibrium to the left of pumping and the direct cascade (18) with a constant ξ to the right. In the opposite case ($\alpha < 1/2$, V_i =const), we find an inverse cascade (18) with constant ξ to the left and equilibrium equipartition to the right of pumping. In both cases, the damping on the flux side is carefully selected to avoid build-up in the spectrum (the damping on the equilibrium side can be then set to zero to establish cleaner scaling). We have chosen $V_i = F_i$ and $V_i =$ const because they qualitatively correspond to the Kolmogorov scaling of the direct energy cascade in incompressible turbulence and to the inverse wave action cascade in deep water turbulence respectively.

Thermal equilibrium at the scales exceeding the pumping scale together with a direct cascade at smaller scales have been predicted and observed [13]. To the best of our knowledge, nobody has seen before an inverse-only cascade together with a thermal equilibrium on the other side of the pumping, neither in hydrodynamic-type systems nor in wave turbulence or shell models. Inverse cascades play a prominent role in geophysics and astrophysics, from creation of planetary jets to Jupiter Great Red Spot and stormy seas. In all known cases inverse cascades appear in systems with at least two conserved quantities that scale differently. All our conserved quantities (8) scale the same in the limit $i \gg 1$. Probably closest to our findings are the results of Tom and Ray [14] who observed an inverse cascade in the limiting case of a shell model with two invariants having the same scaling. Their inverse cascade had normal scaling and run from fast to slow modes; the direct cascade was not resolved, but was likely present.

Our observation poses the question: can one find another class of systems with a single conservation law and

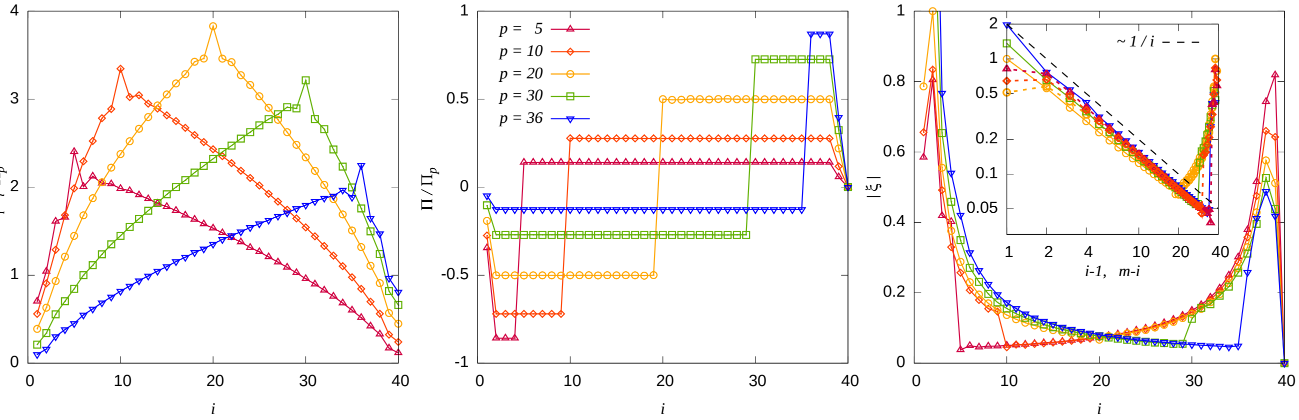


FIG. 1. Compensated spectra, fluxes and skewness for $\alpha = 1/2$ with different pumping locations: $p = 5$, 10, 20, 30, and 36 on the 40 mode interval. Pumping rate is selected to provide the same flux in all cases, $\Pi_p = 67.65$. In all cases damping rates are $\gamma_L = \gamma_R = 1$. Inset reproduces the longer arm of the cascades in log-log scale.

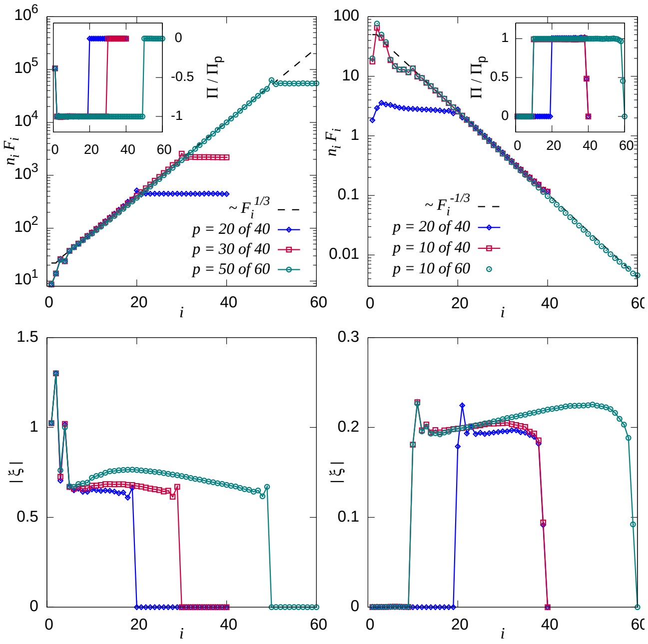


FIG. 2. Compensated spectra, fluxes, and the dimensionless skewness for $\alpha = 0$ (left) and $\alpha = 1$ (right) for systems with different location of pumping. In all cases $\Pi_p = 67.65$. For $\alpha = 0$ damping rates are $\gamma_L = 1.5$ and $\gamma_R = 0$; for $\alpha = 1$ damping rates are $\gamma_L = 0$ and $\gamma_R = 140$ at mode 40 and $\gamma_R = 3500$ at mode 60.

the turbulent spectrum less steep than equilibrium. In weak wave turbulence, this requires the sum of the space dimensionality and the scaling exponent of the three-wave interaction to be less than the frequency scaling exponent [3]. We do not know such a physical system, nor we aware of any fundamental law that forbids its existence. Remark that the connection between the cascade direction, its stability and steepness relative to equipartition has been firmly established in the weak turbulence theory [3, 10]. In all known examples, the formal turbulent solution with a wrong flux sign is not realized; the system chooses instead to stay close to equipartition with a slight deviation that provides for the flux in the right direction [3, 15]. Similarly, when we place pumping and damping at the "wrong" ends of a finite chain, our system heats up, staying close to thermal equilibrium.

It is important that our system is a one-dimensional chain, as well as shell models, so that there is no space and consequently no distinction in the phase volume (number of modes) between infrared and ultraviolet parts of the spectrum. The directions along the chain are only distinguished temporally, i.e. in terms of growth/decay of the typical interaction time. The same combination $V_i^2/F_i \propto \phi^{2\alpha-1}$ determines the i-dependence of the inverse interaction time both for the equilibrium, $V_i b_i^{1/2} = V_i F_i^{-1/2} T^{1/2}$, and for a cascade, $V_i(\Pi/V_iF_i)^{1/3} = (V_i^2/F_i)^{1/3}\Pi^{1/3}$. As the above consideration shows, the cascade proceeds from slow modes to fast modes in Fibonacci turbulence. Similarly in shell models [11, 16, 17] (albeit with parameters and conservation laws distinct from our model), a cascade proceeding from fast modes to slow modes was never observed. It was argued that this is because the fast modes act like thermal noise on the slow ones, which must lead to equilibrium [16]. That this cannot be generally true follows from the existence of the inverse energy cascade in 2D incompressible turbulence and from numerous examples in weak wave turbulence where non-linear interaction time either grows or decays along the cascade. Moreover, the formation of the cascade spectrum proceeds from fast to slow modes (and not necessarily from pumping to damping), according to the information-theory argument [18].

Why is the flux direction unambiguously related to the cascade acceleration in shell models in general and in our model in particular, in distinction from other cases? The argument can be made by considering capacity, a measure that tells at which end the conserved quantity is stored — perturbations are known to run towards that end [3]. For example, the power-law energy density spectrum $\epsilon_k \propto k^{-s}$ in d dimensions has the total energy $\int \epsilon_k \, d^d k$ — at which end it diverges is determined by the sign of $d-s$. This is generally unrelated to the direction of the energy cascade, determined by the sign of s, which tells whether the spectrum is more or less steep than the equipartition. However, in shell models the exponential character of i-dependencies makes the total energy $\sum_i F_i |a_i|^2$ determined by either the last or the first term of the sum, which solely depends on whether $F_i|a_i|^2$ is steeper than equipartition or not, that is by the sign of the flux.

Which direction then the cascade goes in the symmetric case, $V_i = \sqrt{F_i}$? Now the naive cascade solution (18) coincides with thermal equipartition, $F_i n_i$ =const, and the interaction time is independent of the mode number for such n_i. If we start from thermal equilibrium and apply pumping to some intermediate mode, the system develops cascades in both directions. The left panel of the Figure 1 shows that the pumping at site p inside the interval $(1, N)$ generates left and right fluxes in the proportion $\Pi_L/\Pi_R \simeq (N-p)/p$. This seems natural as in the shorter interval the steeper spectrum falls away from the pumping, which must correspond to a larger flux. This means that if we want to keep the flux constant while increasing p or $N-p$, we need to keep constant the ratio $(N-p)/p$.

We end this section with a general remark. Fibonacci Hamiltonian is not symmetric with respect to reversing the order of modes, it sets the preferred direction, which is physically meaningful since the frequencies of two lower modes sum into the frequency of a high one. Yet, as we see in the case $V_i F_i^{-1/2}$ =const, direct and inverse cascades are pretty symmetric. So, it is natural to conclude that indeed the i-dependence of $V_i F_i^{-1/2}$ determines which way cascade goes.

IV. ALONG THE CASCADES AND AWAY FROM EQUILIBRIUM

As we have seen, thermal equilibrium statistics is exactly Gaussian with no correlation between modes, despite strong interaction (which actually establishes equipartition). The reason for the absence of correlation is apparently the detailed balance that cancels them. We do not expect such cancelations in non-equilibrium states. In all cases of strong turbulence known before, the degree of non-Gaussianity increases along a direct cascade and stays constant along an inverse cascade [19, 20]. As we shall show now, non-Gaussianity always increases along the cascades in our one-dimensional chains.

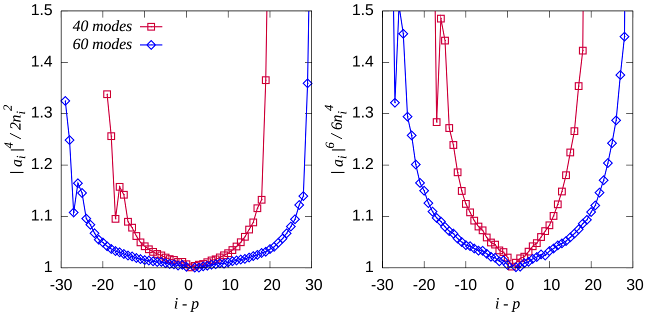


FIG. 3. Fourth and sixth moments for $\alpha = 1/2$ and center pumping in 40-mode system, with $\gamma_L = \gamma_R = 3$, $P = 0.1$, and in 60-mode system with $\gamma_L = \gamma_R = 30$, $P = 1$.

We present first the symmetric case, where the system is close to the equilibrium equipartition with the temperature set by pumping and slowly changing with the mode number: $n_i F_i \approx (PF_p)^{2/3} f(i)$. The slow function $f(i)$ can be suggested by the analogy with the 2D enstrophy cascade [21, 22] as $f(i) \propto \ln^{2/3} F_i \propto i^{2/3}$, counting from the damping region. This gives the dimensionless cumulant $\xi \propto 1/i$. This hypothesis is supported by the right panel of the Figure 1, which shows that ξ grows along both cascades by a power law in i rather than exponentially. Let us stress that count always starts from the dissipation region, where we have the balance condition $\Pi = \gamma_d F_{d+k-1} n_d$ and where $\gamma_d \simeq V_d J_d^{1/3} \simeq V_d(\Pi/V_d F_d)^{1/3}$ according to the dynamical estimate. This sets the nonlinearity parameter of order unity at the damping region and decaying towards pumping; the longer the interval, the smaller is ξ at any fixed distance from the pumping region. The limit of long intervals may then be amenable to an analytical treatment. Indeed, Figure 3 demonstrates that as the interval increases, the higher cumulants remain small over longer and and longer intervals starting from pumping. Despite the model having ultra-local interactions (every mode participates in only three adjacent interacting triplets), the cascade formation is very nonlocal. It is somewhat similar to thermal conduction: if we keep the flux but increase the distance, the distribution gets closer to the thermal equilibrium at every point.

Turning to asymmetric (one-cascade) cases, we see the cumulants higher than third growing with F_i by a power law instead of logarithmic. Rather than look for scaling in the mode number i, we find it more natural to use F_i (playing the role of frequency); at large i one has $F_i \approx \phi^i$, where ϕ is the golden mean. Traditional study of turbulence in general and shell models in particular was focused on the single-mode moments (analog of structure functions), $\langle |a_i|^q \rangle \propto F_i^{-\zeta_q}$, whose anomalous scaling exponents, $\Delta(q) = q\zeta_3/3 - \zeta_q$ give particular measures of how non-Gaussianity grows along the cascade. For $V_i = F_i^\alpha$, the flux law gives $J_i \propto \Pi/V_i F_i$, that is $\zeta_3 = \alpha + 1$. The anomalous scaling is observable in numerics for the single-cascade cases $\alpha = 0$ and $\alpha = 1$, as shown in the right panel of the Figure 6. This seems to be the first case of an anomalous scaling in an inverse cascade, with the anomalous dimensions having the opposite signs to those in direct cascades. The exponents start fairly small but grow fast with q. The anomalous

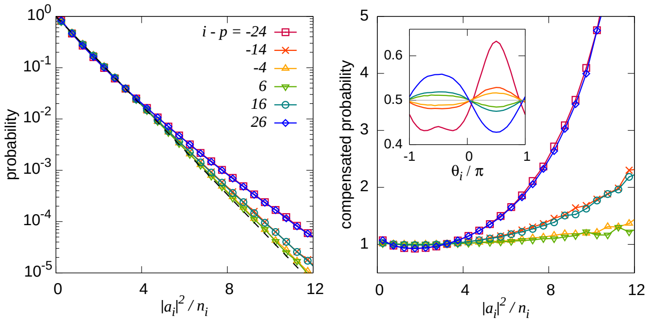


FIG. 4. Probability (left) and deviation of probability from equilibrium (right) for $\alpha = 1/2$. Main panels show probabilities of occupation numbers rescaled to their averages, the inset shows the probability of phase difference, $\theta_i = \varphi_i - \varphi_{i-1} - \varphi_{i-2}$. Refer to the first panel for the line color for different modes. Data are shown for 60-mode system with center pumping and $\gamma_L = \gamma_R = 30$, $P = 1$.

exponents, $\Delta(q)$, can be related to the statistical Lagrangian conservation laws [23, 24] in fluid turbulence; no comparable physical picture was developed for shell models. Without physical guiding, the set of the anomalous exponents is not very informative, all the more that they characterize only one-mode distribution.

Here we suggest a complementary set of three information-theoretic measures, which shed a new light on the turbulent statistics emerging along the cascade. The main distinction of any non-equilibrium state is that it has lower entropy than the thermal equilibrium at the same energy. Turbulence has the entropy that is much lower, which means that a lot of information is processed to excite the turbulence state. We pose the question: where is the information that distinguishes turbulence from equilibrium encoded?

V. WHERE IS THE INFORMATION ENCODED?

First, the information is encoded in a single-mode statistics, which is getting more non-Gaussian deeper in the cascade. This must be reflected in the decay of the one-mode entropy, $S_i = S(x_i) = S(|a_i|/\sqrt{n_i})$, with the growth of $|i-p|$. This can be computed using the multi-fractal formalism: the moments $\langle x_i^q \rangle \propto F_i^{-\zeta_q + q\zeta_2/2}$ in the limit of large $|i-p|$ correspond to the multi-fractal distribution,

$$\mathcal{P}(x_i) \propto \int g(x_i/F_i^h) x_i^{-1} \exp[f(h)\ln F_i]\, dh\,, \tag{20}$$

where $f(h) = \min_q(\zeta_q - q\zeta_2/2 - qh)$, that is $f(h)$ is the Legendre transform of $\zeta(q)$. The entropy is then

$$S_i = -\int dx \mathcal{P} \ln \mathcal{P} \propto [\Delta'(0) - \Delta_2/2] \ln F_i\,.$$

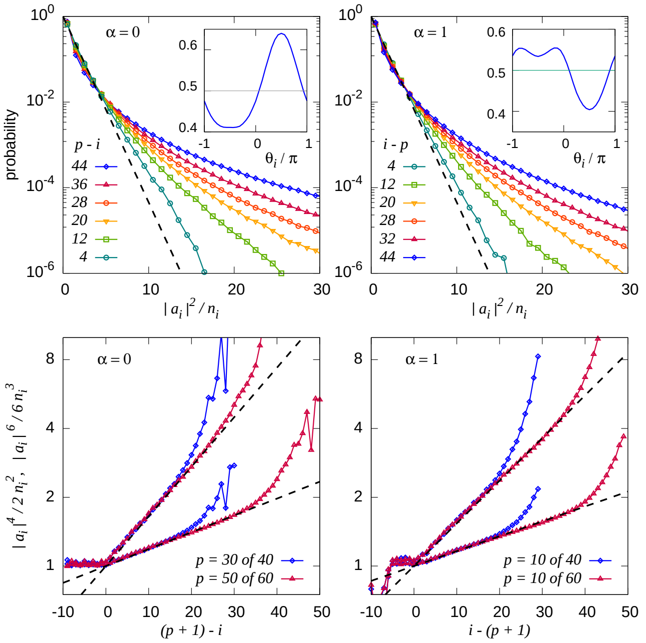


FIG. 5. Probabilities (top) and forth and sixth moments (bottom) for the inverse cascade, $\alpha = 0$ (left), and the direct cascade, $\alpha = 1$ (right). Probabilities for the rescaled occupation numbers are shown in the main panels, while probabilities for the phase difference, $\theta_i = \varphi_i - \varphi_{i-1} - \varphi_{i-2}$, are shown in the insets. The variation between $\mathcal{P}(\theta_i)$ for different i is minor. In all cases $\Pi_p = 67.65$. For $\alpha = 0$, the damping rates are $\gamma_L = 1.5$ and $\gamma_R = 0$; for $\alpha = 1$ the damping rates are $\gamma_L = 0$ and $\gamma_R = 140$ at $i = 40$ and $\gamma_R = 3500$ at $i = 60$. In the top panels the dashed lines indicate the Gaussian probability; in the bottom panels the dashed lines show linear fits to the data.

This decay is logarithmic in frequency F_i, that is linear in i, as indeed can be seen in Figure 6, where i is counted from pumping. Noticing that $\Delta_1 \approx \Delta_2$ and assuming quadratic dependence for $q \leq 3$, we estimate $\Delta'(0) \approx 3\Delta_1/2$ and observe that the dashed lines in the right panel of Figure 6 with the slopes $\Delta_1 \ln \phi$ by the order of magnitude represent the entropy decay in the inertial interval in both direct and inverse cascades.

Second, the information is encoded in the correlations of different modes. It is natural to assume that correlations are strongest for modes in interacting triplets, a_i, a_{i+1}, a_{i+2}. Disentangling of information encoded can be done by using structured groupings [25–27]:

$$\begin{aligned}&\textstyle\sum_{i=1}^n S(a_i) - \sum_{ij} S(a_i, a_j) + \sum_{ijk} S(a_i, a_j, a_k) \qquad (21)\\ &- \textstyle\sum_{ijkl} S(a_i, a_j, a_k, a_l) + \ldots + (-1)^{n+1} S(a_1, \ldots, a_n)\,.\end{aligned}$$

For $n = 1$, this gives the one-mode entropy S_i which measures the total amount of information one can obtain by measuring or computing one-mode statistics. While the entropy itself depends on the units or parametrization, all the quantities (22) for $n > 1$ are independent of units and invariant with respect to simultaneous re-

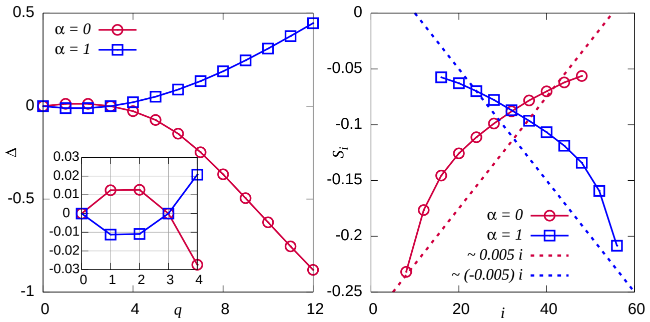


FIG. 6. Left panel: Anomalous exponents computed as $\Delta(q) = q\zeta_3/3 - \zeta_q$. Right panel: Decay of entropy down the cascade for the one-mode complex amplitude normalized by $\sqrt{n_i}$. The dashed lines $S_i - S_p \approx -0.005|i-p|$ have the slopes equal to $\Delta_1 \ln(\phi)$ with Δ_1 shown in the left panel. Direct cascade - blue, inverse cascade - red.

parametrization of every single variable. For $n = 2$, we have the widely used mutual information,

$$I_{ij} = S(a_i) + S(a_j) - S(a_i, a_j)\,,$$

which measures the amount of information one can learn about one mode by measuring another, that is characterizes the correlation between two modes. It is interesting that all pairs in the triplet have comparable mutual information in the direct cascade ($V_i = F_i$), while $I_{i,i+1}$ exceeds noticeably $I_{i,i+2}$ in the inverse cascade ($V_i = 1$), see the upper right panel in Figure 8. One can also define the total (multi-mode) mutual information as the relative entropy between the true joint distribution and the product distribution: $I(a_1, \ldots, a_k) = \sum_{i=1}^{k} S(a_i) - S(a_1, \ldots, a_k)$. It is positive and monotonically decreases upon averaging over any of its arguments. As we see from Figure 8, the changes along the cascade in one-mode entropy and in two-mode and three-mode mutual information are comparable, that is one obtains comparable amount of information about turbulence from these quantities.

To see how much more information one gets by measuring or computing the three modes simultaneously compared to separately by pairs, one needs to use the measure of the irreducible information encoded in triplets, as given by the third member of the hierarchy (22):

$$\begin{aligned} II_i =& S(a_i) + S(a_{i+1}) + S(a_{i+2}) + S(a_i, a_{i+1}, a_{i+2}) \\ & - S(a_i, a_{i+1}) - S(a_i, a_{i+2}) - S(a_{i+1}, a_{i+2}) \\ =& I_{i,i+1} + I_{i,i+2} + I_{i+1,i+2} - I_{i,i+1,i+2} \\ =& I(i, i+1) - I(i, i+1|i+2)\,, \end{aligned} \tag{22}$$

It is called interaction information in the classical statistics and topological entanglement entropy in the quantum statistics [25, 28]. Interaction information measures the influence of the third variable on the amount of information shared between the other two and could be of either sign. Positive $II(X, Y, Z)$ measures the redundancy in the information about Y obtained by measuring X and Z separately, while negative one measures synergy which is the extra information about Y received by knowing X and Z together. While we cannot prove it mathematically, it seems physically plausible that systems with three-mode interaction must demonstrate synergy. Indeed, one finds a strong synergy in weak turbulence: it was shown that $I_{123} \gg I_{12} + I_{23} + I_{13}$ [18], so that $II < 0$ and much more information is encoded in three modes than in the pairs separately. Here we find that the same is true for the cascades close to thermal equilibrium at $V_i = \sqrt{F_i}$ as seen in Figure 7. Indeed, the two-mode mutual information is much smaller than both the one-mode entropy and the absolute value of the interaction information, which is negative.

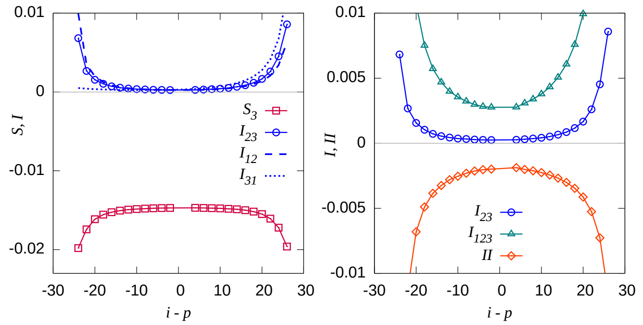


FIG. 7. Deviation of entropies from equilibrium, mutual information, and interaction information for $\alpha = 1/2$ and center pumping for a set of $5 \cdot 10^7$ data point. The same values of entropy were obtained for a set of $2 \cdot 10^7$ data point, that is S_i is saturated. Both I and II show a slight decrease in absolute values with the increase of the ensemble size from $2 \cdot 10^7$ to $5 \cdot 10^7$.

Let us stress that both the mutual information and the interaction information are symmetric, that is they measure the degree of correlation rather than causal relationship or cascade direction.

We compute the entropies and mutual information as follows. First, we obtain the probability distribution in 4D space $(x_{i-2}^2, x_{i-1}^2, x_i^2, \theta_i)$ and integrate it to get corresponding 1D and 2D distributions. Here, $\theta_i = \varphi_i - \varphi_{i-1} - \varphi_{i-2}$, where φ_i is the phase of mode i, and $x_i = |a_i|/\sqrt{n_i}$, while $n_i = \langle |a_i|^2 \rangle$ is the direct average. Mutual information and information interaction are computed directly from entropies, $S = -\Sigma\mathcal{P}\log_2\mathcal{P}$, obtained for these distributions, since all normalization factors cancel out in subtraction. The entropy for an individual mode, however, is presented relative to the Gaussian entropy based on the average occupation number obtained for the binned, staircase distribution for x_i^2. We use the bin sizes $\Delta x_i^2 = 1$ for $\alpha = 0$ and $\alpha = 1$, and $\Delta x_i^2 = 1/2$ for $\alpha = 1/2$. In all cases $\Delta\theta = 2\pi/32$.

Far from equilibrium, we find synergy for the modes close to the pumping and redundancy for damping, see the last panel of Figure 8. That means that the interaction information passes through zero in the inertial interval. There even seems to be a tendency to stick to zero in the inertial interval but this requires further studies with the number of modes exceeding our present abilities. (Our computations are done with a record number of modes, up to 80, while previous studies were mostly done for 20-30. The interaction times decrease exponentially with the mode number, which imposes heavy requirements on the computational time step. On top of that one needs very long runs to collect enough statistics to reliably represent the three-mode probability distribution in four-dimensional space.) With the present set of data we can suggest that most of the information about the three-mode correlation is in the sum of the pair correlations in the triplet. This is more pronounced in the direct cascade than in the inverse cascade. Since the requirements on statistics grow exponentially with the dimensionality, the suggestion that one can get most of information (or at least a large part of it) from lower-dimensional probability distributions is great news for turbulence measurements and modeling. To put it simply, comparable amounts of information can be brought from one-mode and from three-mode measurements in direct and inverse cascades; most of that information can be inferred from two-mode measurements. It remains to be seen to what degree this property of small (asymptotically zero?) interaction information is a universal feature of strong turbulence.

Insets in the Figures 4,5 show the probability distribution of the relative phase, θ_i, which is closely related to the flux (skewness), proportional to $\langle |a_i a_{i-1} a_{i-2}| \sin\theta_i \rangle$. The probability maximum is then at $\pm\pi/2$ for direct and inverse cascades respectively. Also, the i-dependence of the phase distributions is in accordance with the changes in skewness along i. In the two-cascade symmetric case, the distribution is flat (the phases are random) near the pumping, and the phase correlations appear along the cascades, as can be seen comparing the last panel of Figure 1 with the inset in the right panel of Figure 4. In the one-cascade cases, both skewness and the form of the spectrum are practically independent of the mode number, as seen from Figures 2,5.

The fact that the deviations from Gaussianity grow along our inverse cascade, in distinction from all the inverse cascades known before, calls for reflection. We used to think about the anomalous scaling and intermittency in spatial terms: Direct cascades proceed inside the force correlation radius, which imposes non-locality, while in inverse cascades one effectively averages over many small-scale fluctuations, which bring scale invariance [19, 20]. The emphasis on the spatial features was reinforced by the success of the Kraichnan's model of passive tracer turbulence, where it has been shown that the spatial (rather than temporal) structure of the velocity field is responsible for an anomalous scaling and intermittency of the tracer. There is no space in our case, so apparently it is all about time. Indeed, as we have seen, all our cascades propagate from slow to fast modes, which leads to the build-up of non-Gaussianity and correlations. As a result, the entropy of every mode decreases and the inter-mode information grows along the cascade. This diminishes the overall entropy compared to the entropy of the same number of modes in thermal equilibrium with the same total energy.

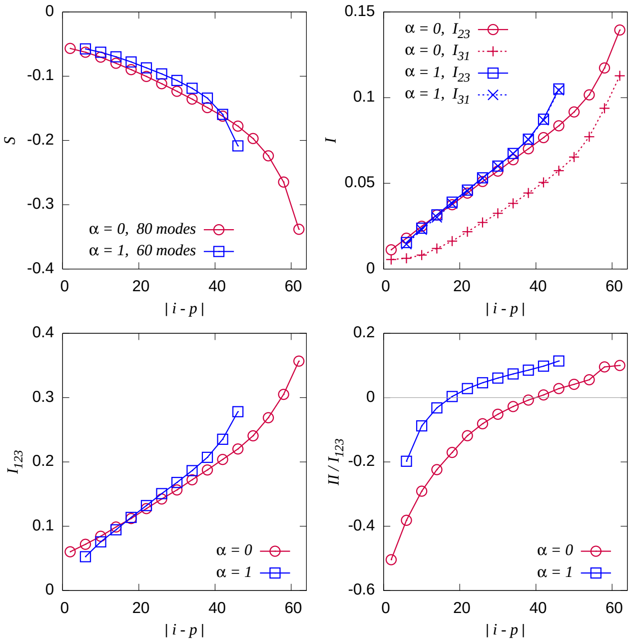


FIG. 8. Deviation of the entropy from equilibrium, the mutual information, and the interaction information (all in bits) for $\alpha = 0$ and $\alpha = 1$ and center pumping. Number of data points $2 \cdot 10^8$ for $\alpha = 0$, 80 modes, $6 \cdot 10^7$ for $\alpha = 0$, 60 modes, and 10^8 for $\alpha = 1$. For the bin size selected, all quantities agree with those obtained in a half-reduced data set.

Despite qualitative similarity, there is a quantitative differences between our direct and inverse cascades. Figures 5,6 show that the one-mode statistics and its moments faster deviate from Gaussian as one proceeds along the inverse cascade than the direct one. And yet one can see from Figures 6,7 that the one-mode entropy is essentially the same in both cascades, as well as the mutual information between two neighboring modes and the three-mode mutual information. The mutual information between non-neighboring modes I_{13} is about twice smaller, as seen in Figure 8. This difference can probably be related to the dynamics, which in our system is the coalescence of two neighboring modes into the next one and the inverse process of decay of one into two. In the dynamical equation (16), only one (first) term is responsible for the direct process (and the direct cascade), while two terms are responsible for the inverse process (and the inverse cascade).

An important distinction between double-cascade and single-cascade turbulence in our system is the dependence on the system size. The degree of non-Gaussianity of the complex amplitudes is fixed in the dissipation regions of the double cascade, so that in the thermodynamic limit the statistics is Gaussian in the inertial intervals. On the contrary, the statistics of the amplitudes is fixed at the forcing scale for a single cascade, and it deviates more and more from Gaussianity as one goes along the cascade.

We end this section by a short remark on the production balance of the total entropy $S = -\langle \ln \rho(a_1, \dots, a_N)\rangle$. Here $\rho(a_1, \dots, a_N)$ is the full N-mode PDF. Since wave interaction does not change the total entropy, then the entropy absorption by the dissipation must be equal to the entropy production by the pumping [18, 29]:

$$P \int \prod_i \frac{da_i da_i^*}{2\rho} \Big|\frac{\partial \rho}{\partial a_p}\Big|^2 = 2 \sum_k \gamma_k, . \tag{23}$$

For a single-cascade cases ($V_i = 1$ and $V_i = F_i$), the energy balance $PF_p = 2\gamma F_d n_d$ means that the left hand side of (23) must be much larger than the Gaussian estimate P/n_p [18]. It may seem to contradict our numerical finding that the pumping-connected mode a_p has its one-mode statistics close to Gaussian. Of course, there are nonzero triple correlation and the mutual information with two neighboring modes in the direction of the cascade. Yet since $\xi \simeq 1$, then the triple moment $J_p \simeq n_p^{3/2}$ both in direct and inverse cascades, so that the contribution to the left hand side of (23) is comparable with P/n_p. We conclude then that even the pumping-connected mode must have strong correlations with many other modes. Since the triple correlation function of non-adjacent modes are zero, such correlations must be encoded in higher cumulants. That deserves further study.

VI. KOLMOGOROV MULTIPLIERS AND SELF-SIMILARITY

Unbounded decrease of entropy along a single cascade prompts one to ask whether the total entropy of turbulence is extensive (that is proportional to the number of modes) or grows slower than linear with the number of modes, so there could be some "area law of turbulence" (like for the entropy of black holes). This question can be answered with the help of the so-called Kolmogorov multipliers, $\sigma_i = \ln |a_i/a_{i-1}|$ [30]. Figure 9 shows that in our cascades the multipliers have universal statistics independent of i, similar to shell models [31–34]. One consequence of the scale invariance of the statistics of the multipliers is that the entropy of the system is extensive, that is proportional to the number of modes. Of course, the entropy depends on the representation. From the information theory viewpoint, the Kolmogorov multipliers realize representation by (almost) independent component, that is allow for maximal entropy. In other words, computing or measuring turbulence in terms of multipliers gives maximal information per measurement (the absolute maximum is achieved by using the flat distribution, that is the variable $u(\sigma)$ defined by $du = P(\sigma)d\sigma$).

The amplitudes are expressed via the multipliers:

$$X_k = \ln x_k = \ln \frac{|a_k|}{\sqrt{n_k}} = \ln x_p + \sum_{i=p+1}^{p+k} \sigma_i + \frac{1}{2} \log \frac{n_p}{n_k} \, .$$

The first term is due to the pumping-connected mode, which correlates weakly with σ_i in the inertial interval. As shown below, the correlation between multipliers decays fast with the distance between them. That suggests that the statistics of the amplitude logarithm at large k must have asymptotically a large-deviation form:

$$\ln \mathcal{P}(X_k) = -kH(X_k/k) \, . \tag{24}$$

Indeed, the three upper curves in the top row of Figure 5 collapse in these variables, as shown in the bottom row of Figure 9. The self-similar distribution of the logarithm of amplitude, (24), is a dramatic simplification in comparison with the general multi-fractal form (20). Technically, it means that $g(x_k/F_k^h) = g(e^{X_k - kh\ln\phi})$ is such a sharp function that the integral in (20) is determined by the single X_k-dependent value, $h(X_k) = X_k/k\ln\phi$. We then identify $f = -H/\ln\phi$.

The self-similarity of the amplitude distribution (plus the independence of the phase distribution on the mode number) is great news, since it allows one to predict the statistics of long cascades (at higher Reynolds number) from the study of shorter ones. In our case, Figure 9 shows that 28-th mode already has the form close to asymptotic. Self-similarity and finite correlation radius of the Kolmogorov multipliers has been also established experimentally for Navier-Stokes turbulence [35]. To avoid misunderstanding, let us stress that the self-similarity is found for the probability distribution of the logarithm of the amplitude, which does not contradict the anomalous scaling of the amplitude moments with the exponents ζ_q determined by the Legendre transform of f or H.

If the multipliers were statistically independent, one would compute $\ln \mathcal{P}(X) = -kH(X/k)$ or ζ_q proceeding from $P(\sigma)$ by a standard large-deviation formalism: $H(y) = \min_z[zy - G(z)]$, where $G(z) = \ln \int d\sigma e^{z\sigma} P(\sigma)$. Such derivation would express $\langle |a_k|^q\rangle$ via $\langle e^{q\sigma_k}\rangle$, which is impossible since the former moments exist for all q, while the latter do not because of the exponential tails of $P(\sigma)$, see also [35, 36].

Therefore, to describe properly the scaling of the amplitudes one needs to study correlations between multipliers. Physically, it is quite natural that the law of the distribution change along the cascade must be encoded in correlations between the steps of the cascade. Indeed, we find that the neighboring multipliers are dependent, albeit weakly, as expressed in their mutual information (traditionally used pair correlation function [32, 33, 35] is not a proper measure of correlation for non-Gaussian statistics). We find that for the inverse cascade, $I(\sigma_i, \sigma_{i+1}) \simeq 0.23$, $II(\sigma_i, \sigma_{i+1}, \sigma_{i+2}) \simeq -0.1$. For the direct cascade, $I(\sigma_i, \sigma_{i+1}) \simeq 0.3$, $II((\sigma_i, \sigma_{i+1}, \sigma_{i+2}) \simeq$

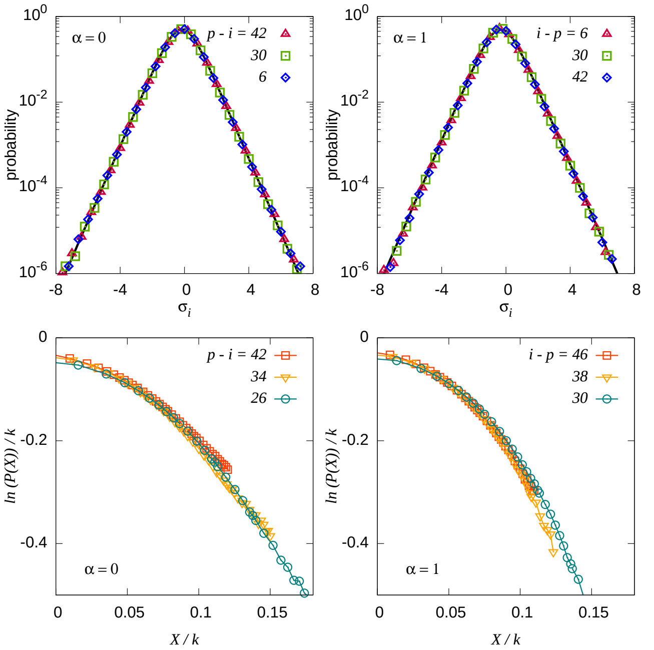


FIG. 9. Top: probability distributions of the Kolmogorov multipliers $\sigma_i = \ln|a_i/a_{i-1}|$ for different positions in the inverse (left) and direct (right) turbulent cascades. Solid lines correspond to the thermal equilibrium $P(\sigma) = 1/2\cosh^2(\sigma - \bar{\sigma})$, where $\bar{\sigma} = -(1/3)\ln\phi$ for the inverse cascade and $\bar{\sigma} = -(2/3)\ln\phi$ for the direct one. Bottom: probability distributions of $X = \ln|a_k|^2$ collapse to the large-deviation form far away from the pumping, that is for large $k = |i - p|$.

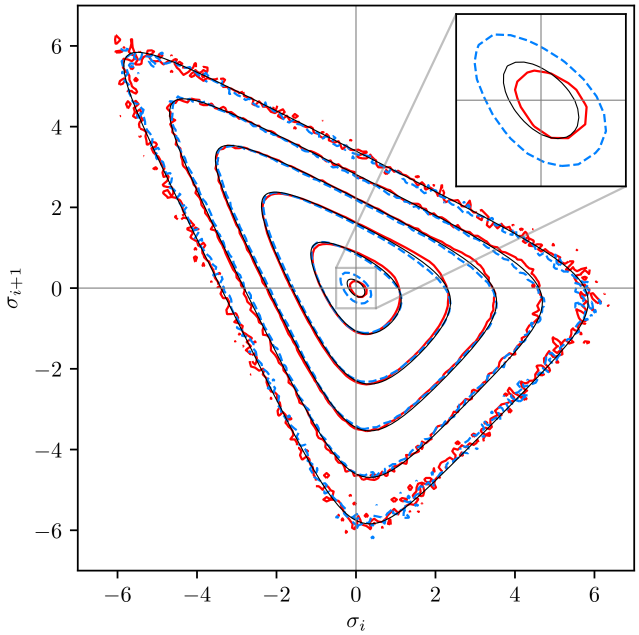


FIG. 10. Joint probability distributions of two neighboring Kolmogorov multipliers shifted to zero means. The contours are at $log_{10}(P) = -0.55, -1, -2, -3, -4, -5$. Inverse cascade ($\alpha = 0$) is red, direct cascade ($\alpha = 1$) is blue, black is the equilibrium distribution (25).

-0.08. No discernible $I(\sigma_i, \sigma_{i+k})$ were found for $k > 1$. While σ_i and σ_{i+2} are practically uncorrelated, there is some small synergy in a triplet.

To appreciate these numbers, let us present for comparison the statistics of the Kolmogorov multipliers in thermal equilibrium. Normalized for zero mean and unit variance, we have

$$P(\sigma) = \int\int_0^\infty dxdy\, e^{-x-y}\delta\left(\sigma - \frac{1}{2}\ln\frac{x}{y}\right) = \frac{1}{2\cosh^2\sigma},$$

$$P(\sigma_i, \sigma_{i+1}) = \frac{8e^{4\sigma_i + 2\sigma_{i+1}}}{\left[1 + e^{2\sigma_i}\left(1 + e^{2\sigma_{i+1}}\right)\right]^3} \ . \qquad (25)$$

That gives $I(\sigma_i, \sigma_{i+1}) = \ln 2 - 1/2 \approx 0.19$.

Figure 9 shows that the equilibrium Gaussian statistics of independent amplitudes perfectly represents the statistics of a single multiplier. The joint PDFs $P(\sigma_i, \sigma_{i+1})$ are shown in Figure 10 for thermal equilibrium and for two cascades. Again, the Gaussian statistics represents turbulence remarkable well. The differences between the three cases are most pronounced around the peak at the origin, while the distant contours are hardly distinguishable. In plain words, the probabilities of strong fluctuations of the multipliers are the same in thermal equilibrium as in turbulence cascades. This is remarkably different from the statistics of the complex amplitudes, which demonstrate most difference between the three cases for strong fluctuations and for high moments. There seems to be a certain duality between fluctuations of the amplitudes and multipliers: strong fluctuations of the multipliers correspond to weakly correlated amplitudes, while strong fluctuations of the amplitudes may require their strong correlations and thus correspond to multipliers close to their mean values. Whether this duality can be exploited for an analytic treatment remains to be seen. The information about the anomalous scaling exponents of the amplitudes in turbulence must be encoded in the correlations between multipliers. Note that the mutual information $I(\sigma_i, \sigma_{i+1})$ for both cascades ($I = 0.23$ and $I = 0.30$) is not that much higher than in thermal equilibrium ($I = 0.19$ bits). Physicists tend to be much excited about any broken symmetry; it is refreshing to notice that relatively small information is needed to encode the broken scale invariance in turbulence. How to decode this information from the joint statistics of multipliers remains the task for the future

VII. DISCUSSION

The most surprising finding of our work is the existence of an inverse-only cascade and its anomalous scaling. In all cases known before, an inverse cascade appears only as an outlet for an extra invariant that cannot be transferred along the direct cascade with other invariant(s). In a truly weak turbulence, when the whole statistics is close to Gaussian, an inverse-only cascade is indeed impossible, since it would require an environment that provides rather than extracts entropy, which contradicts the second law of thermodynamics [18, 29]. Here we have shown that an inverse-only cascade is possible in a strong turbulence. As far as an anomalous scaling is concerned, we relate it to the change of the interaction time along the cascade. All the inverse cascades known before run from fast to slow modes and have a normal scaling. In our case, as in all shell models, cascades always proceed from slow to fast modes. Apparently, this is the reason that non-Gaussianity increases along all our cascades, and an anomalous scaling takes place in both single inverse and single direct cascades. Indeed, proceeding from fast to slow modes (in inverse cascades known before) involves an effective averaging over fast degrees of freedom, which diminishes intermittency. On the contrary, our cascades build up intermittency as they proceed.

Another unexpected conclusion follows from the entropy production balance in a steady turbulent state: even though the marginal statistics of the pumping-connected mode (averaged over all other modes) can be close to Gaussian, the correlations of that mode with other modes cannot be weak.

Most of the present work was devoted to disentangling of the information encoded in strong turbulence. It was predicted that in weak turbulence most of the information is encoded in the three-mode statistics [18], and Figure 7 confirms this prediction. Yet in strong turbulence, we find that as much information is encoded in one-mode as in two-mode statistics, while three-mode statistics does not add much. This could be of practical importance for turbulence studies since it is much more difficult to collect, store and analyze statistics for three-mode and multi-mode distributions. Another important lesson is that measuring or computing mode amplitudes (or velocity structure functions) brings diminishing returns, that is less and less information, as one goes deep into the cascade. The maximal information is encoded in the statistics of the Kolmogorov multipliers. Most of that information is encoded in the statistics of a single multiplier; less than 10% is encoded in the correlation of neighbors. How to decode it is the task for the future.

We wish to thank Yotam Shapira for helpful discussions. The work was supported by the Scientific Excellence Center and Ariane de Rothschild Women Doctoral Program at WIS, grant 662962 of the Simons foundation, grant 075-15-2019-1893 by the Russian Ministry of Science, grant 873028 of the EU Horizon 2020 programme, and grants of ISF, BSF and Minerva. NV was in part supported by NSF grant number DMS-1814619. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF grant number ACI-1548562, allocation DMS-140028.

[1] A. M. Obukhov, *Integral invariants in hydrodynamic systems*, Dokl. Akad. Nauk SSSR, 184:2, (1969).

[2] R. Peierls, *On the kinetic theory of thermal conduction in crystals*, Ann. Phys. (N.Y.) **3**, 1055 (1929).

[3] V. Zakharov, V. Lvov and G. Falkovich, *Kolmogorov Spectra of Turbulence* (Springer 1991).

[4] A. Newell and B. Rumpf, *Wave turbulence*, Annu. Rev. Fluid Mech. **43**, 59 (2011).

[5] S. Nazarenko, *Wave Turbulence* (Springer 2011).

[6] E. Kartashova, L. Piterbarg and G. Reznik, *Weak Nonlinear-interactions of rossby waves on sphere*, Oceanology **29**, 405 (1990).

[7] L. Biferale,v, *Shell models of energy cascade in turbulence*, Ann. Rev. Fluid Mech **35** :441-468 (2003).

[8] V. L'vov E. Podivilov and I. Procaccia, *Hamiltonian structure of the Sabra shell model of turbulence: exact calculation of an anomalous scaling exponent*, EPL **46** 609 (1999).

[9] V. L'vov, E. Podivilov, A. Pomyalov, I. Procaccia and D. Vandembroucq, *Improved shell model of turbulence-text*,Phys. Rev. E **58** 1811,22 (1998).

[10] J.D. Fournier and U. Frisch, *d-Dimensional turbulence*, Phys Rev A **17**, 747 (1978).

[11] T. Gilbert, V. L'vov, A. Pomyalov and I. Procaccia,*Inverse cascade regime in shell models of two-dimensional turbulence*, Phys. Rev. Lett. **89**, 074501 (2002).

[12] A.C. Hindmarsh, ACM Signum Newsletter, **15**(4), 10 (1980), K. Radhakrishnan and A. C. Hindmarsh, the Livermore solver for ordinary differential equations (1993).

[13] E. Balkovsky, G. Falkovich, V. Lebedev and I.Y. Shapiro, *Large-scale properties of wave turbulence*, Phys. Rev. E **52** 4537 (1995).

[14] R. Tom and S. Ray, *Revisiting the SABRA model: Statics and dynamics*, Eur.Phys.Lett, **120** (2017)

[15] G. Falkovich and N. Vladimirova, *Cascades in nonlocal turbulence*, Phys. Rev. E **91**, 041201(R) (2015).

[16] E. Aurell, G. Boffetta, A. Crisanti, P. Frick, G. Paladin, and A. Vulpiani, *Statistical mechanics of shell models for two-dimensional turbulence*, Phys. Rev. E **50**, 4705 (1994).

[17] P. D. Ditlevsen and I. A. Mogensen, *Cascades and statistical equilibrium in shell models of turbulence*, Phys. Rev. E **53**, 4785 (1996).

[18] M. Shavit and G. Falkovich, *Singular measures and information capacity of turbulent cascades*, Phys. Rev. Lett. **125**, 104501 (2020).

[19] G. Falkovich, *Symmetries of the turbulent state*, J. Phys. A: Math. Theor. **42** 123001 (2009).

[20] G. Falkovich, *Cascade and scaling*, Scholarpedia, 3(8) 6088 (2008).

[21] G.Falkovich and V. Lebedev, *Nonlocal vorticity cascade in two dimensions*, Phys Rev E **49**, R1800 (1994); Phys Rev E **50** 3883 (1994).

[22] C. Pasquero and G. Falkovich, *Stationary spectrum of vorticity cascade in two-dimensional turbulence*, Phys Rev E **65**, 056305 (2002).

[23] G. Falkovich, K. Gawedzki and M. Vergassola, *Particles*

and fields in fluid turbulence, Rev. Mod. Phys. (2001).

[24] G. Falkovich and A. Frishman, *Single flow snapshot reveals the future and the past of pairs of particles in turbulence*, Phys. Rev. Lett. **110**, 214502 (2013).

[25] W.J McGill, *Multivariate mutual information*, Psychometrika. **19**, 97-116 (1954).

[26] A.J. Bell, in International workshop on independent component analysis and blind signal separation (2003) p. 921.

[27] Cluster decomposition of the information encoded in different subsets of interacting particles was found in unpublished notes of Onsager (Gregory Eyink, private communication).

[28] A. Kitaev and J. Preskill, *Topological entanglement entropy*, Phys. Rev. Lett. **96**, 110404 (2006).

[29] G. Falkovich and A. Fouxon, *Entropy production and extraction in dynamical systems and turbulence*, New. J. Phys. **6**, 50 (2004).

[30] A. Kolmogorov, *A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number*, J. Fluid Mech. **13** 1 (1962).

[31] R. Benzi, L. Biferale and G. Parisi, *A random process for the construction of multiaffine fields*, Physica D **65** (1993).

[32] G. Eyink, S. Chen, and Q. Chen, *The joint cascade of energy and helicity in three-dimensional turbulence*, J. Stat. Phys. **113**, 719 (2003).

[33] L. Biferale, A. A. Mailybaev, and G. Parisi, *Optimal subgrid scheme for shell models of turbulence*, Phys. Rev. E **95** 043108 (2017).

[34] A.A. Mailybaev, *Hidden scale invariance of intermittent turbulence in a shell model*, Phys. Rev. Fluids **6.1** (2021).

[35] Q. Chen, S. Chen, G. L. Eyink, and K. R. Sreenivasan, *Kolmogorov's third hypothesis and turbulent sign statistics*, Phys. Rev. lett. **90** 254501 (2003).

[36] G. Eyink, S. Chen, Q. Chen, *Gibbsian Hypothesis in Turbulence*, J. Stat. Phys. **113** (2003).

Floquet boundary states in AB-stacked graphite

Hernán L. Calvo,[1,2] Jose E. Barrios Vargas,[3] and Luis E. F. Foa Torres[4]

[1]*Instituto de Física Enrique Gaviola (CONICET) and FaMAF, Universidad Nacional de Córdoba, Argentina*
[2]*Departamento de Física, Universidad Nacional de Río Cuarto, Ruta 36, Km 601, 5800 Río Cuarto, Argentina*
[3]*Departamento de Física y Química Teórica, Facultad de Química, UNAM, 04510 México City, México*
[4]*Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile*

We report on the effect of laser illumination with circularly polarized light on the electronic structure of AB-stacked graphite samples. By using Floquet theory in combination with Green's function techniques, we find that the polarized light induces bandgap openings at the Floquet zone edge $\hbar\Omega/2$, bridged by chiral boundary states. These states propagate mainly along the borders of the constituting layers as evidenced by the time-averaged local density of states and the probability current density in several geometries. Semianalytic calculations of the Chern number suggest that these states are of topological nature, similar to those found in illuminated 2D samples like monolayer and bilayer graphene. These states are promising candidates for the realization of a three-dimensional version of the quantum Hall effect for Floquet systems.

I. INTRODUCTION

Condensed matter physics has provided one of the most fertile and captivating grounds for discoveries over the last few decades:[1] from two-dimensional materials,[2] which were thought to be impossible to exist in nature, to new topological phases of matter[3–7] which have completely reshaped our understanding of old concepts. The use of light has been an instrumental cornerstone in this adventure, being one of the prime tools for unveiling a material or device properties.[8,9] However, beyond this already important role, a new research front aims at using light in an active fashion to actually change the response of a material,[10,11] by opening a gap[10,12–15] or even endowing a material with topological states[10,11,16–19] or a spontaneous orbital magnetization.[20]

Experiments have successfully confirmed the possibility of creating and tuning hybrid electron-photon states,[15,21] also called Floquet-Bloch states, and also the generation of a laser-induced Hall effect in graphene.[22] The name Floquet here is used because the prevalent theory for this type of driven systems: the Floquet theory,[23–26] from which the spectrum, effective Hamiltonians,[27,28] a map of the topological invariants[29,30] and transport properties[31–33] can be computed. It is also worthy to mention that this Floquet picture can be ported, with small changes, to phonon-induced states as in Refs. [34] and [35]. Most of the attention has been devoted to illuminating two-dimensional materials, including graphene,[10,36–38] germanene,[39] silicene,[40] MoS$_2$,[41] and manufactured systems like periodic arrangements of quantum rings.[42] More recently, the interest in Floquet engineering three-dimensional materials such as three-dimensional topological insulators,[43,44] Weyl semimetals[45,46] or Dirac materials[47] has been growing.

Here we focus on laser-illuminated graphite. In contrast with most three-dimensional crystals, graphite has a hierarchical structure of weakly coupled layers making it an archetypal system for learning on the way from two to three dimensions. In two-dimensions, circularly polarized radiation leads to bandgap openings and Floquet edge states that bridge the gap.[10,18,36–38] These topological Floquet edge states are akin to those found in Chern insulators or in the integer quantum Hall effect, they are robust and chiral.[18] By analogy with the physics of the quantum Hall effect which was discovered in two dimensions[48,49] and which has been predicted to be possible in three dimensions,[50] a prediction which has not been verified until very recently,[51] one might then wonder what happens in three-dimensions with the laser-induced states. In this paper, we show that for graphite there are also laser-induced bandgaps at $\pm\hbar\Omega/2$ which turn out to be bridged by surface states. Our calculations, which are based on Green's functions techniques combined with Floquet theory, show that these surface states are chiral, have a topological nature and can form a band of chiral states bridging the bulk gap. In the following we introduce our model, followed by an analysis of bulk graphite, and finite samples with emphasis on the surface states and the associated currents.

II. HAMILTONIAN MODEL AND FLOQUET SPACE SOLUTION SCHEME

Let us introduce our model for graphite under circularly polarized laser illumination. We consider graphene layers in graphite with AB stacking, and we follow Ref. [52] for the tight-binding parameters obtained in the static case (see below). We consider a tight-binding description for graphite given by the generic Hamiltonian

$$\hat{\mathcal{H}} = \sum_{\boldsymbol{r},\boldsymbol{r}'} \gamma_{\boldsymbol{r},\boldsymbol{r}'} \left|\boldsymbol{r}\right\rangle\left\langle\boldsymbol{r}'\right| , \qquad (1)$$

where $\boldsymbol{r}$ and $\boldsymbol{r}'$ denote the positions of the carbon atoms in the lattice, such that the states $|\boldsymbol{r}\rangle$ form a real-space basis. Under this notation, the sum runs over sites connected by the hopping amplitudes $\gamma_{\boldsymbol{r},\boldsymbol{r}'}$, and it also includes the on-site energies through $\gamma_{\boldsymbol{r},\boldsymbol{r}} = \epsilon_{\boldsymbol{r}}$.

The laser field is included within a semiclassical approximation as a time-dependent term in the Hamiltonian. The time-periodic electric field $\boldsymbol{E}(\boldsymbol{r},t)$, with period $T = 2\pi/\Omega$, is included through the gauge $\boldsymbol{E} = -\partial_t \boldsymbol{A}$, where the vector potential takes the form $\boldsymbol{A}(\boldsymbol{r},t) = \mathrm{Re}[\boldsymbol{A}_0 e^{i\Omega(z/c-t)}]$, such that its direction of propagation points perpendicular to the graphene layers, defined in the xy planes. As a consequence, in three dimensional samples there is a variation of the wave along the z direction due to the phase factor $\Omega z/c$ in $\boldsymbol{A}(\boldsymbol{r},t)$. This would become appreciable in samples with transversal lengths higher than $L_z \sim 0.1\lambda$, with $\lambda = 2\pi c/\Omega$ the laser's wavelength. For laser frequencies near the infrared region ($\hbar\Omega \sim 2$ eV) this implies $L_z \sim 620$ Å, which in graphite means a number of ~ 185 transversal layers. As we will assume smaller values for L_z, the z dependence in the vector potential can be neglected in a first approximation. We will work with circularly polarized light, by taking $\boldsymbol{A}_0 = A_0(1, i\tau, 0)$, with $\tau = \pm 1$ the handedness of the polarized light.[53] By means of Peierls' substitution, the vector potential enters in Eq. (1) by adding a time-dependent phase in the hopping amplitudes, namely,

$$\gamma_{\boldsymbol{r},\boldsymbol{r}'} \xrightarrow{\text{laser}} g_{\boldsymbol{r},\boldsymbol{r}'}(t) = \gamma_{\boldsymbol{r},\boldsymbol{r}'} \exp\left[i\frac{2\pi}{\Phi_0}\int_{\boldsymbol{r}'}^{\boldsymbol{r}} \mathrm{d}\boldsymbol{\ell}\cdot\boldsymbol{A}(t)\right], \quad (2)$$

with Φ_0 the magnetic flux quantum and the line integral taken over the straight path connecting sites $\boldsymbol{r}'$ and $\boldsymbol{r}$. Given the specific form of the vector potential, the time-dependent hopping terms entering in the Hamiltonian are given by:

$$g_{\boldsymbol{r},\boldsymbol{r}'}(t) = \gamma_{\boldsymbol{r},\boldsymbol{r}'} \sum_{n=-\infty}^{\infty} i^n \mathcal{J}_n(\zeta_{\boldsymbol{r},\boldsymbol{r}'}) e^{in(\Omega t - \phi_{\boldsymbol{r},\boldsymbol{r}'})}, \quad (3)$$

where we used the Jacobi-Anger expansion for future convenience. In this expression, $\mathcal{J}_n(\zeta_{\boldsymbol{r},\boldsymbol{r}'})$ represents the n-th Bessel function of the first kind, and the adimensional variable $\zeta_{\boldsymbol{r},\boldsymbol{r}'} = 2\pi A_0|\boldsymbol{r}-\boldsymbol{r}'|\sin\theta_{\boldsymbol{r},\boldsymbol{r}'}/\Phi_0$ quantifies the strength of the laser along the carbon bond, characterized by $\boldsymbol{r}-\boldsymbol{r}' = |\boldsymbol{r}-\boldsymbol{r}'|(\sin\theta_{\boldsymbol{r},\boldsymbol{r}'}\cos\phi_{\boldsymbol{r},\boldsymbol{r}'},\, \sin\theta_{\boldsymbol{r},\boldsymbol{r}'}\sin\phi_{\boldsymbol{r},\boldsymbol{r}'},\, \cos\theta_{\boldsymbol{r},\boldsymbol{r}'})$.

A. Floquet theory

In this section we introduce the basics of Floquet theory as used later in this paper. The readers already acquainted with the technical details or focused on the results rather than the techniques may skip this in a first reading.

According to Floquet theory,[23–26] there is a full set of solutions to the time-dependent Schrödinger equation (TDSE) of the form $|\psi(t)\rangle = e^{-i\epsilon t/\hbar}|\phi(t)\rangle$, where the Floquet state $|\phi(t)\rangle$ presents the same periodicity of the Hamiltonian, i.e. $|\phi(t+T)\rangle = |\phi(t)\rangle$. By replacing this ansatz in the TDSE one obtains

$$\hat{\mathcal{H}}_{\mathrm{F}}|\phi(t)\rangle = \epsilon|\phi(t)\rangle, \quad (4)$$

where $\hat{\mathcal{H}}_{\mathrm{F}} = \hat{\mathcal{H}}(t) - i\hbar\partial_t$ is the so-called Floquet Hamiltonian and ϵ its associated quasienergy. The great advantage of Floquet theory is that $\hat{\mathcal{H}}_{\mathrm{F}}$ can be reduced to a time-independent matrix when described in the product space (also called Floquet space) $\mathcal{F} = \mathcal{R}\otimes\mathcal{T}$, with $\mathcal{R}$ the usual Hilbert space and $\mathcal{T}$ the space of time-periodic functions, spanned by the set of orthonormal vectors $\langle t|n\rangle = e^{in\Omega t}$, with n an integer number. Working within the local space representation, a suitable basis for $\mathcal{F}$ is given by the product states $|\boldsymbol{r},n\rangle = |\boldsymbol{r}\rangle\otimes|n\rangle$, representing the lattice site $\boldsymbol{r}$ and the Fourier replica n, together with the inner product rule

$$\langle\boldsymbol{r},n|\boldsymbol{r}',m\rangle = \int_0^T \frac{\mathrm{d}t}{T} e^{i(m-n)\Omega t}\langle\boldsymbol{r}|\boldsymbol{r}'\rangle = \delta_{\boldsymbol{r},\boldsymbol{r}'}\delta_{n,m}. \quad (5)$$

On this basis, the Floquet states in Eq. (4) can be computed as

$$|\phi(t)\rangle = \sum_n e^{in\Omega t}|\phi_n\rangle \xrightarrow{\mathcal{F}} |\phi\rangle = \sum_{\boldsymbol{r},n}\phi_n(\boldsymbol{r})|\boldsymbol{r},n\rangle, \quad (6)$$

with $\phi_n(\boldsymbol{r}) = \langle\boldsymbol{r},n|\phi\rangle$ the amplitude of the Floquet state at site $\boldsymbol{r}$ and replica n. Importantly, the matrix elements of the Floquet Hamiltonian $[H_{\mathrm{F}}]^{n,m}_{\boldsymbol{r},\boldsymbol{r}'} = \langle\boldsymbol{r},n|\hat{\mathcal{H}}_{\mathrm{F}}|\boldsymbol{r}',m\rangle$ are in this basis

$$[H_{\mathrm{F}}]^{n,m}_{\boldsymbol{r},\boldsymbol{r}'} = \int_0^T \frac{\mathrm{d}t}{T} e^{i(m-n)\Omega t} H_{\boldsymbol{r},\boldsymbol{r}'}(t), + n\hbar\Omega\delta_{\boldsymbol{r},\boldsymbol{r}'}\delta_{nm}, \quad (7)$$

where the inner product includes the average over one driving cycle, thus Eq. (4) written in this composite space becomes a time-independent eigenvalue problem. Once the Floquet eigenstates $|\phi\rangle$ are obtained in $\mathcal{F}$, it is possible to return to the usual Hilbert space $\mathcal{R}$ and calculate the expectation value of any observable from the general solution $|\psi(t)\rangle$ of the TDSE. In particular, we are interested in the probability density $\hat{\rho}(\boldsymbol{r}) = |\boldsymbol{r}\rangle\langle\boldsymbol{r}|$, whose time-averaged expectation value with respect to some eigenstate of the TDSE writes

$$\rho(\boldsymbol{r}) = \int_0^T \frac{\mathrm{d}t}{T}\langle\hat{\rho}(\boldsymbol{r})\rangle = \sum_n |\phi_n(\boldsymbol{r})|^2. \quad (8)$$

We are also interested in the probability current density, defined as $\hat{J}(\boldsymbol{r},t) = -i[\hat{\mathcal{H}}(t),\hat{\rho}(\boldsymbol{r})]/\hbar$. Its time-averaged expectation value can be written in terms of the Floquet Hamiltonian

$$J(\boldsymbol{r}) = \frac{2}{\hbar}\sum_{\boldsymbol{r}'}\sum_{n,m}\mathrm{Im}\{\phi_n^*(\boldsymbol{r})[H_{\mathrm{F}}]^{n,m}_{\boldsymbol{r},\boldsymbol{r}'}\phi_m(\boldsymbol{r}')\}. \quad (9)$$

Since the averaged probability current at site $\boldsymbol{r}$ needs to be zero due to probability conservation, we will use this quantity to check that there is no charge accumulation/loss after completing one period of the driving field. More interestingly, from this expression we can extract the bond current as[54]

$$J(\boldsymbol{r},\boldsymbol{r}') = \frac{2}{\hbar}\sum_{n,m}\mathrm{Im}\{\phi_n^*(\boldsymbol{r})[H_{\mathrm{F}}]^{n,m}_{\boldsymbol{r},\boldsymbol{r}'}\phi_m(\boldsymbol{r}')\}, \quad (10)$$

which, as we will show later on, gives a clear picture on the chiral nature of the resulting eigenstates of the illuminated system.

As we already mentioned, the periodic time-dependence enters in Eq. (2) as an additional phase that the electron picks up when it "hops" from site $\boldsymbol{r}'$ to site $\boldsymbol{r}$. The Floquet Hamiltonian can then be calculated from Eq. (7) as

$$[H_{\mathrm{F}}]^{n,m}_{\boldsymbol{r},\boldsymbol{r}'} = \gamma^{(m-n)}_{\boldsymbol{r},\boldsymbol{r}'} + n\hbar\Omega\delta_{\boldsymbol{r},\boldsymbol{r}'}\delta_{nm}, \quad (11)$$

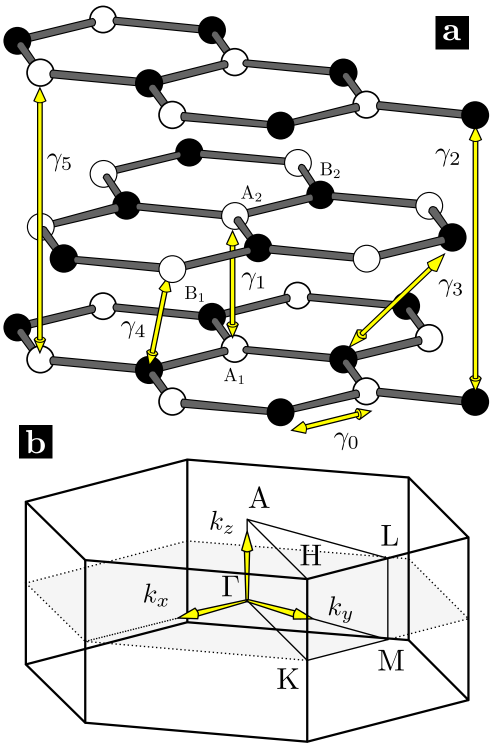


Figure 1. Graphite with AB stacking. (a) Schematic view of three adjacent graphene layers. The hopping amplitudes are marked by yellow arrows (see text). (b) Graphite's first Brillouin zone with the high-symmetry points.

where the hopping amplitudes are defined as Fourier components of the time-dependent ones appearing in Eq. (3), i.e.,

$$\gamma^{(n)}_{\boldsymbol{r},\boldsymbol{r}'} = \int_0^T \frac{\mathrm{d}t}{T} g_{\boldsymbol{r},\boldsymbol{r}'}(t) e^{in\Omega t} = \gamma_{\boldsymbol{r},\boldsymbol{r}'} i^n \mathcal{J}_n(\zeta_{\boldsymbol{r},\boldsymbol{r}'}) e^{in\phi_{\boldsymbol{r},\boldsymbol{r}'}}, \tag{12}$$

and this can be *interpreted* as the probability amplitude for the electron to hop from site $\boldsymbol{r}'$ to site $\boldsymbol{r}$, together with the absorption ($n > 0$) or emission ($n < 0$) of $|n|$ photons.

So far we have not specified the tight-binding Hamiltonian of Eq. (1), so the above discussion is somewhat general as far as the light propagates along the z direction. In graphite with AB stacking, the unit cell contains four basis sites: A_1 and B_1 in the lower graphene layer (LL), A_2 and B_2 in the upper graphene layer (UL), see Fig. 1. The sites in the UL are displayed in such a way that the A_2 site is aligned with the A_1 site of the LL. This implies the following choice for the basis vectors in the unit cell:

$$\mathrm{LL} \rightarrow \quad \boldsymbol{\delta}_{A_1} = (0,0,0)\,, \qquad \boldsymbol{\delta}_{B_1} = (0,a_0,0)\,, \tag{13}$$

$$\mathrm{UL} \rightarrow \quad \boldsymbol{\delta}_{A_2} = (0,0,c_0)\,, \qquad \boldsymbol{\delta}_{B_2} = \left(\frac{\sqrt{3}a_0}{2}, \frac{a_0}{2}, c_0\right), \tag{14}$$

where $a_0 = 1.42$ Å is the carbon-carbon distance in graphene and $c_0 = 3.35$ Å is the separation between two adjacent layers. The graphite's Bravais lattice can then be described by the primitive vectors

$$\boldsymbol{a}_1 = \left(\frac{\sqrt{3}a_0}{2}, \frac{3a_0}{2}, 0\right), \; \boldsymbol{a}_2 = \left(-\frac{\sqrt{3}a_0}{2}, \frac{3a_0}{2}, 0\right)$$
$$\boldsymbol{a}_3 = (0,0,2c_0)\,. \tag{15}$$

The translational invariance along the three directions given by the primitive vectors allows us to decompose the static Hamiltonian in Eq. (1) as the following operator representing the Bloch Hamiltonian:

$$\hat{\mathcal{H}}_{\boldsymbol{k}} = \sum_{\boldsymbol{R}} \hat{\mathcal{V}}_{\boldsymbol{R}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}}, \tag{16}$$

where $\boldsymbol{R}$ denotes the position of the nearest-neighbor unit cells to the one placed at the origin. In the general description of the unit cell position through $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$, with n_i integer numbers, the lattice connectivity given by the hopping parameters determined in Ref. [52] implies that the possible values for n_i in $\boldsymbol{R}$ are $n_i = \{-1,0,1\}$. To reconcile the notation, we notice that the hopping operator $\hat{\mathcal{V}}_{\boldsymbol{R}}$ represents the bonds going from site $\boldsymbol{r}' = \boldsymbol{\delta}(\boldsymbol{r}')$ to site $\boldsymbol{r} = \boldsymbol{R} + \boldsymbol{\delta}(\boldsymbol{r})$, where $\boldsymbol{\delta}(\boldsymbol{r})$ indicates the basis vector associated with $\boldsymbol{r}$.

In the case of graphite with AB stacking, we consider for the static case the following parameters:[52] $\gamma_0 = 3.16$ eV connecting nearest-neighbor in-plane sites (A_1B_1 and A_2B_2), $\gamma_1 = 0.39$ eV for A_1A_2, $\gamma_2 = -0.02$ eV connecting B-sites (B_1B_1 and B_2B_2) between consecutive cells along $\boldsymbol{a}_3$, $\gamma_3 = 0.315$ eV for B_1B_2, $\gamma_4 = 0.044$ eV for A_1B_2 and B_1A_2, and $\gamma_5 = 0.038$ eV connecting A-sites (A_1A_1 and A_2A_2) between consecutive cells along $\boldsymbol{a}_3$. This can be easily understood, for example, by inspecting the matrix elements of the Bloch Hamiltonian with respect to the site basis $\{|i\rangle\}$, with $i = 1, \ldots, 4$ for (A_1, B_1, A_2, B_2):

$$\boldsymbol{H}_{\boldsymbol{k}} = \begin{pmatrix} \epsilon_0 + \Delta + \gamma_5 f_5 & \gamma_0 f_1 & \gamma_1 f_4 & \gamma_4 f_2 f_4 \\ \gamma_0 f_1^* & \epsilon_0 + \gamma_2 f_5 & \gamma_4 f_1^* f_4 & \gamma_3 f_3 f_4 \\ \gamma_1 f_4^* & \gamma_4 f_1 f_4^* & \epsilon_0 + \Delta + \gamma_5 f_5 & \gamma_0 f_2 \\ \gamma_4 f_2^* f_4^* & \gamma_3 f_3^* f_4^* & \gamma_0 f_2^* & \epsilon_0 + \gamma_2 f_5 \end{pmatrix}, \tag{17}$$

where $\epsilon_0 = -0.024$ eV is the Fermi energy and $\Delta = -0.008$ eV is the energy shift between inequivalent carbon atoms. The functions $f_i = f_i(\boldsymbol{k})$ carry information about the directions in which the basis sites in the unit cell are connected with its neighbors, and are defined as:

$$\begin{aligned}
f_1 &= 1 + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_2},\\
f_2 &= 1 + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\boldsymbol{k}\cdot(\boldsymbol{a}_1-\boldsymbol{a}_2)} = e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} f_1^*,\\
f_3 &= 1 + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2} + e^{i\boldsymbol{k}\cdot(\boldsymbol{a}_1-\boldsymbol{a}_2)} = e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2} f_1,\\
f_4 &= 1 + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_3},\\
f_5 &= 2\cos(\boldsymbol{k}\cdot\boldsymbol{a}_3).
\end{aligned}$$

So, for example, in the matrix element $[\boldsymbol{H}_{\boldsymbol{k}}]_{12}$, the function f_1 accounts for the intracell connection $\boldsymbol{\delta}_{\mathrm{B}_1} \to \boldsymbol{\delta}_{\mathrm{A}_1}$ and the intercell connections $\boldsymbol{\delta}_{\mathrm{B}_1} \to \boldsymbol{a}_1 + \boldsymbol{\delta}_{\mathrm{A}_1}$ and $\boldsymbol{\delta}_{\mathrm{B}_1} \to \boldsymbol{a}_2 + \boldsymbol{\delta}_{\mathrm{A}_1}$. Similarly, f_2 in $[\boldsymbol{H}_{\boldsymbol{k}}]_{34}$ takes into account those bonds connecting $\boldsymbol{\delta}_{\mathrm{B}_2}$ with $\boldsymbol{\delta}_{\mathrm{A}_2}$. With this notation, it is clear that the coupling between different graphene layers enters in the 2×2 off-diagonal blocks, which are modulated by either f_4 or f_4^*. In the diagonal blocks, on the other hand, there are in-plane connections given by γ_0 and on-site energy corrections due to the coupling to neighbor cells along $\boldsymbol{a}_3$. Additionally, the difference in the involved directions given by f_1 in the LL and f_2 in the UL, respectively, comes from the choice of the unit cell basis sites. Notice, in particular, that $|f_1| = |f_2| = |f_3|$.

If we now turn on the laser, one should notice that the vector potential does not break translational invariance, so it is possible to combine Eq. (16) with Eq. (7) by introducing a superindex $(m-n)$ in the Bloch Hamiltonian which accounts for the replicas $m \to n$ it connects. This implies that all hoppings belonging to $\hat{\mathcal{H}}_{\boldsymbol{k}}^{(m-n)}$ need to be transformed as $\gamma_{\boldsymbol{r},\boldsymbol{r}'} \to \gamma_{\boldsymbol{r},\boldsymbol{r}'}^{(m-n)}$, and we obtain the following structure

$$\hat{\mathcal{H}}_{\mathrm{F},\boldsymbol{k}} = \sum_{n,m}\left[\hat{\mathcal{H}}_{\boldsymbol{k}}^{(m-n)} + n\hbar\Omega\hat{\mathcal{I}}\delta_{n,m}\right]\otimes|n\rangle\langle m|\,, \qquad (18)$$

for the Floquet-Bloch Hamiltonian, defined in the $\mathcal{F}$-space. Here $\hat{\mathcal{I}}$ represents the identity operator in the reduced space of the unit cell and $|n\rangle$ corresponds to the Fourier replica n.

Summarizing, the construction of $\hat{\mathcal{H}}_{\mathrm{F},\boldsymbol{k}}$ follows two simple steps: (1) the identification of the static Bloch Hamiltonian of Eq. (16), and (2) the Fourier decomposition of all their matrix elements once the laser has been incorporated. Notice that there is, however, a subtlety in going from step 1 to step 2: as the time-dependent hopping phases [cf. Eq. (3)] depend on both the magnitude and direction of the bond connecting sites $\boldsymbol{r}'$ and $\boldsymbol{r}$, this information needs to be given in step 1 even if in the static case such a dependence is not present.

Following the above steps, the matrix elements of the Floquet-Bloch Hamiltonian for bulk graphite can be compactly written in terms of the hopping amplitudes between the different basis sites $i,j = \{\mathrm{A}_1, \mathrm{B}_1, \mathrm{A}_2, \mathrm{B}_2\}$ as

$$[H_{\mathrm{F},\boldsymbol{k}}]_{i,j}^{n,m} = \sum_{\boldsymbol{R}} \gamma_{\boldsymbol{R}+\boldsymbol{\delta}_i,\boldsymbol{\delta}_j}^{(n-m)} e^{i\boldsymbol{k}\cdot\boldsymbol{R}} + n\hbar\Omega\delta_{i,j}\delta_{n,m}. \qquad (19)$$

Notice that not all lattice vectors $\boldsymbol{R}$ contribute to the sum on the right hand side, as we assume some finite range for the allowed hopping parameters.

III. ILLUMINATED BULK GRAPHITE

The purpose of this section is to give an explicit calculation of the Floquet Hamiltonian in illuminated graphite, such that the role of the laser field is evidenced as modifications in the band structure of the static material. This will allow us to identify, in turn, the band crossing regions where boundary states induced by the laser may appear.

As starting point, in Fig. 2(a), we show the dispersion relation for bulk graphite in the absence of laser illumination. We can see how the highest valence (1v) and the lowest conduction (1c) bands cross at the K symmetry point. These bands are quadratic in shape (a reminiscence of bilayer graphene's band structure), and cross each other at two different points: one of them along the Γ-K path while the other exactly at the K-point (see inset).[55] The breaking of the electron-hole (e-h) symmetry is clearly visible along the whole spectrum and is produced by the hoppings γ_2, γ_4, and γ_5. Along the A-H-L path [top face of the Brillouin zone in Fig. 1(b)] the energy bands become doubly degenerate. Inspecting the Bloch Hamiltonian in Eq. (17), this band degeneracy can be easily understood since f_4 becomes exactly zero, meaning that the layers are completely decoupled along this path and, in addition, $|f_1| = |f_2|$. Exactly at the H point, there is a gap $\Delta\epsilon \simeq 124$ meV due to γ_2, γ_5 and Δ.

For illuminated graphite, one should notice that an infinite number of replicas develop in the quasienergy spectrum associated with the Floquet-Bloch Hamiltonian. We are, however, interested in the changes that the laser field produces on the static spectrum shown in Fig. 2(a). A convenient way to visualize this is to use a colorscale that represents the weight of the $\boldsymbol{k}$-eigenstates on the $n=0$ Fourier replica, i.e.

$$\bar{w}_{\boldsymbol{k}} = \sum_{\boldsymbol{r}} |\phi_{\boldsymbol{k},0}(\boldsymbol{r})|^2, \qquad (20)$$

where the sum runs over the basis sites composing the unit cell, i.e. $\boldsymbol{r} = \{\boldsymbol{\delta}_{\mathrm{A}_1}, \boldsymbol{\delta}_{\mathrm{B}_1}, \boldsymbol{\delta}_{\mathrm{A}_2}, \boldsymbol{\delta}_{\mathrm{B}_2}\}$. Comparing the above expression with Eq. (8), $\bar{w}_{\boldsymbol{k}}$ represents the fraction of the (time-averaged) probability density which is distributed along the $n=0$ replica. Notice that in the static case $\bar{w}_{\boldsymbol{k}} = \sum_{\boldsymbol{r}} \rho_{\boldsymbol{k}}(\boldsymbol{r}) = 1$ since no other replicas are involved. We set the strenght of the laser through $\zeta_0 = 2\pi A_0 a_0/\Phi_0 = 0.0568$, such that $\zeta_{\boldsymbol{r},\boldsymbol{r}'} = \zeta_0|\boldsymbol{r}-\boldsymbol{r}'|\sin\theta_{\boldsymbol{r},\boldsymbol{r}'}/a_0$, and the frequency as $\hbar\Omega = 2.2$ eV. In this regime, no strong modifications of the entire band structure are expected and one can, in turn, safely truncate the full Floquet space by taking an adequate number of replicas such that the observed spectrum converges. For the chosen parameters, appreciable changes induced by the laser only appear around the K and H symmetry points where the bands come close to each other, so we can focus in the gray shaded rectangles of Fig. 2(a). This is plotted in Fig. 2(b) in the vicinity of the K point along the path Γ-K-M and in Fig. 2(c) for the vicinity of the H point along the path A-H-L, respectively. The main features in these plots are the bandgap openings that appear around the boundaries of the Floquet zone (FZB), defined at $\epsilon = \pm\hbar\Omega/2$. Although not clearly visible, there is also a bandgap opening around the center of the Floquet zone (FZC) at $\epsilon = 0$. The large difference in the

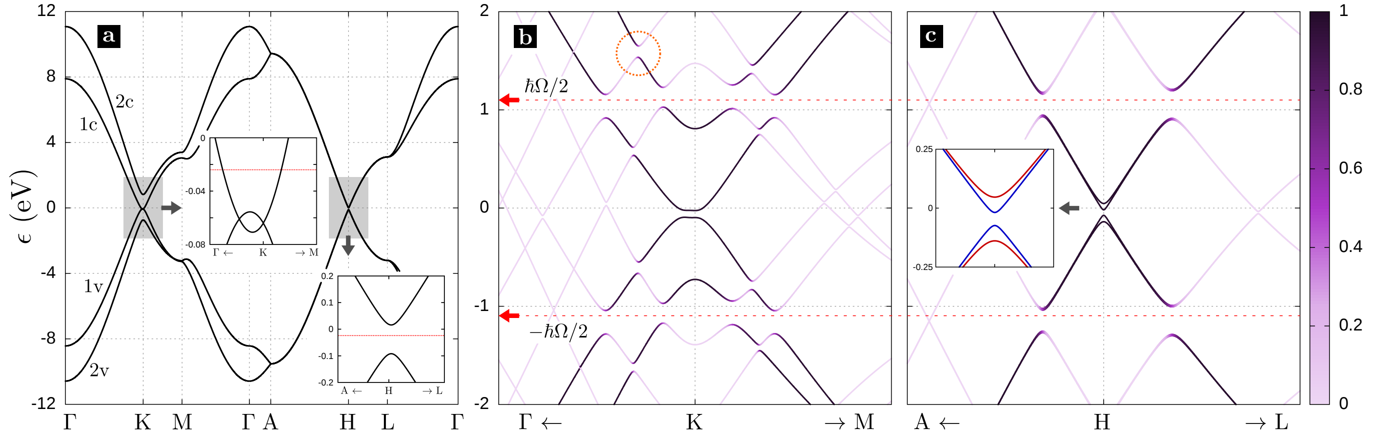


Figure 2. Bulk graphite's dispersion relations. (a) Static case where no laser is applied. The labels in the $\boldsymbol{k}$-axis (horizontal) correspond to the symmetry points depicted in Fig. 1(b). The valence bands are labeled as 1v and 2v, while the conduction bands are labeled as 1c and 2c. The insets are zoom regions around the K and H symmetry points and the Fermi energy is depicted in dotted red. (b) and (c) are zoom regions around the K and H symmetry points [gray shaded rectangles in (a)], respectively, for circularly polarized light with $\zeta_0 = 0.04$ Å$^{-1}\times 1.42$ Å $= 0.0568$ and $\hbar\Omega = 2.2$ eV. The inset in (c) is a zoom around $\epsilon = 0$ and shows the LL-bands in solid red and the UL-bands in solid blue (see text). The colorscale represents the weight of the $\boldsymbol{k}$-states on the zeroth Fourier replica, according to Eq. (20).

magnitude of the two gaps obeys a simple reason: the gap in the FZB region depends linearly on the laser's strength, while for the FZC gap such a dependence is quadratic.[14]

As it happens in two-dimensional samples with circularly polarized light,[10,14] the bandgap openings are a known consequence of the breaking of the time-reversal symmetry, which in this case extends to three-dimensional graphite. In Fig. 2(b), we can also distiguish some avoided crossings above and below the FZB gaps, between different e-h band partners. Take for example the one marked by the dotted circle, which corresponds to the crossing between the 2c–0 and the 1v–1 bands, where "–n" means that it belongs to the nth-Fourier replica in the limit $\zeta_0 \to 0$. We can see, however, that this is not a fully developed gap since for that energy range the 1c–0 band (to the left) is barely affected by the laser.

Interestingly, in Fig. 2(c), the band degeneracy observed along the A-H-L trajectory in Fig. 2(a) for the static case is removed by the laser (see inset). Although the lower and the upper layers are still decoupled, the combination of the broken sublattice symmetry, due to the on-site energies, i.e.,

$$\epsilon_{\mathrm{A}_1} = \epsilon_{\mathrm{A}_2} = \epsilon_0 + \Delta - 2\gamma_5,$$
$$\epsilon_{\mathrm{B}_1} = \epsilon_{\mathrm{B}_2} = \epsilon_0 - 2\gamma_2,$$

together with the handedness of the circularly polarized waves, allows one to distinguish between LLs and ULs, since these are mirror images of each other. This is depicted in the inset of Fig. 2(c), where we use red for the LL-bands and blue for the UL-bands. If we change the handedness of the laser field, then the bands behavior is indeed inverted (i.e., "red becomes blue" and viceversa), as expected from the $z \to -z$ inversion operation.

It is important to notice that for the gap at the FZC to be greater than that at the FZB one needs to be in the strong coupling regime. Therefore, as the laser intensity is smoothly increased from zero, the first visible feature, without the complications of heating and non-equilibrium effects present in the strong coupling regime,[56] should be the gap at the FZB. Because of this, and also to keep within the validity range of our assumptions, we focus from now on in the FZB modes. Later on, when inspecting the bond currents in Sec. V, we will set a stronger laser intensity where the FZC gap becomes clearly visible.

IV. LASER INDUCED BOUNDARY STATES

Having finished our program with illuminated bulk graphite, our next step is to check for laser-induced boundary states. To do it we introduce a boundary and inspect whether midgap states appear or not. We take one of the three directions of the lattice given by the primitive vectors of Eq. (15) as finite, while keeping translational invariance along the other two. For example, we could define a 'slab' geometry along the $\hat{\boldsymbol{a}}_1$ direction by taking a lattice with $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$, such that $1 \leq n_1 \leq N_1$, and $\{n_2, n_3\} \in \mathbb{Z}$. The problem then is that one should take a sufficiently large width as to prevent a considerable overlap between the expected boundary states, if these are present at the borders of the sample. This brings with it an important numerical effort since this geometry increases the dimension of the effective Hamiltonian to be diagonalized.

Perhaps a more convenient strategy to circumvent this issue is to refer to the time-averaged local density of states (LDoS) $\mathcal{N}_{\boldsymbol{r},\boldsymbol{k}}(\epsilon)$, which characterizes the weight of the $\boldsymbol{k}$-state at quasienergy ϵ on the site $\boldsymbol{r}$ along the broken direction. In the context of Floquet theory, this quantity can be written as[57]

$$\mathcal{N}_{\boldsymbol{r},\boldsymbol{k}}(\epsilon) = -\frac{1}{\pi}\lim_{\eta\to 0^+}\mathrm{Im}\Big[\langle \boldsymbol{r}, 0|\, \hat{\mathcal{G}}_{\mathrm{F},\boldsymbol{k}}(\epsilon + i\eta)\, |\boldsymbol{r}, 0\rangle\Big], \quad (21)$$

with $\hat{\mathcal{G}}_{\mathrm{F},\boldsymbol{k}}$ the Floquet-Green operator associated with $\hat{\mathcal{H}}_{\mathrm{F},\boldsymbol{k}}$,

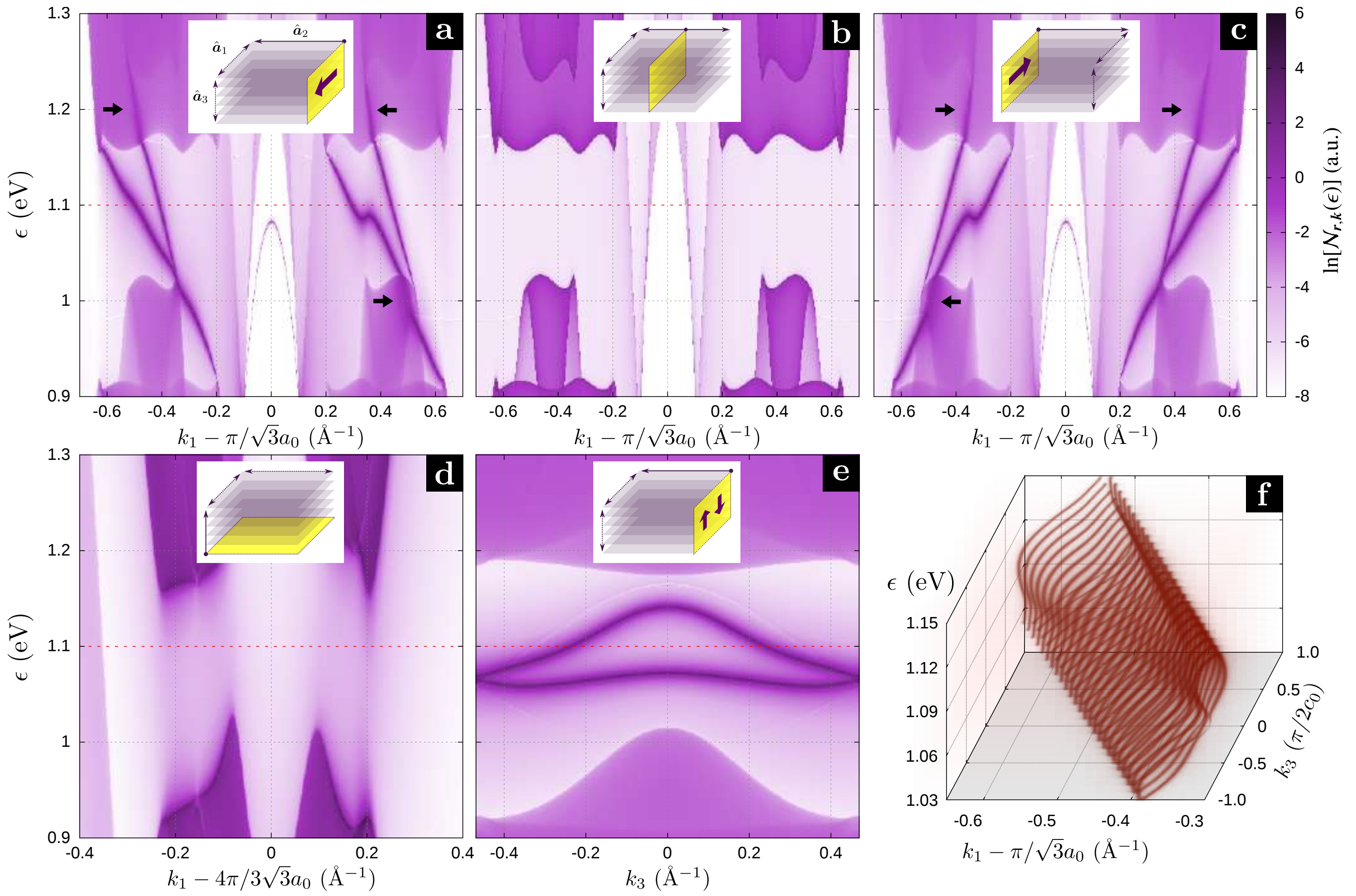


Figure 3. Quasienergy and $\boldsymbol{k}$ resolved local density of states $\mathcal{N}_{\boldsymbol{r},\boldsymbol{k}}(\epsilon)$ in logarithmic scale for illuminated graphite. The LDoS is evaluated in a sample with N unit cells along $\hat{\boldsymbol{a}}_2$ in (a)-(c) and (e) and along $\hat{\boldsymbol{a}}_3$ in (d). The insets schematically illustrate the regions in which the LDoS is being evaluated. Panels (a), (b) and (c) show the LDoS evaluated at $n_2 = 1$, $N/2$, and N, respectively, for $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1$; in (d) the LDoS is evaluated at $n_3 = 1$ for $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1$, while in (e) $n_2 = 1$ and $\boldsymbol{k} = (\pi/\sqrt{3}a_0 + 0.4\ \text{Å}^{-1})\hat{\boldsymbol{a}}_1 + k_3\hat{\boldsymbol{a}}_3$. We use an extremely large value for N, i.e. $N \sim 2^{25}$, such that the sample can be considered as semi-infinite in (a), (c)-(e), while in (b) the sample can be understood as infinite. In (f) we show the LDoS at $n_2 = 1$ as a function of energy and $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1 + k_3\hat{\boldsymbol{a}}_3$. The laser parameters coincide with those in Fig. 2 and the FZB at $\epsilon = 1.1$ eV is denoted by red dashed lines

.

i.e.,

$$\hat{\mathcal{G}}_{\mathrm{F},\boldsymbol{k}}(\epsilon) = \left[\epsilon\hat{\mathcal{I}}_{\mathrm{F}} - \hat{\mathcal{H}}_{\mathrm{F},\boldsymbol{k}}\right]^{-1}. \tag{22}$$

The advantage of this method relies in that one still operates in the original dimension of the *truncated* Floquet space, i.e., $\dim \mathcal{F} = 4 \times (2n_r + 1)$, where $n_r \geq 0$ denotes the highest taken value for the Fourier replica and we consider the replicas going from $-n_r$ to n_r. The recursive Green's function method allows us to calculate the effective Hamiltonian of the unit cell placed at different positions within the sample,[58] by including the self-energy corrections that account for the presence of all subsequent unit cells. This involves a decimation procedure which is explained in detail in Ref. [59].

In Fig. 3, we show the illuminated graphite LDoS for different slab geometries as a function of the quasienergy ϵ and wavevector $\boldsymbol{k}$. Panels (a)–(c) and (e) refer to a sample which is finite along $\hat{\boldsymbol{a}}_2$, containing N_2 unit cells. In this case the corresponding Bravais lattice is rectangular, and given by primitive vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$. Therefore the primitive unit vectors of the reciprocal lattice coincide with those of the real lattice, and the wavevector can be written as $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1 + k_3\hat{\boldsymbol{a}}_3$. We evaluate the LDoS at the positions $n_2 = 1$ in (a) and (e), $N_2/2$ in (b) and N_2 in (c), respectively. In panels (a)–(c), we take the wave vector as $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1$ and fixed $k_3 = 0$, while in panel (e) we use $\boldsymbol{k} = k_3\hat{\boldsymbol{a}}_3$ and fixed $k_1 = \pi/\sqrt{3}a_0 + 0.4\ \text{Å}^{-1}$. The insets illustrate the regions where the LDoS is being evaluated: yellow rectangles denote the evaluation region and grey rectangles represent the graphene layers. In Fig. 3(d), we consider another geometry, where the sample is finite along $\hat{\boldsymbol{a}}_3$ and we evaluate the LDoS at the $n_3 = 1$ unit cell (see inset). In this case, the corresponding Bravais lattice is triangular, and we evaluate the LDoS for $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1$. As we use a huge value ($\sim 2^{25}$) for the amount of unit cells along the broken direction, the sample can be taken as semi-infinite in Figs. 3(a) and (c)-(e), while in Fig. 3(b) the sample is effectively infinite.

Figure 3 shows the laser induced gap around $\epsilon = \hbar\Omega/2 =$

1.1 eV (see red dashed lines) and four states crossing the gap in (a) and (c), while these peaks in the LDoS disappear in (b). In panel (d), there is a clear gap induced by the laser at the FZB, and no peaks crossing this region can be observed. These are clear signals of the presence of laser induced boundary states, located at those surfaces perpendicular to the graphene layers [although not shown, figures similar to (a)–(c) are obtained for a finite sample along $\hat{\boldsymbol{a}}_1$]. The shape of the bands in Fig. 3(b) suggests that the laser produces two gaps centered around different quasienergies, which could be attributed to the four band structure observed in Fig. 2. The effective gapped region corresponds to the intersection between the two gaps, and outside this region these states may strongly mix with the bands [see black arrows in Figs. 3(a) and 3(c)]. From the slope of the trajectories defined by the LDoS peaks in Figs. 3(a) and 3(c), we can infer that these states propagate along the $\hat{\boldsymbol{a}}_1$ direction and with opposite velocities, depending on the border which is being evaluated. Specifically, the peaks shown in (a) can be attributed to states localized around the $n_2 = 1$ border that propagate along $-\hat{\boldsymbol{a}}_1$, while the peaks in (b) correspond to states localized around the $n_2 = N_2$ border which propagate along $+\hat{\boldsymbol{a}}_1$, see violet arrows in the inset schemes. In Fig. 3(e), we can observe the evolution of the localized states as we move the wavevector along the stacking direction, i.e. $\boldsymbol{k} = (\pi/\sqrt{3}a_0 + 0.4\,\text{Å}^{-1})\hat{\boldsymbol{a}}_1 + k_3\hat{\boldsymbol{a}}_3$, in the same spatial region as in Fig. 3(a), i.e. $n_2 = 1$. The peaks reveal some dispersion (non-negligible slope), meaning that the boundary states also propagate along the stacked layers. However, for a given border, these peaks stay in the middle of the gap without crossing it, and the slopes developed by them take both positive and negative values (a similar behavior occurs for $n_2 = N_2$). This means that the sign of the group velocity along the stacking direction is not restricted to the border in which the state is localized, so the two directions (say, positive and negative) may coexist in the same border (see violet arrows in the inset).

Notice that a similar behavior is obtained in monolayer graphene,[18] where the circularly polarized laser induces chiral edge-states. By "chiral" is meant that the direction of propagation of the state depends on both the edge in which it is localized and the laser's handedness. In this sense, all the previous analysis indicates that in illuminated graphite there are also localized chiral states. We can continue this analogy and infer whether the observed localized states in graphite can be characterized by a topological invariant. This is presented in Appendix A, where we calculate the Chern number associated with the FZB for a simplified model of graphite that retains the leading hoppings γ_0, γ_1, and γ_3, and neglects all remaining (static) parameters. We are interested in the localized states generated by the mixing of the $n = 0$ and $n = 1$ replicas, so we truncate the Floquet space to these subspaces. Although higher-order mixings are also possible,[30] the associated gaps decrease very fast for the considered small laser intensity, and these contributions can be neglected for the purpose of the present discussion. Under this approximation it is possible to derive analytic expressions for the eigenenergies of the bulk Hamiltonian of Eq. (17), which allows us to identify the crossings between conduction and valence bands that belong to the $n = 0$ and $n = 1$ replicas, respectively. The main conclusion is that the contribution to the Chern number for a fixed value of k_3 is given by the number of bands that cross the FZB, multiplied by the sign τ of the polarization (thereby the chiral nature of these states). For the chosen frequency $\hbar\Omega = 2.2$ eV, this number results to be 4τ, in full agreement with the bulk-boundary correspondence, since the crossing bands are 1c–0, 1v–1, 2c–0 and 2v–1. Since we are computing only the contribution from the FZB gap to the topological invariant, it is implicitly assumed that the contributions from the bands below does not change. Considering all the contributions involves a more complex procedure as presented in Ref. [30] and is beyond our present scope.

So, what is new in this three-dimensional system? The first obvious difference with monolayer graphene is that now, rather than edge-states, what the LDoS peaks reveal are surface states located perpendicular to the planes defined by the graphene layers. To get an idea on how these surface states look, in Fig. 3(f) we evaluate the LDoS at $n_2 = 1$ for $\boldsymbol{k} = k_1\hat{\boldsymbol{a}}_1 + k_3\hat{\boldsymbol{a}}_3$ to picture out its shape in the *two* directions where translational invariance holds. This was done by fixing the quasienergy in steps of 0.005 eV within the range $1.03\text{ eV} \leq \epsilon \leq 1.15\text{ eV}$. We use a transparency scale (arbitrary units) to visualize all $\boldsymbol{k}$-points where the LDoS takes a large value, such that the obtained curves define what can be thought of as the 'skeleton' of the surface states. In fact, a close inspection for all energy steps when sweeping both k_1 and k_3 reveals two peaks which are separated each other, i.e. each peak defines an open trajectory. This suggests the presence of two surface states (in the shown region) which can be imagined as the natural dimensional extension of the chiral edge-states in graphene when adding an infinite number of layers. Another difference with monolayer graphene is that here the number of chiral states per value of k_3 is doubled, since now the gap comprises the crossing between four energy bands, due to the four basis sites in the unit cell. This, however, may change depending on the value of the chosen frequency. When $\hbar\Omega/2 \lesssim 0.25\gamma_0$ and $k_3 \sim 0$, it may happen that the bands that cross at the FZB are only 1c–0 and 1v–1, so the expected Chern number is in this case 2τ (see Appendix A).

In Fig. 4 we extract the maxima of the peaks of Fig. 3(f) and separate them in panels (a) and (b) to appreciate the surface states individually. The same is done in (c) and (d) for the LDoS evaluated at $n_2 = N_2$. The lines thus correspond to those boundary states that form the surface state for a fixed energy. We use the same energies as in Fig. 3(f) but these are distinguished through a colorscale ranging from red ($\epsilon = 1.15$ eV) to yellow ($\epsilon = 1.03$ eV). The shown plots thus resemble maps of equipotential lines (quasienergies) associated to the surface states. From the colorscale it is possible then to infer the group velocity of these states. Since

$$\boldsymbol{v}_{\text{g}}(\boldsymbol{k}) = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\epsilon_{\boldsymbol{k}}, \tag{23}$$

the group velocity points from yellow to red and perpendicular to the equipotential lines. From the plots it is easy to see that, in almost all cases, $\boldsymbol{v}_{\text{g}}$ points along the horizontal axis, i.e. $\hat{\boldsymbol{a}}_1$.

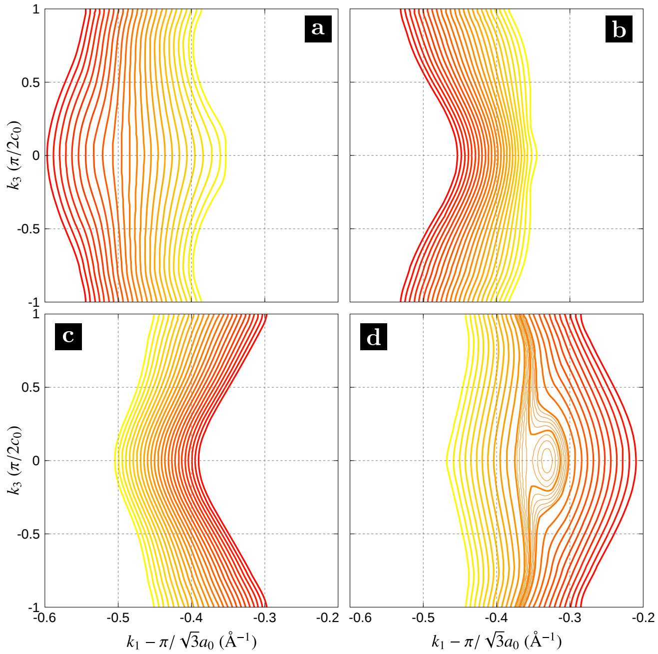


Figure 4. Laser induced surface states extracted from the LDoS. (a) and (b) show the two peaks of Fig. 3(f), respectively, corresponding to the LDoS at $n_2 = 1$. Similarly, (c) and (d) show the LDoS peaks when evaluated at $n_2 = N$. The used colorscale goes from red ($\epsilon = 1.15$ eV) to yellow ($\epsilon = 1.03$ eV).

In Fig. 4(d) there is, however, a particular region where a local minimum develops,[60] and the $\hat{\boldsymbol{a}}_3$ component of the group velocity dominates over the $\hat{\boldsymbol{a}}_1$ component at least locally. In any case, the mirror symmetry around $k_3 = 0$ implies that

$$\boldsymbol{v}_{\mathrm{g}}(k_1, k_3) \cdot \hat{\boldsymbol{a}}_3 = -\boldsymbol{v}_{\mathrm{g}}(k_1, -k_3) \cdot \hat{\boldsymbol{a}}_3, \tag{24}$$

meaning that for a given Fermi energy within the gapped region the overall velocity points along $\hat{\boldsymbol{a}}_1$ only. With this in mind, we can again conclude that these states are chiral, since the group velocity points towards opposite directions regarding to which border the surface state belongs.

All the above findings therefore enforce the idea that the physics behind illumination on graphite is, to some extent, similar to that of monolayer (or bilayer) graphene. The additional dimension present in this case contributes with a weak component of the group velocity along the new direction, which averages to zero when populating the system to the FZB. This is possibly due to the large separation between the stacked layers (c_0), as compared to the first-neighbor distance (a_0). The obtained surface states are rather continuations of graphene's edge-states in the stacking direction so in this sense one could say that these move through the boundary of the sample in an orderly manner. To which extent this is true is a question whose answer requires the evaluation of the LDoS along the boundary when both the $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ directions are finite. This obviously difficults the calculation of the LDoS as the effective dimension over which one operates is now dim $\mathcal{F} = 4N_3 \times (2n_r + 1)$, where N_3 is the number of unit cells along $\hat{\boldsymbol{a}}_3$. For small samples ($N_3 \sim 10$) this can be done in the same way we did before (i.e. an exact calculation), but for larger samples the previous strategy becomes very demanding (computationally speaking) and we employ an approximation scheme based on a decomposition into normal modes similar to that used in Refs. [61] and [34]. Although in graphite this decomposition scheme is not exact due to the next-nearest-neighbor couplings γ_2 and γ_5, deviations from the exact result can be considered as a small perturbation acting only on $n_3 = 1$ and $n_3 = N_3$, which can be neglected in large samples. This we explain in further details in App. B.

In Fig. 5, we show the LDoS evaluated at $n_2 = 1$ for different sample sizes, given by the number N_3 of unit cells along $\hat{\boldsymbol{a}}_3$. Panels (a)–(d) are the maps in the same k_1 region as in Fig. 3(a). In (a) and (b), we used the standard decimation procedure as in all previous calculations, while in (c) and (d), we used the normal mode decomposition explained in Appendix B. A comparison between Figs. 5(b) and 5(c) for $N_3 = 10$ shows that the used decomposition, though not exact, yields an accurate LDoS even in relatively small samples.

As expected, we can see that the peaks of Fig. 3 are also present in this case, maintaining the same chiral behavior as before. This is somewhat obvious when regarding the LDoS as decomposed by normal modes along $\hat{\boldsymbol{a}}_3$. Since this decomposition takes discrete values of k_3, c.f. Eq. (B1), the LDoS for a fixed k_3 is similar to that of Fig. 3(a), and the final LDoS is given as the sum of all mode contributions. For the considered region in the maps, then, the number of chiral edge-states crossing the gap simply goes as $4N_3$, as anticipated by the total Chern number of App. A. This is easy to see when N_3 is small, as it happens in Figs. 5(a)–5(c). For $N_3 = 100$, however, such a counting is no longer possible in Fig. 5(d) even if we would be able to increase the map resolution indefinitely. The reason for this is a rather subtle effect we did not comment so far. All LDoS peaks we have shown have, in fact, a finite width, which is independent of the chosen regularization energy η of Eq. (21).[62] To understand the origin of this width, notice that the localized states around $\epsilon = \hbar\Omega/2$ are produced by the coupling between $n = 0$ and $n = 1$ replicas. However, other extended states belonging to other replicas may be present within the gapped region. Strictly speaking, there is no real gap in the FZB where the localized states develop. However, we refer to the opening of $n = 0$ and $n = 1$ bands as a "gap" since the contributions coming from other replicas to the time-averaged LDoS are quite small. In other words, only when the replicas $n = 0$ and $n = 1$ are considered, the band opening at the FZB is a real gap. The observed width in the LDoS peaks then signals a small mixing term between the localized states (formed as a superposition of the $n = 0$ and $n = 1$ replicas) and extended states from other Floquet replicas (in this case the main contribution comes from $n = -1$ and $n = 2$). This implies that the localized states decay into the bulk upon absorption or emission of photons, in a characteristic time proportional to the inverse of the energy width of the peaks. Therefore, when the number of localized states is small, the mean level spacing is larger than their widths, and the system "recognizes" its finite size along $\hat{\boldsymbol{a}}_3$. When increasing N_3, at some point the level spacing becomes com-

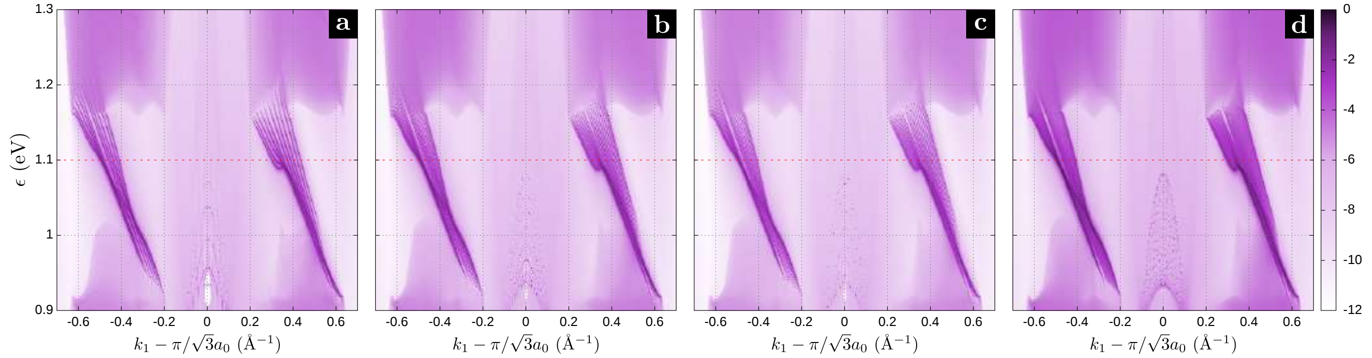


Figure 5. LDoS evaluated at $n_2 = 1$ for broken $\hat{\boldsymbol{a}}_2$ and $\hat{\boldsymbol{a}}_3$ directions. (a) and (b) are k vs ϵ LDoS maps for $N_3 = 5$ and 10, respectively, obtained from an exact calculation. (c) and (d) show the approximated LDoS through the normal mode decomposition for $N_3 = 10$ and 100, respectively. In all plots we normalized the densities to its maximum value and used a logarithmic scale. The red dashed lines at $\epsilon = 1.1$ eV denote the FZB.

parable to the energy width, and the system is no longer able to discern its finite size, so it behaves as a bulk in the staking direction. This originates the formation of localized states bands of Fig. 5(d), which may well be taken as surface states even in this limit of relatively small N_3.

V. LASER INDUCED PROBABILITY CURRENTS IN FINITE SYSTEMS

Another interesting effect that we would like to address is the fact that chiral states, by having a well-defined direction of propagation, are able to transport a probability current along the sample. This was shown in the context of illuminated monolayer graphene, where the laser-induced probability current appears either along the borders of the sample[63] or surrounds different types of defects like vacancies and adatoms.[64] Interestingly, such edge-states and their associated currents are able to be accessed by measuring the magnetic field they produce.[63] In graphite, therefore, similar effects can be naturally expected. To illustrate this, we consider a finite graphite sample consisting in a few hexagonal layers along the stacking direction. According to the discussion in Sec. II A, the quantity of interest, rather than the site current $J(\boldsymbol{r})$, is the bond current $J(\boldsymbol{r}, \boldsymbol{r}')$ given by Eq. (10).[54] The carbon bonds where this current is non-zero are thus given by those sites coupled by the Floquet Hamiltonian.

In order to identify the role of the laser illumination on the chirality of these currents, we also calculate the circulation of the bond currents through the lateral borders of the sample. This can be computed as the following discrete version of the line integral of the bond currents:

$$C_\alpha = \sum_{\boldsymbol{r},\boldsymbol{r}' \in S} J_\alpha(\boldsymbol{r}, \boldsymbol{r}'), \tag{25}$$

where α labels the Floquet state and the sum runs over all sites belonging to the border of the layers. In Fig. 6(a) we show the obtained circulation of the probability current for the hexagonal sample shown in Fig. 6(b). The dots in the plot are the obtained quasienergies from the eigenvalue equation, and we used a grayscale to indicate the weight of the Floquet eigenstate $|\alpha\rangle$ on the $n = 0$ replica, but with a minor change with respect to Eq. (20), i.e.

$$\bar{w}'_\alpha = 1 - 2\left|\bar{w}_\alpha - \frac{1}{2}\right|, \text{ where } \bar{w}_\alpha = \sum_{\boldsymbol{r}} |\langle \boldsymbol{r}, 0|\alpha\rangle|^2. \tag{26}$$

The idea behind this modification is to highlight the superposition of the $n = 0$ replica with the remaining ones: when $\bar{w}_\alpha = 1$, the Floquet state has full weight on the $n = 0$ replica, so there is no mixing with higher-order replicas ($\bar{w}'_\alpha = 0$), and when $\bar{w}_\alpha = 0$ the state has no weight on $n = 0$ and so again there is no mixing. The maximum value $\bar{w}'_\alpha = 1$ is reached when $\bar{w}_\alpha = 1/2$, meaning that the probability to find the system in the $n = 0$ replica is equal to that of finding it in all other replicas. Roughly speaking, $\bar{w}'_\alpha$ serves to infer where photon emission/absorption processes are more likely to occur. For the calculations, we used 366 carbon atoms per layer, so the total dimension of the truncated Floquet space is dim $\mathcal{F} = 366 \times 6 \times 5$, where the six corresponds to the number of layers and we considered five Floquet replicas, i.e. $n = -2, \ldots, 2$. For this example then, diagonalization of the Floquet Hamiltonian is a problem that can be treated exactly. However, as we discussed before in the context of the LDoS, for larger samples such a calculation may become seriously hard and one should move to the normal mode decomposition of Appendix B. We here took $\hbar\Omega = 8$ eV and $\zeta_0 = 0.71$ for the laser's parameters. Though these parameters may exceed standard values, we use them as to illustrate the effect in a relatively small sample. The same effects would be obtained for smaller parameters when used in larger samples, specially the size of the hexagonal layers, were a mode decomposition is not available.

In order to support the obtained circulation of the bond currents we show, in Fig. 7, the LDoS for the same geometry as that used in Fig. 3(a). This allows us to identify the boundary states appearing at the FZB and FZC gaps in this regime of laser parameters.

The main feature of Fig. 6(a) are the peaks of C_α in the vicinity of the FZB, defined at $\epsilon = \pm\hbar\Omega/2$ (red dashed lines). In *all* states within this region, the probability density circu-

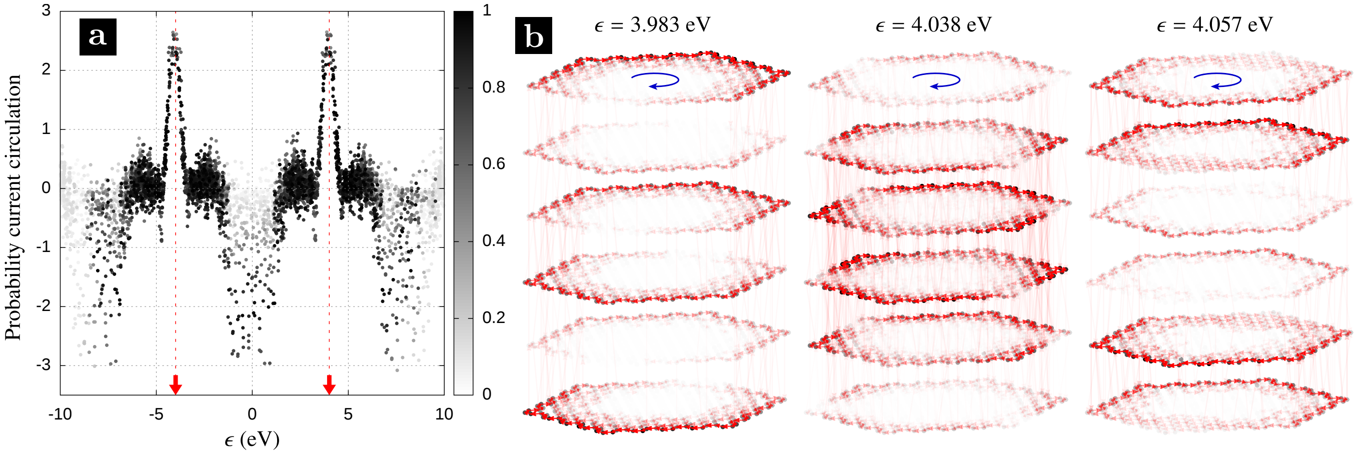


Figure 6. (a) Circulation of the time-averaged probability current as a function of the quasienergy in arbitrary units. The gray scale emphasizes the mixing of the $n = 0$ replica (see text). The red dashed lines at $\epsilon = \pm 4$ eV denote the FZB. (b) Examples of laser induced probability currents in few-layer graphite. The chosen geometry for the layers is hexagonal and we used $N_3 = 3$, i.e. six graphene layers. The position of the carbon atoms is represented by dots and we use a gray scale to denote the time-averaged probability density of Eq. (8). The corresponding probability bond currents (red arrows) are plotted in a transparency scale according to their magnitude. Blue arrows indicate the overall direction of the bond currents. The laser parameters were changed to $\hbar\Omega = 8$ eV and $\zeta_0 = 0.5\ \text{Å}^{-1} \times 1.42\ \text{Å} = 0.71$.

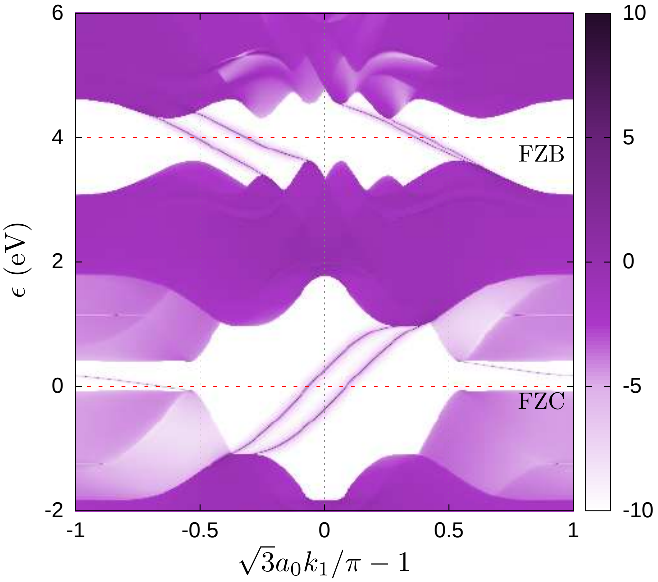


Figure 7. LDoS (in logarithmic scale) evaluated at $n_2 = 1$ for broken $\hat{\boldsymbol{a}}_2$ direction as in Fig. 3a. The used laser parameters coincide with those of Fig. 6. Red dashed lines at $\epsilon = 0$ and 4 eV denote the FZC and FZB, respectively.

lates through the boundaries of the hexagonal layers, with a given handedness. Obviously, if we change the sign of the circularly polarized waves, the direction of the bond currents is inverted, and with it the sign of the circulation. In addition, the mixing of the $n = 0$ replica in the peaks is large (i.e., $\bar{w}'_\alpha \sim 1$), which indicates a correlation between circulation and photon emission/absorption processes. In other words, the illuminated electrons are more likely to circulate in the energy regions where the interaction with the laser field becomes relevant. This also happens around the FZC, defined at $\epsilon = 0$. Although here the band crossings are more complicated than in the FZB, we can appreciate a negative circulation, though not all states are participating in this peak. In fact, we can identify some states with small (or even positive) circulation, which accordingly are weakly mixed. This can be attributed to the boundary states appearing in the FZC gap of Fig. 7, where we can see two states with positive group velocity and a single state with negative group velocity. Of course, the comparison between Figs. 6(a) and 7 can only be taken as qualitative, since for the LDoS we used a semi-infinite sample along the $\hat{\boldsymbol{a}}_2$ direction while for the bond currents we used a finite system.

In Fig. 6(b), we show the bond currents and probability densities for three Floquet states whose eigenenergies lie close to $\hbar\Omega/2$. The bond currents $J(\boldsymbol{r}, \boldsymbol{r}')$ are shown in red arrows that go from $\boldsymbol{r}'$ to $\boldsymbol{r}$, and we use a transparency scale to indicate its relative magnitude to the maximum current. For each carbon atom, we also calculated the time-averaged probability density $\rho(\boldsymbol{r})$ given in Eq. (8) and is shown through a gray scale. The resulting Floquet states around this energy region are clearly localized at the boundaries of the hexagonal layers. As Eq. (10) suggests, the bond currents are expected to be nonzero in those sites where $\rho(\boldsymbol{r})$ is appreciable, so they are also confined to the boundaries of the layers.

The bond currents' features discussed in this section are clear fingerprints of the chiral nature of the laser induced localized states. Interestingly, some differences appear when comparing these states with those found in illuminated monolayers. In fact, the magnitude of the bond currents in Fig. 6(b) is not constant along the full border of the hexagonal layers, but it rather alternates between successive layers. This is provided by some small, but non negligible, bond currents pointing along the stacking direction. Although this effect does not break the chirality of the localized states, the probability current displays nontrivial patterns due to the interlayer hopping amplitudes.

VI. SUMMARY AND FINAL REMARKS

To sum up, illumination by a circularly polarized laser on graphite generates boundary states. These boundary states turn out to be chiral, may form bands bridging the gap, and bear similarities and differences with those found in graphene. In the limit of large samples we show that a normal mode decomposition is applicable along the vertical direction. This provides a useful tool to reduce the 3D system onto a set of decoupled 2D subsystems where the z component of the wavevector enters as a fixed parameter. Under this decomposition scheme we were able to calculate the corresponding Chern number, which can be linked to the number of bands that intersect at the symmetry point $\epsilon = \hbar\Omega/2$. We highlight, however, two interesting features which we attribute to the extra dimension of the sample. First, we observe a smooth transition in the local density of states that goes from separable peaks (bundles) to the formation of bands of surface states, which evidence the three-dimensional nature of the sample even for relatively small N_3 values. This is attributed to a photon assisted decay of the localized states into extended states that belong to higher-order replicas. Second, the calculated probability currents may display intrincate patterns due to the small component along the stacking direction.

Regarding other possible stacking orders for graphite it should be noticed that, in principle, each crystal structure could present a topological structure of its own. However, given the hierarchical layered structure of graphite, we expect that in this case the main features observed for AB stacking should be kept. Notwithstanding, this is beyond the scope of our study, which remains non-exhaustive in this respect, motivating further investigations on illuminated multilayered systems.

We hope that the obtained results may stimulate further experimental research in strong light-matter interaction in graphite and related systems.

Acknowlegdments.– This work was supported by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Secretaría de Ciencia y Tecnología – Universidad Nacional de Córdoba (SECYT–UNC). HLC is member of CONICET. LEFFT acknowledges funding from FondeCyT (Chile) under project number 1170917 and 1170921, and by the EU Horizon 2020 research and innovation program under the Marie-Sklodowska-Curie Grant Agreement No. 873028 (HYDROTRONICS Project). JEBV acknowledges funding from PAIP Facultad de Química UNAM (grant 5000-9173).

Appendix A: Chern number calculation

In this section we sketch the calculation of the Chern number associated with the light induced band-gap openings around the crossing region $\epsilon = \hbar\Omega/2$, i.e., the Floquet zone boundary (FZB). To such end, we derive analytic expressions for the energies of those bands crossing at the FZB, under the mode decomposition scheme presented in App. B. This implies that in the present model we consider the most relevant hopping terms γ_0, γ_1 and γ_3, and neglect all remaining ones in the bulk Hamiltonian of Eq. (17). The contributions c_p to the Chern numbers are given by each band crossing p taking place at the FZB, when the laser is turned off, and can be obtained by reducing the Hamiltonian to those bands participating in the crossing. This yields a 2×2 effective Hamiltonian of the form

$$\hat{\mathcal{H}}_{p,\mathrm{eff}} = \boldsymbol{h}_p \cdot \hat{\boldsymbol{\sigma}}, \tag{A1}$$

with $\hat{\boldsymbol{\sigma}}$ the vector of Pauli matrices and $\boldsymbol{h}_p$ the associated vector to the p-crossing. The corresponding expression for c_p is the following[5]

$$c_p = \frac{1}{4\pi} \int \mathrm{d}^2\boldsymbol{k}\, \hat{\boldsymbol{h}}_p \cdot \left(\partial_{k_x} \hat{\boldsymbol{h}}_p \times \partial_{k_y} \hat{\boldsymbol{h}}_p \right), \tag{A2}$$

where the integral is taken over the first Brillouin zone for k_3 fixed as in Eq. (B1) and $\hat{\boldsymbol{h}}_p$ is the unit vector associated with $\boldsymbol{h}_p$. In order to obtain $\boldsymbol{h}_p$, we start with the bulk Floquet Hamiltonian of Eq. (19), and truncate the Floquet space to replicas $n = 0$ and $n = 1$. This can be computed as the following matrix:

$$\boldsymbol{H}_{\mathrm{F}} = \begin{pmatrix} \boldsymbol{H}^{(0)} - w & \boldsymbol{H}^{(1)} \\ \boldsymbol{H}^{(-1)} & \boldsymbol{H}^{(0)} + w \end{pmatrix}, \tag{A3}$$

where we shifted the energy origin to $w = \hbar\Omega/2$ so that the crossings we are interested in are placed at $\epsilon = 0$. The structure of these block matrices obey the form given in Eq. (17), i.e.

$$\boldsymbol{H}^{(n)} = \begin{pmatrix} 0 & \gamma_{12}^{(n)} & \gamma_{13}^{(n)} & 0 \\ \gamma_{21}^{(n)} & 0 & 0 & \gamma_{24}^{(n)} \\ \gamma_{31}^{(n)} & 0 & 0 & \gamma_{34}^{(n)} \\ 0 & \gamma_{42}^{(n)} & \gamma_{43}^{(n)} & 0 \end{pmatrix}. \tag{A4}$$

For the calculation of the hopping terms, we assume $\zeta_0 \ll 1$ so the phase introduced by the vector potential in Eq. (2) can be linearized as

$$e^{i2\pi(\boldsymbol{r}-\boldsymbol{r}')\cdot\boldsymbol{A}(t)/\Phi_0} \simeq 1 + i\zeta_0 \cos(\Omega t - \tau\phi_{\boldsymbol{r},\boldsymbol{r}'}), \tag{A5}$$

where $\tau = \pm 1$ denotes the laser's handedness. Following Eq. (2), we notice that in all cases we have $|\boldsymbol{r} - \boldsymbol{r}'| \sin\theta_{\boldsymbol{r},\boldsymbol{r}'} = a_0$, and hence $\zeta_{\boldsymbol{r},\boldsymbol{r}'} = \zeta_0$. Recalling that in the construction of the Floquet Hamiltonian we multiply these terms by $\exp(in\Omega t)$ and take the time-integral over one period, this yields for the above equation:

$$\delta_{n,0} + i\frac{\zeta_0}{2} e^{in\tau\phi_{\boldsymbol{r},\boldsymbol{r}'}} (\delta_{n,-1} + \delta_{n,1}). \tag{A6}$$

Therefore, the hopping terms can be specified by:

$$\gamma_{21}^{(0)} = \gamma_0 \left(1 + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right),$$

$$\gamma_{21}^{(1)} = i\frac{\zeta_0\gamma_0}{2} \left(e^{+i\tau\frac{1}{2}\pi} + e^{-i(\boldsymbol{k}\cdot\boldsymbol{a}_1 + \tau\frac{5}{6}\pi)} + e^{-i(\boldsymbol{k}\cdot\boldsymbol{a}_2 + \tau\frac{1}{6}\pi)} \right),$$

$$\gamma_{12}^{(1)} = i\frac{\zeta_0\gamma_0}{2} \left(e^{-i\tau\frac{1}{2}\pi} + e^{+i(\boldsymbol{k}\cdot\boldsymbol{a}_1 + \tau\frac{1}{6}\pi)} + e^{+i(\boldsymbol{k}\cdot\boldsymbol{a}_2 + \tau\frac{5}{6}\pi)} \right),$$

together with

$$\begin{aligned}\gamma_{31}^{(n)} &= \gamma_1\left(1+e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_3}\right)\delta_{n,0},\\ \gamma_{43}^{(n)} &= \gamma_{12}^{(n)}e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1},\\ \gamma_{42}^{(n)} &= \frac{\gamma_3}{\gamma_0}\gamma_{21}^{(n)}e^{+i\boldsymbol{k}\cdot\boldsymbol{a}_2}\left(1+e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_3}\right),\end{aligned}$$

and the general rule $\gamma_{ij}^{(n)} = [\gamma_{ji}^{(-n)}]^*$. With all these terms specified, we now construct the above Floquet Hamiltonian, and diagonalize the blocks $\boldsymbol{H}^{(0)}$. This gives the following eigenenergies:

$$\epsilon_{1,\mathrm{c}} = \sqrt{\frac{\alpha}{2}-\sqrt{\frac{\alpha^2}{4}-\beta}},\quad \epsilon_{2,\mathrm{c}} = \sqrt{\frac{\alpha}{2}+\sqrt{\frac{\alpha^2}{4}-\beta}},$$

for the conduction bands, while for the valence bands we have $\epsilon_{p,\mathrm{v}} = -\epsilon_{p,\mathrm{c}}$ for $p = 1, 2$, as in this model the e-h symmetry is preserved when γ_2, γ_4 and γ_5 are neglected. The terms in the above expressions are given by:

$$\begin{aligned}\alpha &= |\gamma_{21}|^2+|\gamma_{31}|^2+|\gamma_{42}|^2+|\gamma_{43}|^2,\\ \beta &= |\gamma_{21}|^2|\gamma_{43}|^2+|\gamma_{31}|^2|\gamma_{42}|^2-2\mathrm{Re}\left(\gamma_{13}\gamma_{34}\gamma_{42}\gamma_{21}\right),\end{aligned}$$

where we simplified the notation by taking $\gamma^{(0)} \to \gamma$, i.e., all hoppings in α and β correspond to the zeroth Fourier component. When including the w-term in these bands, we obtain the following two crossings:

$$\epsilon_{1,\mathrm{c}} - w = \epsilon_{1,\mathrm{v}} + w, \quad \text{and} \quad \epsilon_{2,\mathrm{c}} - w = \epsilon_{2,\mathrm{v}} + w. \tag{A7}$$

So we have that the bands that participate in the crossings are the conduction bands associated to the $n = 0$ replica and the valence bands for the $n = 1$ replica. The above mentioned e-h symmetry implies that the crossing conditions are simply given by $\epsilon_{p,\mathrm{c}} = w$, where $p = 1, 2$ now labels each band crossing.

The following step is to reduce the Floquet Hamiltonian to the found crossings. What we obtain then is the effective Hamiltonian as

$$\hat{\mathcal{H}}_{p,\mathrm{eff}} = \begin{pmatrix}\epsilon_{p,\mathrm{c}}-w & \gamma_{\mathrm{c,v}}^{(p,1)}\\ \gamma_{\mathrm{v,c}}^{(p,-1)} & \epsilon_{p,\mathrm{v}}+w\end{pmatrix}, \tag{A8}$$

where $\gamma_{\mathrm{v,c}}^{(p,-1)}$ and $\gamma_{\mathrm{c,v}}^{(p,1)}$ are obtained after applying the transformation matrix $\boldsymbol{U}$ that diagonalizes $\boldsymbol{H}^{(0)}$ on the coupling matrices $\boldsymbol{H}^{(\pm 1)}$ between the replicas, i.e.

$$\gamma_{\mathrm{v,c}}^{(p,-1)} = \langle p,\mathrm{v},1|\,\boldsymbol{U}^\dagger\boldsymbol{H}^{(-1)}\boldsymbol{U}\,|p,\mathrm{c},0\rangle\,, \tag{A9}$$

$$\gamma_{\mathrm{c,v}}^{(p,+1)} = \langle p,\mathrm{c},0|\,\boldsymbol{U}^\dagger\boldsymbol{H}^{(+1)}\boldsymbol{U}\,|p,\mathrm{v},1\rangle\,. \tag{A10}$$

In this way, we can identify the components of the $\boldsymbol{h}_p$ vector multiplying the Pauli matrices in Eq. (A1) as

$$\boldsymbol{h}_p = \left(\mathrm{Re}[\gamma_{\mathrm{v,c}}^{(p,-1)}], \mathrm{Im}[\gamma_{\mathrm{v,c}}^{(p,-1)}], \epsilon_{p,\mathrm{c}}-w\right). \tag{A11}$$

What follows in the calculation of c_p are the k_x and k_y derivatives, together with the integration over the first Brillouin zone. We carried out this numerically and obtained the Chern number depicted in Fig. 8(a). By way of comparison, we also show in (b) the LDoS for a semi-infinite geometry along the $\boldsymbol{a}_2$ direction, evaluated at $n_2 = 1$ and $\epsilon = \hbar\Omega/2$, as a function of k_1, with $k_3 = 0$. For the relevant parameter region of Ω, we find a perfect agreement between the number of chiral states and the calculated Chern number, i.e., bulk-boundary correspondence is verified. In addition, it is possible to observe that the chirality of the localized states is determined by the laser handedness, since the inversion $\tau \to -\tau$ naturally changes the sign of the Chern number. From the obtained result, we conclude that the contributions to the Chern number is τ times the number of bands crossing at the FZB. In fact, we distinguish four different regions for the Chern number, which coincide with those cases in which either the $p = 1$ or $p = 2$ bands cross this energy. For example, for $\hbar\Omega/2 \lesssim 0.25\,\gamma_0$, only the $p = 1$ bands can fulfill the crossing condition, so $c_1 = 2\tau$ and $c_2 = 0$. The opposite happens for $2.5\,\gamma_0 \lesssim \hbar\Omega/2 \lesssim 3.5\,\gamma_0$, where $c_1 = 0$ and $c_2 = 2\tau$. In this sense, we can say that the Chern number (and with it the number of localized states for a given k_3) signals the number of band crossings taking place at the FZB.

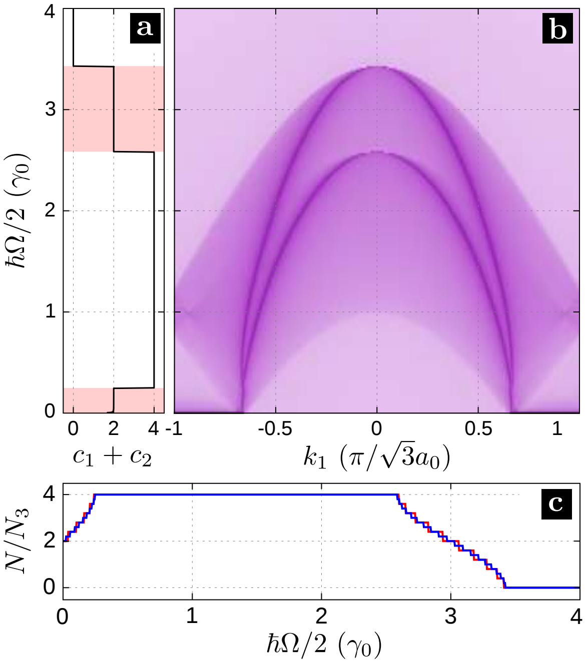


Figure 8. Localized states and Chern numbers as a function of the driving frequency Ω. (a) Contributions to the Chern number from the crossings $p = 1$ and $p = 2$. The red shaded areas denote transition regions where the Chern number may vary depending on the value of k_3. The laser intensity is $\zeta_0 = 0.01$ and we took $k_3 = 0$. (b) LDoS for broken $\boldsymbol{a}_2$ direction and evaluated at $n_2 = 1$ for $\epsilon = \hbar\Omega/2$. (c) Number of bands crossing at the FZB, divided by N_3, for $N_3 = 5$ (red) and $N_3 = 10$ (blue), in the limit $\zeta_0 = 0$.

Of course, the above analysis is valid under the assumption that the contributions from the stacking direction can be decomposed into normal modes, such that k_3 given by Eq. (B1) can be fairly taken as a fixed parameter. In this case, we notice that the interlayer hoppings depend on k_3, and therefore the Chern number can change from one normal mode to another. When adding up all contributions coming from the normal modes, the total Chern number

$$c_{\mathrm{FZB}} = \sum_{p=1}^{2} \sum_{n=1}^{N_3} c_{p,n}, \tag{A12}$$

varies in a similar way as in Fig. 8(a), but with the following differences: (1) such a quantity needs to be multiplied by N_3. (2) Around the transition region centered at $\hbar\Omega/2 = 0$, the total Chern number varies in a staggered way from $2\tau N_3$ to $4\tau N_3$, while around $\hbar\Omega/2 = 3\gamma_0$ this number changes from $4\tau N_3$ to 0. This behavior is shown in Fig. 8(c), where we calculate the number of bands that cross at the FZB as a function of the driving frequency for $N_3 = 5$ (solid red) and $N_3 = 10$ (solid blue).

We recognize that, in order to deal with a semi-analytic calculation for the Chern number, we worked in a simplified model of graphite where next-nearest-neighbor couplings and energy shifts between inequivalent carbon atoms were disregarded. The inclusion of these terms would only complicate such a calculation, though the main result would remain the same, namely, each band crossing at the FZB contributes with a factor 2τ to the total Chern number. With this in mind, we only expect some differences in the number of crossings near the transition regions of Fig. 8 (red shaded areas), as the bands experience slight modifications when including these terms. However, for the considered frequency value $\hbar\Omega = 2.2$ eV we used along this work, the Chern number would remain the same regardless of the value of k_3, so for finite samples along the stacking direction we expect $c_{\mathrm{FZB}} = 4\tau N_3$.

Appendix B: Normal mode decomposition

In this section we discuss the employed normal mode decomposition in the calculation of the local density of states shown in Fig. 4. Let us consider the bulk Hamiltonian $\hat{\mathcal{H}}_{\boldsymbol{k}}$ of Eq. (16) whose matrix representation is given in Eq. (17). We first break translational invariance along the z-direction, by considering N_3 unit cells along $\hat{\boldsymbol{a}}_3$. The set of allowed values for k_3 is no longer a continuum, and we expect some discrete set which we propose to be given by

$$k_3 = \frac{n\pi}{c_0(2N_3+1)}, \qquad n = 1, 2, \ldots, N_3. \tag{B1}$$

In this way, we obtain that the functions that depend on k_3 take the following values:

$$f_4 = 1 + e^{2i\varphi_n}, \text{ and } f_5 = 2\cos(2\varphi_n), \tag{B2}$$

where $\varphi_n = n\pi/(2N_3+1)$. Now, for every value n and fixed $\boldsymbol{k} = (k_x, k_y)$ we can diagonalize $\hat{\mathcal{H}}_{\boldsymbol{k}} \to \hat{\mathcal{H}}_{\boldsymbol{k}}^n$, where the superscript indicates that k_3 is given by n. This yields 4 eigenenergies and their corresponding eigenkets, i.e.

$$\hat{\mathcal{H}}_{\boldsymbol{k}}^n \left|\phi_{\alpha,\boldsymbol{k}}^n\right\rangle = \epsilon_{\alpha,\boldsymbol{k}}^n \left|\phi_{\alpha,\boldsymbol{k}}^n\right\rangle, \qquad \alpha = 1, \ldots, 4, \tag{B3}$$

where the eigenket can be written in terms of the site basis $i = \{\mathrm{A}_1, \mathrm{B}_1, \mathrm{A}_2, \mathrm{B}_2\}$ as

$$\left|\phi_{\alpha,\boldsymbol{k}}^n\right\rangle = \sum_i \phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i) \left|i\right\rangle, \tag{B4}$$

and $\phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i) = \langle \boldsymbol{\delta}_i | \phi_{\alpha,\boldsymbol{k}}^n \rangle$. What we do now is to translate these coefficients into a new space of dimension $4N_3$, given by the amount of units cells spanned along $\hat{\boldsymbol{a}}_3$. This is accomplished by transforming these coefficients as follows:

$$\phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i, n_3) = \frac{2}{\sqrt{N_3+1}} \phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i) \begin{cases} \sin\left[\dfrac{(2n_3-1)n\pi}{2N_3+1}\right] e^{-i\varphi_n/2}, & i \in \mathrm{LL} \\[2ex] \sin\left[\dfrac{2n_3 n\pi}{2N_3+1}\right] e^{+i\varphi_n/2}, & i \in \mathrm{UL} \end{cases} \equiv \phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i) \begin{cases} a_{n,n_3}, & i \in \mathrm{LL} \\ b_{n,n_3}, & i \in \mathrm{UL} \end{cases}, \tag{B5}$$

where the index $n_3 = 1, \ldots, N_3$ denotes the unit cell in the finite system. We can therefore construct the following states in this new space as

$$\left|\Phi_{\alpha,\boldsymbol{k}}^n\right\rangle = \sum_{n_3=1}^{N_3} \sum_i \phi_{\alpha,\boldsymbol{k}}^n(\boldsymbol{\delta}_i, n_3) \left|i, n_3\right\rangle. \tag{B6}$$

The idea then is to test such a transformation in the full Hamiltonian that arises when translational invariance along $\hat{\boldsymbol{a}}_3$ is broken. In terms of the $\{|n_3\rangle\}$ basis, this Hamiltonian presents the following structure:

$$\hat{\mathcal{H}}_{\boldsymbol{k}} = \sum_{n_3=1}^{N_3} \boldsymbol{h} \otimes |n_3\rangle\langle n_3| + \sum_{n_3=1}^{N_3-1} \left(\boldsymbol{v} \otimes |n_3+1\rangle\langle n_3| + \text{h.c.}\right),$$

where the block matrices $\boldsymbol{h}$ and $\boldsymbol{v}$ represent the intra- and inter-cell couplings, respectively, and are defined as:

$$\boldsymbol{h} = \begin{pmatrix} \epsilon_0 + \Delta & \gamma_0 f_1 & \gamma_1 & \gamma_4 f_2 \\ \gamma_0 f_1^* & \epsilon_0 & \gamma_4 f_1^* & \gamma_3 f_3 \\ \gamma_1 & \gamma_4 f_1 & \epsilon_0 + \Delta & \gamma_0 f_2 \\ \gamma_4 f_2^* & \gamma_3 f_3^* & \gamma_0 f_2^* & \epsilon_0 \end{pmatrix}, \tag{B7}$$

and

$$\boldsymbol{v} = \begin{pmatrix} \gamma_5 & 0 & \gamma_1 & \gamma_4 f_2 \\ 0 & \gamma_2 & \gamma_4 f_1^* & \gamma_3 f_3 \\ 0 & 0 & \gamma_5 & 0 \\ 0 & 0 & 0 & \gamma_2 \end{pmatrix} . \tag{B8}$$

If we now apply this Hamiltonian into the proposed state given by Eq. (B6), we obtain:

$$\hat{\mathcal{H}}_{\boldsymbol{k}} \left|\Phi^n_{\alpha,\boldsymbol{k}}\right\rangle = (\epsilon^n_{\alpha,\boldsymbol{k}} \hat{\mathcal{I}} + \hat{\mathcal{V}}) \left|\Phi^n_{\alpha,\boldsymbol{k}}\right\rangle , \tag{B9}$$

where $\hat{\mathcal{I}}$ is the identity operator in this extended space and

$$\hat{\mathcal{V}} = \gamma_5 \left(\hat{P}_{\mathrm{A}_1,1} + \hat{P}_{\mathrm{A}_2,N_3} \right) + \gamma_2 \left(\hat{P}_{\mathrm{B}_1,1} + \hat{P}_{\mathrm{B}_2,N_3} \right) ,$$

where we defined the projectors $\hat{P}_{i,n_3} = |i, n_3\rangle\langle i, n_3|$. The matrix associated with this operator is therefore diagonal, and the nonzero elements are only in the first ($n_3 = 1$) and last ($n_3 = N_3$) unit cells. The proposed decomposition scheme, therefore, is not exact due to the next-nearest-neighbor hoppings γ_5 and γ_2 appearing in $\hat{\mathcal{V}}$. However, as both γ_2 and γ_5 are much smaller than γ_0, the operator $\hat{\mathcal{V}}$ can be taken as a small perturbation on $\hat{\mathcal{H}}_{\boldsymbol{k}}$ when we increase N_3, such that it can be disregarded in a first approximation. This implies that the energies $\epsilon^n_{\alpha,\boldsymbol{k}}$, obtained from a 4×4 Hamiltonian matrix, are in fact a good approximation to the exact eigenenergies, which would be obtained from a $4N_3 \times 4N_3$ matrix.

It is important to notice that the above presented decomposition can be extended straightforwardly to incorporate the circularly polarized light. What changes in this case is that the static Bloch Hamiltonian in Eq. (B3) should be replaced by the Bloch-Floquet Hamiltonian of Eq. (18), and the state $|\phi^n_{\alpha,\boldsymbol{k}}\rangle$ is now defined in the $\mathcal{F}$-space, whose dimension is $4(2n_r + 1)$ (recall that $2n_r + 1$ is the amount of considered Floquet replicas). Additionally, for the LDoS of Fig. 4, translational invariance is not only broken along $\hat{\boldsymbol{a}}_3$, but also in $\hat{\boldsymbol{a}}_2$. For a given k_3, specified by n, we can calculate an effective local density $\mathcal{N}_{\boldsymbol{r},n}$ by following the decimation procedure discussed in detail in Ref. [59]. This procedure consists in the recursive calculation of the self-energy correction on the site located at $\boldsymbol{r}_0 = n_2 \boldsymbol{a}_2 + \boldsymbol{\delta}_i$, due to the presence of the other sites in the lattice. Once we obtain $\mathcal{N}_{\boldsymbol{r},n}$, the final LDoS at site $\boldsymbol{r} = \boldsymbol{r}_0 + n_3 \boldsymbol{a}_3$ can be obtained as:

$$\mathcal{N}_{\boldsymbol{r}} = \sum_n \mathcal{N}_{\boldsymbol{r}_0,n} \begin{cases} |a_{n,n_3}|^2, & i \in \mathrm{LL} \\ |b_{n,n_3}|^2, & i \in \mathrm{UL} \end{cases} . \tag{B10}$$

The relevance of this decomposition scheme relies on the fact that it effectively reduces the dimension of the involved Hamiltonians, and thus the computation time demanded by the calculation of either the system's eigenenergies or the LDoS. This scheme, in turn, yields a very good approximation to the exact solutions for large values of N_3, such that surface effects due to the perturbation $\hat{\mathcal{V}}$ can be neglected. It is precisely in this limit where the exact calculation becomes highly demanding and, in most of cases, almost impossible to carry out.

[1] J. D. Martin, "When condensed-matter physics became king," Phys. Today **72**, 30 (2019).

[2] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," Science **306**, 666 (2004).

[3] "Topology on top," Nat. Phys. **12**, 615 (2016).

[4] D. Castelvecchi, "The strange topology that is reshaping physics," Nature **547**, 272 (2017).

[5] M. Z. Hasan and C. L. Kane, "Colloquium: Topological insulators," Rev. Mod. Phys. **82**, 3045 (2010).

[6] D. Xiao, M.-C. Chang, and Q. Niu, "Berry phase effects on electronic properties," Rev. Mod. Phys. **82**, 1959 (2010).

[7] J. K. Asbóth, L. Oroszlány, and A. Pályi, A Short Course on Topological Insulators, Lecture Notes in Physics, Vol. 919 (Springer International Publishing, Cham, 2016).

[8] S.-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C.-C. Lee, S.-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, "Discovery of a Weyl fermion semimetal and topological Fermi arcs," Science **349**, 613 (2015).

[9] M. Zahid Hasan, S.-Y. Xu, I. Belopolski, and S.-M. Huang, "Discovery of Weyl fermion semimetals and topological Fermi arc states," Annu. Rev. of Condens. Matter Phys. **8**, 289 (2017).

[10] T. Oka and H. Aoki, "Photovoltaic Hall effect in graphene," Phys. Rev. B **79**, 081406(R) (2009).

[11] N. H. Lindner, G. Refael, and V. Galitski, "Floquet topological insulator in semiconductor quantum wells," Nat. Phys. **7**, 490 (2011).

[12] S. V. Syzranov, M. V. Fistul, and K. B. Efetov, "Effect of radiation on transport in graphene," Phys. Rev. B **78**, 045407 (2008).

[13] F. J. López-Rodríguez and G. G. Naumis, "Analytic solution for electrons and holes in graphene under electromagnetic waves: Gap appearance and nonlinear effects," Phys. Rev. B **78**, 201406(R) (2008).

[14] H. L. Calvo, H. M. Pastawski, S. Roche, and L. E. F. Foa Torres, "Tuning laser-induced band gaps in graphene," Appl. Phys. Lett. **98**, 232103 (2011).

[15] Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik, "Observation of Floquet-Bloch states on the surface of a topological insulator," Science **342**, 453 (2013).

[16] T. Kitagawa, T. Oka, A. Brataas, L. Fu, and E. Demler, "Transport properties of nonequilibrium systems under the application of light: Photoinduced quantum Hall insulators without Landau levels," Phys. Rev. B **84**, 235108 (2011).

[17] A. Gómez-León and G. Platero, "Floquet-Bloch Theory and Topology in Periodically Driven Lattices," Phys. Rev. Lett. **110**, 200403 (2013).

[18] P. M. Perez-Piskunow, G. Usaj, C. A. Balseiro, and L. E. F. Foa Torres, "Floquet chiral edge states in graphene," Phys. Rev. B **89**, 121401(R) (2014).

[19] M. S. Rudner and N. H. Lindner, "Floquet topological insulators: from band structure engineering to novel non-equilibrium quantum phenomena," (2019), arXiv:1909.02008 [cond-mat.mes-hall].

[20] M. S. Rudner and J. C. W. Song, "Self-induced Berry flux and spontaneous non-equilibrium magnetism," Nat. Phys. **15**, 1017 (2019).

[21] F. Mahmood, C.-K. Chan, Z. Alpichshev, D. Gardner, Y. Lee, P. A. Lee, and N. Gedik, "Selective scattering between Floquet-Bloch and Volkov states in a topological insulator," Nat. Phys. **12**, 306 (2016).

[22] J. W. McIver, B. Schulte, F.-U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, "Light-induced anomalous Hall effect in graphene," Nat. Phys. **16**, 38 (2020).

[23] J. H. Shirley, "Solution of the Schrödinger Equation with a Hamiltonian Periodic in Time," Phys. Rev. **138**, B979 (1965).

[24] H. Sambe, "Steady states and quasienergies of a quantum-mechanical system in an oscillating field," Phys. Rev. A **7**, 2203 (1973).

[25] S. Kohler, J. Lehmann, and P. Hänggi, "Driven quantum transport on the nanoscale," Phys. Rep. **406**, 379 (2005).

[26] G. Platero and R. Aguado, "Photon-assisted transport in semiconductor nanostructures," Phys. Rep. **395**, 1 (2004).

[27] J. C. Sandoval-Santana, V. G. Ibarra-Sierra, J. L. Cardoso, A. Kunold, P. Roman-Taboada, and G. G. Naúmis, "Dynamical Algebra: Method for Finding the Exact Effective Hamiltonian of Time-Driven Quantum Systems (Ann. Phys. 8/2019)," Ann. Phys. (Berlin) **531**, 1970029 (2019).

[28] M. Vogl, M. Rodriguez-Vega, and G. A. Fiete, "Effective Floquet Hamiltonian in the low-frequency regime," Phys. Rev. B **101**, 024303 (2020).

[29] M. S. Rudner, N. H. Lindner, E. Berg, and M. Levin, "Anomalous edge states and the bulk-edge correspondence for periodically-driven two dimensional systems," Phys. Rev. X **3**, 031005 (2013).

[30] P. M. Perez-Piskunow, L. E. F. Foa Torres, and G. Usaj, "Hierarchy of Floquet gaps and edge states for driven honeycomb lattices," Phys. Rev. A **91**, 043625 (2015).

[31] A. Kundu and B. Seradjeh, "Transport Signatures of Floquet Majorana Fermions in Driven Topological Superconductors," Phys. Rev. Lett. **111**, 136402 (2013).

[32] L. E. F. Foa Torres, P. M. Perez-Piskunow, C. A. Balseiro, and G. Usaj, "Multiterminal Conductance of a Floquet Topological Insulator," Phys. Rev. Lett. **113**, 266801 (2014).

[33] E. A. Rodríguez-Mena and L. E. F. Foa Torres, "Topological signatures in quantum transport in anomalous Floquet-Anderson insulators," Phys. Rev. B **100**, 195429 (2019).

[34] H. L. Calvo, J. S. Luna, V. Dal Lago, and L. E. F. Foa Torres, "Robust edge states induced by electron-phonon interaction in graphene nanoribbons," Phys. Rev. B **98**, 035423 (2018).

[35] S. Chaudhary, A. Haim, Y. Peng, and G. Refael, "Phonon-induced Floquet second-order topological phases protected by space-time symmetries," (2019), arXiv:1911.07892 [cond-mat.mes-hall].

[36] H. Dehghani, T. Oka, and A. Mitra, "Dissipative Floquet topological systems," Phys. Rev. B **90**, 195429 (2014).

[37] H. Dehghani, T. Oka, and A. Mitra, "Out-of-equilibrium electrons and the Hall conductance of a Floquet topological insulator," Phys. Rev. B **91**, 155422 (2015).

[38] G. Usaj, P. M. Perez-Piskunow, L. E. F. Foa Torres, and C. A. Balseiro, "Irradiated graphene as a tunable Floquet topological insulator," Phys. Rev. B **90**, 115423 (2014).

[39] M. Tahir, Q. Y. Zhang, and U. Schwingenschlögl, "Floquet edge states in germanene nanoribbons," Sci. Rep. **6**, 31821 (2016).

[40] M. Ezawa, "Photoinduced Topological Phase Transition and a Single Dirac-Cone State in Silicene," Phys. Rev. Lett. **110**, 026603 (2013).

[41] A. Huamán and G. Usaj, "Floquet spectrum and two-terminal conductance of a transition-metal dichalcogenide ribbon under a circularly polarized laser field," Phys. Rev. B **99**, 075423 (2019).

[42] V. K. Kozin, I. V. Iorsh, O. V. Kibis, and I. A. Shelykh, "Periodic array of quantum rings strongly coupled to circularly polarized light as a topological insulator," Phys. Rev. B **97**, 035416 (2018).

[43] B. M. Fregoso, J. P. Dahlhaus, and J. E. Moore, "Dynamics of tunneling into non-equilibrium edge states," Phys. Rev. B **90**, 155127 (2014).

[44] H. L. Calvo, L. E. F. Foa Torres, P. M. Perez-Piskunow, C. A. Balseiro, and G. Usaj, "Floquet interface states in illuminated three-dimensional topological insulators," Phys. Rev. B **91**, 241404(R) (2015).

[45] J. González and R. A. Molina, "Macroscopic Degeneracy of Zero-Mode Rotating Surface States in 3d Dirac and Weyl Semimetals under Radiation," Phys. Rev. Lett. **116**, 156803 (2016).

[46] X.-X. Zhang, T. T. Ong, and N. Nagaosa, "Theory of photoinduced Floquet Weyl semimetal phases," Phys. Rev. B **94**, 235137 (2016).

[47] S. K. Firoz Islam and A. A. Zyuzin, "Photoinduced interfacial chiral modes in threefold topological semimetal," Phys. Rev. B **100**, 165302 (2019).

[48] K. von Klitzing, G. Dorda, and M. Pepper, "New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance," Phys. Rev. Lett. **45**, 494 (1980).

[49] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized Hall Conductance in a Two-Dimensional Periodic Potential," Phys. Rev. Lett. **49**, 405 (1982).

[50] B. I. Halperin, "Possible States for a Three-Dimensional Electron Gas in a Strong Magnetic Field," Jpn. J. Appl. Phys. **26**, 1913 (1987).

[51] F. Tang, Y. Ren, P. Wang, R. Zhong, J. Schneeloch, S. A. Yang, K. Yang, P. A. Lee, G. Gu, Z. Qiao, and L. Zhang, "Three-dimensional quantum Hall effect and metal–insulator transition in ZrTe$_5$," Nature **569**, 537 (2019).

[52] M. Dresselhaus and G. Dresselhaus, "Intercalation compounds of graphite," Adv. Phys. **30**, 139 (1981).

[53] Throughout the manuscript we will take $\tau = 1$, unless otherwise indicated.

[54] T. N. Todorov, "Tight-binding simulation of current-carrying nanostructures," J. Phys.: Condens. Matter **14**, 3049 (2002).

[55] C.-M. Cheng, C.-J. Hsu, J.-L. Peng, C.-H. Chen, J.-Y. Yuh, and K.-D. Tsuei, "Tight-binding parameters of graphite determined with angle-resolved photoemission spectra," Appl. Surf. Sci. **354**, 229 (2015).

[56] S. A. Sato, J. W. McIver, M. Nuske, P. Tang, G. Jotzu, B. Schulte, H. Hübener, U. De Giovannini, L. Mathey, M. A. Sentef, A. Cavalleri, and A. Rubio, "Microscopic theory for the light-induced anomalous Hall effect in graphene," Phys. Rev. B **99**, 214302 (2019).

[57] L. E. F. Foa Torres, "Mono-parametric quantum charge pumping: Interplay between spatial interference and photon-assisted tunneling," Phys. Rev. B **72**, 245339 (2005).

[58] For example, if the slab is finite along $\hat{\boldsymbol{a}}_1$, we can evaluate $\mathcal{N}_{\boldsymbol{r},\boldsymbol{k}}$ for $\boldsymbol{r} = n_1\boldsymbol{a}_1 + \boldsymbol{\delta}_i$, with $n_1 = 1, \ldots, N_1$ and $\boldsymbol{\delta}_i$ the basis sites.

[59] H. L. Calvo, P. Perez-Piskunow, H. M. Pastawski, S. Roche, and L. E. F. Foa Torres, "Non-perturbative effects of laser illumination on the electrical properties of graphene nanoribbons," J. Phys.: Condens. Matter **25**, 144202 (2013).

[60] We increased the resolution in steps of $\epsilon = 0.001$ eV within the range 1.085 eV $\leq \epsilon \leq$ 1.095 eV to identify this local minimum.

[61] C. G. Rocha, L. E. F. Foa Torres, and G. Cuniberti, "ac transport

in graphene-based Fabry-Pérot devices," Phys. Rev. B **81**, 115435 (2010).

[62] In all LDoS calculations, we use $\eta = 1 \times 10^{-5}$ eV, as this reaches the thermodynamic limit for the number $N \sim 2^{25}$ of considered unit cells.

[63] J. P. Dahlhaus, B. M. Fregoso, and J. E. Moore, "Magnetization signatures of light-induced quantum hall edge states," Phys. Rev. Lett. **114**, 246802 (2015).

[64] D. A. Lovey, G. Usaj, L. E. F. Foa Torres, and C. A. Balseiro, "Floquet bound states around defects and adatoms in graphene," Phys. Rev. B **93**, 245434 (2016).

Quantum Hall edge states under periodic driving: a Floquet induced chirality switch

A. Huamán,[1,2] L. E. F. Foa Torres,[3] C. A. Balseiro,[1,2] and Gonzalo Usaj[1,2]

[1]*Centro Atómico Bariloche and Instituto Balseiro, Comisión Nacional de Energía Atómica (CNEA)–Universidad Nacional de Cuyo (UNCUYO), 8400 Bariloche, Argentina*
[2]*Instituto de Nanociencia y Nanotecnología (INN), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)–CNEA, 8400 Bariloche, Argentina*
[3]*Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile*

(Dated: December 15, 2020)

We report on the fate of the quantum Hall effect in graphene under strong laser illumination. By using Floquet theory combined with both a low energy description and full tight-binding models, we clarify the selection rules, the quasienergy band structure, as well as their connection with the two-terminal and multi-terminal conductance in a device setup as relevant for experiments. We show that the well-known dynamical gaps that appear in the Floquet spectrum at $\pm\hbar\Omega/2$ lead to a switch-off of the quantum Hall edge transport for different edge terminations except for the armchair one, where two terms cancel out exactly. More interestingly, we show that near the Dirac point changing the laser polarization (circular right or circular left) controls the Hall conductance, by allowing to switch it on or off, or even by flipping its sign, thereby reversing the chirality of the edge states. This might lead to new avenues to fully control topologically protected transport.

I. INTRODUCTION

Forty years ago, the discovery of the precise quantization of the Hall conductance in a two-dimensional electron gas under extreme conditions [1] opened the doors to a new chapter in condensed matter physics [2]. Elegant topological arguments [3, 4] explained the precision of the Hall plateaus in practical devices under high perpendicular magnetic fields, while also pointing to new deeper and unifying concepts. Over the last two decades, the use of such topological arguments rapidly expanded [5–7] allowing the discovery of, for example, topological insulators in two [8] and three dimensions [9] and Weyl semimetals [10]. Amid the ever growing family of topological phases, the quantum Hall (QH) effect remains as a paradigmatic case where the topological edge states enjoy the highest degree of robustness, a fact that is nowadays exploited in the new international system of units [11].

Besides the plethora of manifestations of topological states in (or near) equilibrium conditions, another growing research front aims at using light to change the properties of a material by generating hybrid electron-photon states (also called Floquet-Bloch states) with different spectral and topological properties [12–18]. The latter has become an emerging research front within the so-called quantum materials [19]. Fascinating experiments have unveiled the Floquet-Bloch states [20, 21] and a much awaited consequence: the light-induced Hall effect [22]. While in photonic systems or ultracold matter the experiments allow to reach high driving [23] frequencies, which in turn allow suitable theoretical approximations [24, 25], the sweet spot for laser-illuminated Dirac materials corresponds to the (theoretically more challenging) mid-infrared where $\hbar\Omega$ (a few hundreds meV) is much smaller than the bandwidth [26, 27].

Here we address the question of how the QH effect in graphene is affected by laser illumination. Previous studies have mainly focused on the effect of light on the intricacies of the Hofstadter butterfly of different lattices [28–31], the bulk properties of the irradiated Landau levels and topological invariants [28], and related dynamics [32]. By computing the spectrum and the topological invariants, laser-induced modifications on the Hofstadter butterfly and topological properties were recognized. The Hall conductivity in presence of both illumination and an external magnetic field (but without dissipation terms) was also discussed but by means of a generalized Kubo formula [33]. The subject still remains controversial, as the issue of how to properly account for the occupation of the Floquet bands [34–40], specially in this bulk regime when dissipation effects need to be included, has demonstrated to be a difficult task.

In this work we tackle two aspects that are unavoidable in condensed matter experiments: (*i*) the regime of photon energies ($\hbar\Omega$) much smaller than the bandwidth—in particular, we consider $\Omega \leq \omega_c$, where ω_c is the cyclotron frequency; and (*ii*) a multi-terminal device geometry with a laser spot applied to the central part and address the conductance measured in such configurations [41]. Specifically, we study the spectrum and the time-averaged conductance both in the two-terminal and multi-terminal case as required for Hall measurements. To such end we use atomistic models within a scattering configuration with a central illuminated spot, thereby allowing for the occupations to be well defined far away in the leads. The spectrum is first analyzed by using the continuous Dirac model that properly describes the low energy properties of the system. This is done for both zigzag and armchair edge terminations. Such an approach allows us to clearly identify the main features of the Landau-Floquet edge modes. These results are further verified using a more complete tight-binding model which is later used for our transport calculations. The latter are implemented by means of a generalization of the coherent Landauer-Büttiker approach to the Floquet picture [42–44]. In this Floquet scattering picture, the leads are not illuminated and have well defined occupations.

We find that in certain experimentally accessible parameter regions laser illumination leads to important effects including the switch-off of the Hall conductance, the splitting of the Hall

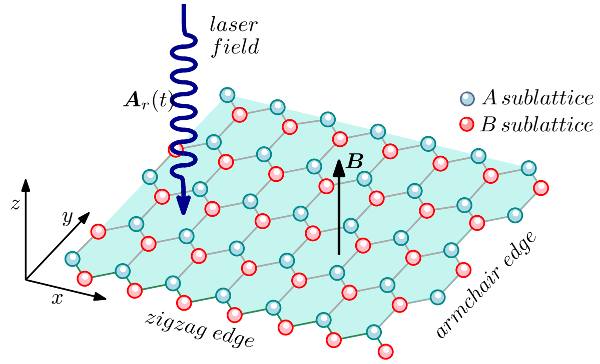


Figure 1. (Color online) Geometry used in the Dirac model. The homogeneous magnetic field and the laser are normal to the graphene monolayer. The zigzag and armchair edges are indicated, highlighting the fact that the former contains only C atoms from a given sublattice (B), while the latter contains both.

plateaus and even a change in the chirality of the propagating states. Interestingly, the Hall conductance is switched off for all edge terminations except for the armchair one, where two contributions cancel out exactly.

This paper is organized as follows. In Sec. II we describe the Landau-Floquet states in graphene within the framework of the Dirac (linear) model. In Sec. III we study the Landau-Floquet bands in a tight binding model for zigzag and armchair ribbons. Sec. IV shows two-terminal conductance simulations which are clarified by visualizing the scattering states. Sec. V shows the simulations of the Hall conductance in a six-terminal configuration. Finally, we summarize our results in Sec. VI.

II. LOW ENERGY HAMILTONIAN

The low energy properties of graphene can be described using the following Hamiltonian [45, 46],

$$\hat{\mathcal{H}}_0 = v_F \left(\tau_z \otimes \sigma_x \,\hat{\boldsymbol{x}} + \tau_0 \otimes \sigma_y \,\hat{\boldsymbol{y}}\right) \cdot \boldsymbol{p} \,, \tag{1}$$

where σ_i (τ_i) with $i = x, y, z$ are Pauli matrices describing the pseudospin (valley) degree of freedom, τ_0 is the 2×2 identity matrix, $\boldsymbol{p} = p_x \,\hat{\boldsymbol{x}} + p_y \,\hat{\boldsymbol{y}}$ is the momentum operator and v_F is the Fermi velocity. The wavefunction Ψ has then four components, $\Psi = [\psi_{AK}, \psi_{BK}, \psi_{AK'}, \psi_{BK'}]^{\mathrm{T}}$, with amplitudes describing the two inequivalent valleys in the Brillouin zone around $K = (4\pi/3\sqrt{3}a_0, 0)$ and $K' = -K$ (the first two amplitudes correspond to K and the remaining ones to K'). The parameter a_0 is the distance between nearest neighbor carbon atoms. The presence of a perpendicular magnetic field, $\boldsymbol{B} = B\,\hat{\boldsymbol{z}}$, can be described by the well-known Peierls substitution, $\boldsymbol{p} \rightarrow \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}$, with $\boldsymbol{A}$ the corresponding vector potential ($-e$ is the electron charge, $e > 0$).

A. The Floquet approach

The illumination with a laser field (applied perpendicularly to the graphene plane) can be modeled as a time-dependent term in the Hamiltonian. Furthermore, as long as the laser is monochromatic (as it will be the case throughout this work), this term is periodic in time and hence it can be treated within the Floquet theory [47–49]. We briefly describe now this approach before going into its application to our problem.

For a periodic time-dependent Hamiltonian $\hat{\mathcal{H}}(t)$, where $\hat{\mathcal{H}}(t + T) = \hat{\mathcal{H}}(t)$ with the period $T = 2\pi/\Omega$, Floquet theory assures the existence of a complete set of solutions of the form $|\Psi_\alpha(t)\rangle = \mathrm{e}^{-i\varepsilon_\alpha t/\hbar}\,|\phi_\alpha(t)\rangle$ with $|\phi_\alpha(t)\rangle = |\phi_\alpha(t+T)\rangle$. Replacing this solution in the time-dependent Schrödinger equation one obtains: $\hat{\mathcal{H}}_F\,|\phi_\alpha(t)\rangle = \varepsilon_\alpha\,|\phi_\alpha(t)\rangle$, where $\hat{\mathcal{H}}_F = \hat{\mathcal{H}}(t) - i\hbar\,\partial_t$ is called the Floquet Hamiltonian. Thus, we get an eigenvalue equation in the composite space $\mathcal{R}\otimes\mathcal{T}$ (also called Floquet space) where $\mathcal{R}$ is the usual Hilbert space and $\mathcal{T}$ the space of $T-$periodic functions spanned by $\exp{(im\Omega t)}$. The integer m is called the *replica* index. The change in the replica index in a process going from a state with, say, m to $m{+}n$ can be assimilated to a number of photon excitations [50].

Our Hamiltonian can be written as the sum of a time-independent term and a time-dependent one involving the interaction with the laser: $\hat{\mathcal{H}}(t) = \hat{\mathcal{H}}_0 + \hat{V}(t)$. By using a Peierls' substitution the time dependent term can be written as $\hat{V}(t) = \frac{e\,v_F}{c}\left(\tau_z \otimes \sigma_x \,\hat{\boldsymbol{x}} + \tau_0 \otimes \sigma_y \,\hat{\boldsymbol{y}}\right) \cdot \boldsymbol{A}_r(t)$, where the vector potential

$$\boldsymbol{A}_r(t) = A_0 \left[\cos\alpha \cos\Omega t \,\hat{\boldsymbol{x}} + \sin\alpha\cos(\Omega t - \varphi)\,\hat{\boldsymbol{y}}\,\right] \tag{2}$$

describes the radiation field. For $\varphi = 0$, this radiation field is linearly polarized, in which case α is the polarization angle; whereas for $\varphi = \pi/2$ ($-\pi/2$), and $\alpha = \pi/4$, the radiation is right-handed (left-handed) circularly polarized. The Fourier components of $\hat{V}(t)$, $\hat{V}_n = \frac{1}{T}\int_0^T \hat{V}(t)\mathrm{e}^{-in\Omega t}dt$, introduce elements connecting the different Floquet replicas in the Floquet Hamiltonian.

In the low energy approximation there is a further simplification: since the momentum $\boldsymbol{p}$ enters linearly, the perturbation $\hat{V}(t)$ is monochromatic so that its Fourier expansion will have only one harmonic (there is no such a simplification in the tight-binding model, see Sec. III). Notice also that any spatial modulation of the laser beam is considered to be larger than all other relevant length scales and hence ignored. This implies that the perturbation cannot mix states that are spatially orthogonal. Finally, we write $\hat{V}(t)$ as

$$\hat{V}(t) = (\mathcal{V}\,e^{i\Omega t} + \mathcal{V}^\dagger\,e^{-i\Omega t}), \tag{3}$$

with

$$\mathcal{V} = \frac{\eta\hbar\Omega}{2}\left[\cos\alpha\,\tau_z \otimes \sigma_x + \sin\alpha\,e^{-i\varphi}\,\tau_0 \otimes \sigma_y\right]. \tag{4}$$

Here we have defined the dimensionless parameter $\eta = ev_F A_0/c\hbar\Omega$ that characterizes the intensity of the perturbation. With this, our eigenvalue equation reduces to

$$(\hat{\mathcal{H}}_0+m\hbar\Omega)\,|\phi_{\alpha,m}\rangle+\mathcal{V}^\dagger|\phi_{\alpha,m+1}\rangle+\mathcal{V}|\phi_{\alpha,m-1}\rangle=\varepsilon_\alpha\,|\phi_{\alpha,m}\rangle\;, \tag{5}$$

where $|\phi_\alpha(t)\rangle=\sum_m e^{im\Omega t}\,|\phi_{\alpha,m}\rangle$. From this it is clear that the laser field can only couple replicas m_i and m_f such that $m_f=m_i\pm1$. For the application of the Floquet formalism, we will expand $|\phi_{\alpha,m}\rangle$ in a basis of eigenfunctions $|\boldsymbol{\chi}_n\rangle$ of the static system (i.e., with the magnetic field alone)

$$|\phi_{\alpha,m}\rangle=\sum_n w_{mn}^{(\alpha)}\,|\boldsymbol{\chi}_n\rangle\,. \tag{6}$$

The eigenstates $|\boldsymbol{\chi}_n\rangle$ can be those corresponding to a system with an edge (see Sec. II B and II C) or to an infinite (bulk) sample (see Sec. II D). From hereon the letter m will be reserved to indicate Floquet replicas.

B. Zigzag Floquet Hall states

We start our analysis with the most relevant case of zigzag edges (see Fig. 1). Since in this case the two valleys are not coupled by the boundary condition, we can consider only one of them, say the K valley, and use a simpler two-component spinor notation. Since this is a generic feature of all terminations except for the armchair one, this case can be considered as the most general. The armchair edge will be analyzed separately later on. As a basis to expand the Floquet space we use the corresponding QH zigzag edge states, which are given by (see Appendix A for details)

$$\boldsymbol{\chi}^s_{\nu_n k}(y)=\frac{1}{\sqrt{C_{\nu_n k}}}\begin{pmatrix} D_{\nu_n}(\xi)\\ s\sqrt{\nu_n}D_{\nu_n-1}(\xi)\end{pmatrix}, \tag{7}$$

where $\xi=\sqrt{2}(y/\ell_B-k\ell_B)$, $\ell_B=\sqrt{\hbar c/eB}$ is the magnetic length, $\varepsilon_n(k)=s\,\hbar\omega_c\sqrt{\nu_n(k)}$ is the energy of the Hall state, where $s=\pm1$ refers to the electron and hole bands, respectively, and $\omega_c=\sqrt{2}v_F/\ell_B$ is the cyclotron frequency, $D_\nu(x)$ is the Parabolic Cylinder function of index ν, k is the crystal momentum along the x axis and $C_{\nu k}$ is a normalization constant. Here, $n\geq1$ enumerates the positive energy levels, for a given k, in ascending order. Notice that the plane wave factor along the x axis (see Appendix A) can be safely ignored as the perturbation does not mix states with different k.

Since the laser field is monochromatic [cf. Eq. (2)], the Floquet matrix $\mathcal{H}_F$—which is a representation of Eq. (5) in the $\mathcal{R}\otimes\mathcal{T}$ space—is an infinite block tridiagonal matrix. With our choice of basis given in Eq. (6), the diagonal blocks are itself diagonal. Because we are interested in the effect of the laser field on a few edge states around the Dirac point ($\varepsilon=0$), we will truncate $\mathcal{H}_F$ and retain $2N$ Landau levels, N above and N below the Dirac point and $2M+1$ Floquet replicas. Hence $n=1,\ldots,N$ and $m=-M,\ldots,M$. The matrix element between states $|\chi^{\tilde{s}}_{\nu_{\tilde{n}}k}\rangle$ and $|\chi^s_{\nu_n k}\rangle$ in the m and $m-1$ Floquet replicas, respectively, is simply given by

$$\langle\chi^{\tilde{s}}_{\nu_{\tilde{n}}k}|\mathcal{V}|\chi^s_{\nu_n k}\rangle=\frac{\eta\hbar\Omega}{2}\left[sf(\alpha,\varphi)R_{\nu_n\nu_{\tilde{n}}}+\tilde{s}f(-\alpha,\varphi)R_{\nu_{\tilde{n}}\nu_n}\right], \tag{8}$$

where $f(\alpha,\varphi)=\cos\alpha-i\sin\alpha\,e^{-i\varphi}$ and

$$R_{\nu\nu'}=\frac{\ell_B\sqrt{\nu}}{\sqrt{2\,C_{\nu k}C_{\nu' k}}}\int_{-\sqrt{2}k\ell_B}^{\infty}\mathrm{d}\xi\,D_{\nu-1}(\xi)D_{\nu'}(\xi)\,. \tag{9}$$

Similar calculations can be performed in the K' valley using the eigenfunctions given by Eq. (A4).

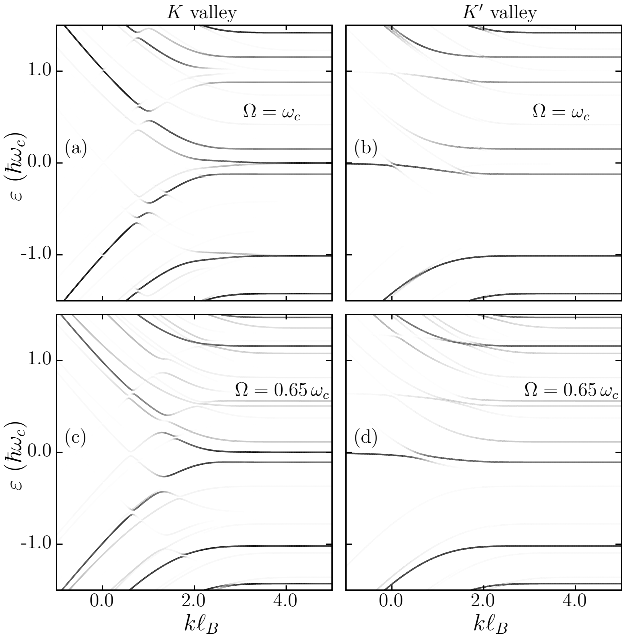


Figure 2. Quasienergy spectrum projected on the $m=0$ Floquet replica [$\rho_0(\varepsilon,k)$, solid black lines] as a function of the dimensionless wavevector $k\ell_B$ along a zigzag edge irradiated with a circularly polarized laser. For (a) and (b) [(c) and (d)] we use $\eta=0.2\,\omega_c/\Omega$ ($\eta=0.3\,\omega_c/\Omega$). Subplots (a) and (b) correspond to the resonant case ($\Omega=\omega_c$) for valleys K and K', respectively. Similarly, (c) an (d) correspond to a non resonant photon energy ($\Omega=0.65\omega_c$). Here five Floquet channels ($-2\leq m\leq2$) were used.

Let us now consider the case of a laser field with positive (counter clockwise) circular polarization: $\alpha=\pi/4$ and $\varphi=\pi/2$. Then we have $f(\pi/4,\pi/2)=0$ and $f(-\pi/4,\pi/2)=\sqrt{2}$, and thus the right hand side of Eq. (8) reduces to $\eta\hbar\Omega\,\tilde{s}R_{\nu_{\tilde{n}}\nu_n}/\sqrt{2}$. Figure 2 shows the quasienergy dispersion of the Floquet Hall edge states, weighted by their projection on the $m=0$ Floquet replica. These dispersion relations were obtained by numerically calculating the following spectral density

$$\rho_0(\varepsilon,k)=-\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}_0\left[\varepsilon+i0^+-\mathcal{H}_F(k)\right]^{-1}\;, \tag{10}$$

where the trace Tr_0 is taken only over the $m=0$ subspace. Here we used five Floquet replicas ($M=2$), twelve Landau levels ($N=6$) and two different photon energies: (i) resonant with the first bulk Landau level ($\Omega=\omega_c$) and (ii) off-resonant ($\Omega=0.65\,\omega_c$). We include also the results for the K' valley which were obtained in a similar fashion.

The main new features in the spectrum that are apparent from the Fig. 2 are: (i) the splitting of the bulk Landau levels and the lack of electron-hole symmetry, both analyzed in detail in Sec. II D; (ii) the appearance of multiple dynamical gaps (or, more precisely, avoided crossings) of different order in η. In particular, the first order ones at $\pm\hbar\Omega/2$ in the K valley arises from the resonant coupling between the lowest electron and the highest hole edge states and, as we will show when discussing the transport properties, lead to the (almost complete) suppression of the QH conductance; (iii) the bending of the otherwise flat zero energy state ($\nu = 0$) of the K' valley, which results from the direct coupling to the lowest Landau level of the electron band $|\boldsymbol{\chi}^{+}_{\nu_1 k}\rangle$ (in the Floquet picture it corresponds to coupling to the $m = -1$ replica). This leads to an edge mode with a polarization dependent dispersion that it is *always* a counter-propagating mode, in the sense that it has the opposite velocity that the edge states with the same sign of quasienergy. This, in turn, causes a change of the sign of the Hall conductance, as we discuss in Section V; (iv) in a small quasienergy region above $\varepsilon = \hbar\omega_c$ there is an effective reduction of the number of edge states as the one coming from the K' valley is shifted upwards. In a finite sample, the same happens for the K valley on the other edge. This leads to the emergence of a $4e^2/h$ feature in the two terminal conductance as discussed in Section IV.

The size $\Delta_{\rm dym}$ of a dynamical gap is given, to first order in η, by the matrix element between the two states involved in the avoided crossing [Eq. (8)]. In the case of the gap at $\varepsilon = \hbar\Omega/2$, they are $|\boldsymbol{\chi}^{+}_{\nu_1 k}\rangle$ and $|\boldsymbol{\chi}^{-}_{\nu_1 k}\rangle$ in the $m = 0$ and $m = 1$ replicas, respectively, so that

$$\Delta_{\rm dym} \simeq \eta\,\hbar\Omega\,\sin\alpha |R_{\nu_1\nu_1}|\,, \tag{11}$$

where $\nu_1(k) = (\Omega/2\omega_c)^2$ defines the value of k where the resonant condition is fulfilled. Eq. (11) is not restricted to a circularly polarize laser. In fact, it must be noted that in the case of a linearly polarized beam, $\Delta_{\rm dym}$ depends on the relative orientation of the electric field and the edge, being zero if the electric field is parallel to the edge (α=0).

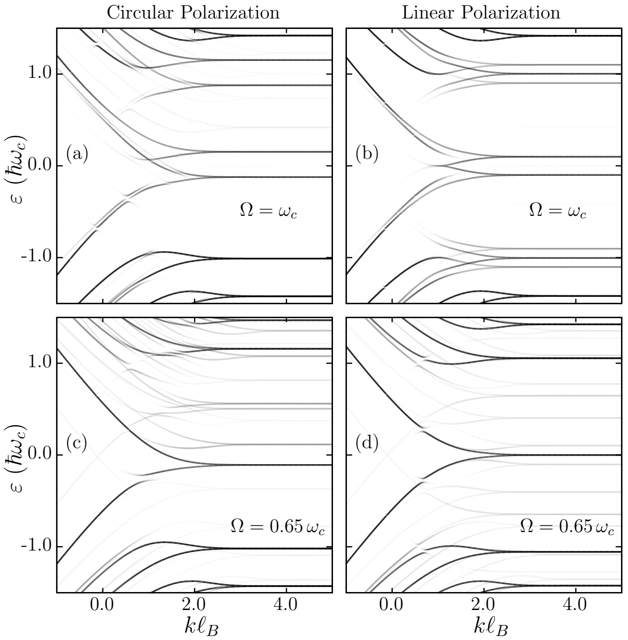


Figure 3. Quasienergy spectrum projected onto the replica $m = 0$ for an irradiated armchair edge. The laser is circularly (linearly) polarized in (a) and (c) [(b) and (d)]. In (a) and (b) the laser is in resonance ($\Omega = \omega_c$), with $\eta = 0.2\,\omega_c/\Omega$; whereas in (c) and (d) it is out of resonance ($\Omega = 0.65\omega_c$), with $\eta = 0.3\,\omega_c/\Omega$. Five Floquet replicas ($-2 \le m \le 2$) were included. It is clear, in comparison with Fig. 2, the absence of a first order dynamical gap at $\pm\hbar\Omega/2$. Higher order gaps due to the coupling with $m = \pm 2$ replicas are present.

C. Armchair Floquet Hall states

Now we apply a similar treatment to the case of armchair edges. For this special termination, however, one needs to take into account both valleys at the same time [45], as the boundary condition mixes them. The eigenfunctions for the static system are presented in Appendix A. They are now four-component spinors

$$\boldsymbol{\chi}^{s}_{\nu_n k}(x) = \frac{1}{\sqrt{C_{\nu_n k}}}\begin{bmatrix} -is\tau_n D_{\nu_n}(\xi)e^{iKx} \\ -\tau_n\sqrt{\nu_n}D_{\nu_n-1}(\xi)e^{iKx} \\ is\sqrt{\nu_n}D_{\nu_n-1}(\xi)e^{-iKx} \\ D_{\nu_n}(\xi)e^{-iKx}\end{bmatrix}, \tag{12}$$

with $\xi = \sqrt{2}(x/\ell_B - k\ell_B)$. The eigenenergies are given by $\varepsilon_n = s\,\hbar\omega_c\sqrt{\nu_n}$ with $s = \pm 1$, the meaning of s being the same as in the zigzag case. The new parameter $\tau_n = (-1)^{n+1}$ indicates the *branch* ν_n belongs to [see Appendix A for details as well as for the explicit form of $C_{\nu_n k}$]. Using Eq. (12) and the interaction matrix Eq. (4) we have the following matrix element that enters in $\mathcal{H}_F$

$$\langle\boldsymbol{\chi}^{\tilde{s}}_{\nu_{\tilde{n}} k}|\,\mathcal{V}\,|\boldsymbol{\chi}^{s}_{\nu_n k}\rangle = i\frac{\eta\hbar\Omega}{2}(\tilde{s}s + \tau_{\tilde{n}}\tau_n)\times \\ [sf(-\alpha,\phi)R_{\nu_{\tilde{n}}\nu_n} - \tilde{s}f(\alpha,\phi)R_{\nu_n\nu_{\tilde{n}}}]\,. \tag{13}$$

This gives an interesting selection rule: $\langle\boldsymbol{\chi}^{\tilde{s}}_{\nu_{\tilde{n}} k}|\,\mathcal{V}\,|\boldsymbol{\chi}^{s}_{\nu_n k}\rangle = 0$ if $\tilde{s}s = -\tau_{\tilde{n}}\tau_n$. In particular, for the $n = \tilde{n} = 1$ edge mode the coupling between the conduction (c, $s = 1$) and valence (v, $\tilde{s} = -1$) bands vanishes

$$\langle\boldsymbol{\chi}^{v}_{\nu_1 k}|\,\mathcal{V}\,|\boldsymbol{\chi}^{c}_{\nu_1 k}\rangle = 0\,. \tag{14}$$

Because this matrix element is responsible for the opening of a dynamical gap at $\pm\hbar\Omega/2$, we do not expect such a gap in the armchair case—of course, this argument refers to a first order gap; higher order (smaller) gaps in fact exist at these points. Note that, in general, there is no coupling between electron and hole levels (i.e. $s\tilde{s} = -1$) belonging to the same solution branch (i.e. $\tau_{\tilde{n}}\tau_n = 1$). Similarly, $\langle\boldsymbol{\chi}^{v}_{\nu_{\tilde{n}} k}|\mathcal{V}|\boldsymbol{\chi}^{v}_{\nu_n k}\rangle = \langle\boldsymbol{\chi}^{c}_{\nu_{\tilde{n}} k}|\mathcal{V}|\boldsymbol{\chi}^{c}_{\nu_n k}\rangle = 0$ if $\tau_n = -\tau_{\tilde{n}}$. i.e., same electron or hole character and different solution branch. All this implies that, considering only first order couplings, the armchair edges have more symmetries than the zigzag ones, leading to a simpler Floquet spectrum.

The corresponding weighted Floquet spectrum, calculated with the projected spectral density $\rho_0(\varepsilon, k)$, is shown in Fig. 3

for a circular ($\varphi = 2\alpha = \pi/2$) and linear ($\alpha = \varphi = 0$) polarization of the laser field. As in the previous section we use five Floquet replicas ($M = 2$). The most striking difference with Fig. 2 is the lack of first order dynamical gaps at $\pm\hbar\Omega/2$, in complete agreement with Eq. (14). Moreover, when the photon energy is out of resonance [Figs. 3(c) and 3(d)], the weighted Floquet bands are similar to the ones of the static system, except for an energy shift of certain Landau levels. It is interesting that certain gaps appear at the crossing of the static system ($m = 0$) with replicas $m = \pm 2$. Since our linear model only couples directly Floquet channels differing in one photon, these gaps are of second order and thus smaller than those seen in Fig. 2 at $\varepsilon = \pm\hbar\Omega/2$. The flat states for $k\ell_B \gg 2$ correspond to the bulk Landau levels, so that their shifting and splitting follows the pattern of the latter, which we now discuss.

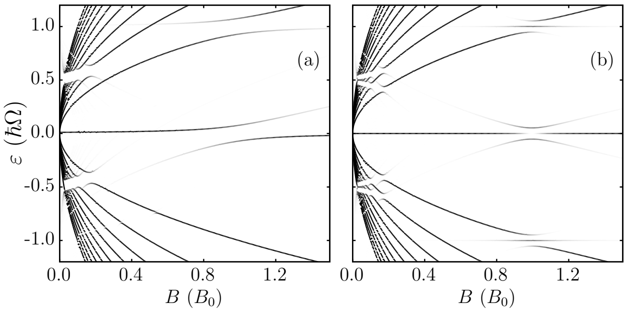


Figure 4. Weighted bulk Landau-Floquet spectral density $\rho_0(\varepsilon, B)$ as a function of the magnetic field for a fixed value of Ω. In (a) and (b) the laser field is circularly ($\alpha = \pi/4$, $\phi = \pi/2$) and linearly ($\alpha = \pi/2$, $\phi = 0$) polarized, respectively. Here $\eta = 0.15$ and $B_0 = \Omega^2\hbar c/(2ev_F^2)$.

D. Bulk selection rules

To better understand some of the features observed in Figs. 2 and 3, it is useful to analyze the bulk case. For that we calculate $\rho_0(\varepsilon)$ [cf. Eq. (10)] with $\mathcal{H}_F$ written in the basis of the bulk eigenfunctions (see Appendix A). The matrix elements of $\mathcal{H}_F$ between the Floquet bulk eigenfunctions can be calculated by considering each valley separately as they are decoupled in bulk. Using the solutions $|\boldsymbol{\chi}_{lk}\rangle$ given in Eq. (A2) for the K valley (and omitting the superscript K) we obtain the following matrix elements

$$\langle\boldsymbol{\chi}_{nk}|\mathcal{V}|\boldsymbol{\chi}_{lk}\rangle = \frac{\eta\hbar\Omega}{4\sqrt{1-(\delta_{n0}+\delta_{l0})/2}}\times \left[f(\alpha,\varphi)\mathrm{sgn}(l)\delta_{|n|,|l|-1} + f(-\alpha,\varphi)\mathrm{sgn}(n)\delta_{|n|,|l|+1}\right]. \quad (15)$$

In the right-handed circularly polarized case, $f(\alpha,\varphi) = 0$ and this matrix element is proportional to $\delta_{|n|,|l|+1}$. In this way we obtain the selection rule $|l| = |n| - 1$, where l is the Landau index of the state with an extra absorbed photon [$(l, m+1) \leftrightarrow (n, m)$ transition in the Floquet space]. For the opposite circular polarization the $|l| = |n| + 1$ rule applies. When the polarization is linear, Eq. (15) dictates that $||n| - |l|| = 1$. The matrix elements between eigenfunctions in the K' valley are the same as those in Eq. (15), and thus the same selections rules apply.

These selection rules can be clearly seen when we plot the Floquet spectral density $\rho_0(\varepsilon)$ as a function of B for a fixed value of $\hbar\Omega$, as shown in Fig. 4. The calculations were carried out with twenty Landau levels ($N = 10$,), five Floquet replicas ($M = 2$) and $\eta = 0.15$. To properly scale the spectrum it is useful to define an auxiliary magnetic field $B_0 = \hbar c\Omega^2/(2ev_F^2)$ so that $\omega_c/\Omega = \sqrt{B/B_0}$ and the energy of the m-th Floquet replica is simply $\varepsilon_n^{(m)} = \left(\mathrm{sgn}(n)\sqrt{\frac{B}{B_0}|n|} + m\right)\hbar\Omega$. Figure 4(a) shows the case of a circular polarization. First we notice that, as B approaches zero, there are dynamical gaps of almost constant size that tend to center around $\varepsilon = \pm\hbar\Omega/2$. These are reminiscences of the well-known dynamical gaps of irradiated graphene that appear in the absence of any magnetic field (with size roughly equal to $\eta\hbar\Omega$) [26, 51]. Another interesting feature is the appearance of anti-crossings near $B = B_0$, that is, when the laser field is in resonance with the transition between the zero and the first Landau level. In the case of the anti-crossing near $\varepsilon = 0$, it originates from the degeneracy of the Floquet states $|\boldsymbol{\chi}_{0k}, 0\rangle$ and $|\boldsymbol{\chi}_{1k}, -1\rangle$, as dictated by the selection rules—here we use the notation $|\boldsymbol{\chi}_{\nu k}, m\rangle$ to indicate the $m-$th Floquet replica the state $|\boldsymbol{\chi}_{\nu k}\rangle$ belongs to—while for the one at $\varepsilon = \hbar\Omega$ it corresponds to the near degeneracy between $|\boldsymbol{\chi}_{1k}, 0\rangle$ and $|\boldsymbol{\chi}_{0k}, 1\rangle$. Note that because of the selection rules there is no coupling between $|\boldsymbol{\chi}_{-1k}, 0\rangle$ and $|\boldsymbol{\chi}_{0k}, -1\rangle$ and so $\rho_0(\varepsilon) \neq \rho_0(-\varepsilon)$ (there is no electron-hole symmetry). This is consistent with the results of Refs. [30, 31] where the full Hofstadter butterfly spectrum (tight-binding model) was analyzed. Finally, we mention that the series of gaps near $\hbar\Omega/2$ arises from the anti-crossings between the Floquet states $|\boldsymbol{\chi}_{nk}, 0\rangle$ and $|\boldsymbol{\chi}_{1-n,k}, 1\rangle$ with $n = 1, 2, \cdots$, while those near $-\hbar\Omega/2$ appear at the crossings of $|\boldsymbol{\chi}_{-nk}, 0\rangle$ and $|\boldsymbol{\chi}_{n+1,k}, -1\rangle$.

As we have already mentioned, for a linearly polarized laser the selection rules require $||n| - |l|| = 1$, with both n and l entering symmetrically. This implies that $\rho_0(\varepsilon) = \rho_0(-\varepsilon)$ as it is clear from Fig. 4(b). It is interesting to analyse in particular the triple crossing that occurs near $\varepsilon = 0$ for $B = B_0$ (resonance condition, $\Omega = \omega_c$). In that case, $|\boldsymbol{\chi}_{0k}, 0\rangle$, $|\boldsymbol{\chi}_{-1k}, 1\rangle$ and $|\boldsymbol{\chi}_{1k}, -1\rangle$ become degenerate while the selection rules allow the coupling between $|\boldsymbol{\chi}_{0k}, 0\rangle$ and each of the other two states, with a matrix element $\eta\hbar\Omega\, e^{\pm i\alpha}/2\sqrt{2}$, respectively. Within this restricted subspace, a straightforward diagonalization gives the eigenvalues $\lambda_0 = 0$ and $\lambda_\pm = \pm\eta\hbar\Omega/2$. The eigenvector corresponding to λ_0 is $(|\boldsymbol{\chi}_{-1k}, 1\rangle + |\boldsymbol{\chi}_{1k}, -1\rangle)/\sqrt{2}$ which does not have any weight on the $m = 0$ replica, as it is evident from the lack of spectral weight shown in Fig. 4(b). For the other eigenvalues $\lambda_\pm$ we have that the corresponding eigenvectors are $(\sqrt{2}|\boldsymbol{\chi}_{0k}, 0\rangle \mp e^{i\alpha}(|\boldsymbol{\chi}_{-1k}, 1\rangle + |\boldsymbol{\chi}_{1k}, -1\rangle)/2$, both with the same weight (1/2) on the $m = 0$ replica. A similar calculation explains the features observed at $\varepsilon = \pm\hbar\Omega$ in Fig. 4(b). Finally, the small gaps near $\hbar\Omega/2$ that appear at

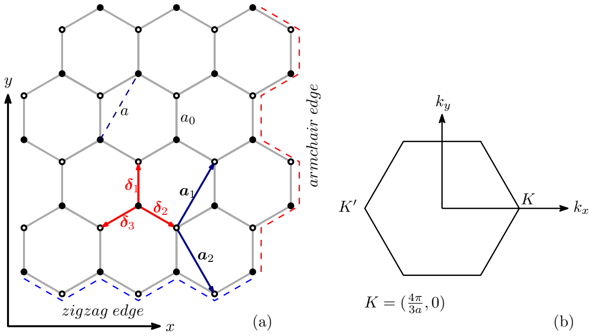


Figure 5. (Color online) (a) Geometry of the problem in the tight-binding model, with the lattice $\boldsymbol{a}_j$ and nearest neighbors $\boldsymbol{\delta}_j$ vectors indicated, whose modules are a and a_0, respectively ($a = \sqrt{3}\,a_0$). We refer all our calculations to this configuration, and thus a zigzag (armchair) ribbon has a translational symmetry along the x (y) direction. (b) First Brillouin Zone and the two non equivalent valleys K and K'.

low B result from the crossing of $|\boldsymbol{\chi}_{nk}, 0\rangle$ with $|\boldsymbol{\chi}_{-(n+1)k}, 1\rangle$ and $|\boldsymbol{\chi}_{-(n-1)k}, 1\rangle$ ($n = 1, 2, \cdots$), and similarly for the gaps near $-\hbar\Omega/2$. It is worth mentioning here that in the absence of a magnetic field there is no laser induced dynamical gap in $\rho_0(\varepsilon)$ for a linearly polarized laser but a pseudo-gap that closes linearly in energy at exactly $\varepsilon = \hbar\Omega/2$ [26]. This can be seen in Fig. 4(b) as B goes to zero, where the size of such gaps become smaller until they vanish at $B = 0$. In this case, the presence of this pseudo-gap is reveled by the Landau-Floquet states that appear nearly pinned at $\varepsilon = \hbar\Omega/2$.

III. TIGHT BINDING MODEL

The Dirac model is suitable only for describing the low energy excitations near the Dirac point ($\varepsilon = 0$), where the energy dispersion is almost conical. A better and more complete description is given by a tight-binding (TB) model, where the p_z carbon orbitals in graphene are described by $\hat{\mathcal{H}} = \sum_{\langle i,j\rangle} t_{ij}\,\hat{c}_i^\dagger \hat{c}_j + h.c.$. Here $\hat{c}_j$ is a destruction operator at the position j, the notation $\langle i, j\rangle$ implies that the summation is carried over nearest neighbors only, separated by a distance $a_0 = 1.42$Å (see Fig. 5), while the hopping t_{ij} is independent of the site: $t_{ij} = t = -2.8\,$eV. The effect of an external field described by the vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$ is included as before via the Peierls substitution, which in the TB approach is given by

$$t_{ij} \to t_{ij} \times \exp\left[\left(\frac{ie}{\hbar c}\right)\int_{\boldsymbol{r}_i}^{\boldsymbol{r}_j} \boldsymbol{A}(\boldsymbol{r}, t)\cdot d\boldsymbol{r}\right]. \qquad (16)$$

A. Landau levels

Let us first briefly describe the well-known effects of an homogeneous magnetic field on a graphene ribbon with either zigzag or armchair edges. Figure 5 shows that with our choice of axes the ribbon has translation symmetry along the x (y) direction for a zigzag (armchair) edge. This symmetry allows us to introduce a Bloch function with a crystal momentum k_x or k_y along the relevant symmetry direction. In order to preserve this symmetry we choose the gauge $\boldsymbol{A}(x) = -By\,\hat{\boldsymbol{x}}$ $[\boldsymbol{A}(x) = Bx\,\hat{\boldsymbol{y}}]$ for the zigzag (armchair) ribbon. Hence, the Peierls substitution for the zigzag ribbon is explicitly given by $t_{ij} = t\,\exp[-i\,\zeta(x_j - x_i)(y_j + y_i)/a_0^2]$, where $\zeta = \pi\Phi/\Phi_0$, $\Phi = Ba_0^2$ and $\Phi_0 = hc/e$ is the flux quantum.

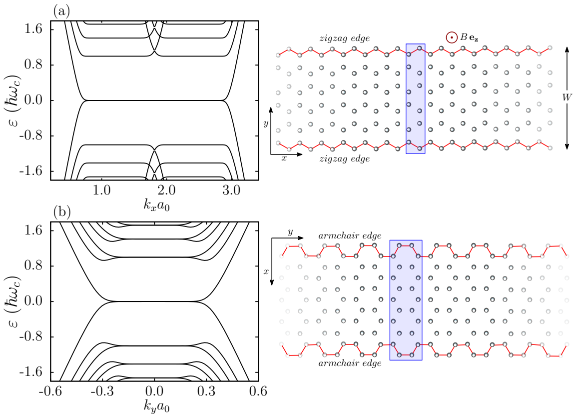


Figure 6. (Color online) Left panels: Energy dispersion of the Landau levels for (a) a zigzag and (b) an armchair ribbon with widths $W = 224\,a_0$ and $W = 75\sqrt{3}\,a_0$, respectively, and $\zeta = 0.003$ (equivalent to $\hbar\omega_c \simeq 0.46\,$eV and $\ell_B \simeq 13\,a_0$). Right panels: Geometry of the ribbons with the corresponding unit cells highlighted by the purple box. The magnetic field $B\hat{\boldsymbol{z}}$ is applied normally to the ribbon's plane.

Figure 6(a) shows the Landau bands for a zigzag ribbon of a width of 300 atoms ($W = 224a_0$) and $\zeta = 0.003$. There are bulk Landau levels (flat bands), as well as dispersive edge states due to the confinement imposed by the ribbon. The dispersion relations in each Dirac point is in very good agreement with those found with the Dirac model [compare with Fig. 19(a) and 19(b) in Appendix A]. It is worth mentioning here that not all the zero energy states in this geometry are bulk Landau levels. There is also a trivial dispersionless edge mode that appear on zigzag ribbons in the absence of a magnetic field (see Appendix A for a further discussion). These modes, being dispersionless, are not affected by the Lorentz force.

A similar calculation done for the armchair case leads to the Landau spectrum shown in Fig. 6(b). Here the ribbon is 302 atoms wide ($W = 75\sqrt{3}\,a_0$). In contrast with a zigzag ribbon, the two dispersionless states at $\varepsilon = 0$ are fully located in bulk, and are identified with the $n = 0$ Landau level.

B. Floquet states

We now add a time dependent laser field using the vector potential given in Eq. (2), which is assumed to be homogeneous throughout space. The integral in Eq. (16)

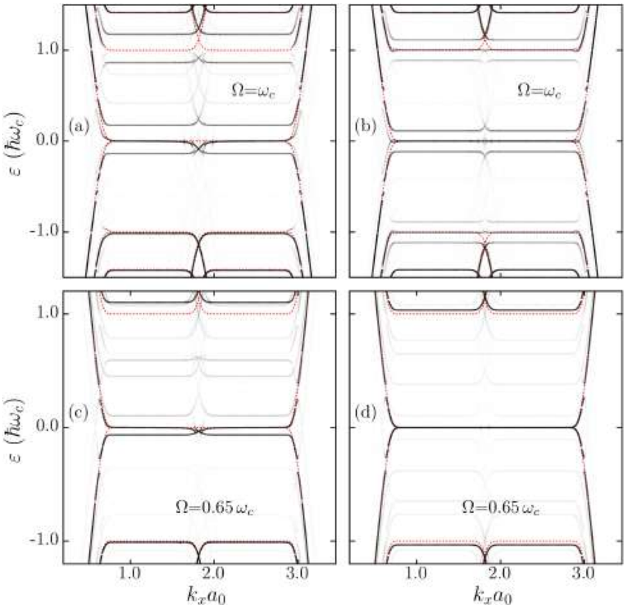


Figure 7. (Color online) Landau-Floquet spectral density $\rho_0(\varepsilon, k)$ in the presence of an homogeneous magnetic field ($\zeta = 0.003$) and a laser with intensity $z = 0.025$. The polarization is circular ($\varphi = 2\alpha = \pi/2$) for (a) and (c), and linear ($\alpha = \pi/2,\ \varphi = 0$) for (b) and (d). Sub-plots (a) and (b) correspond to the resonant case with $\Omega = \omega_c$, whereas in (c) and (d) the laser field is off-resonant with $\Omega = 0.65\,\omega_c$. Five Floquet replicas ($-2 \leq m \leq 2$) were used . The ribbon is 300 atoms wide ($W = 224\,a_0$).

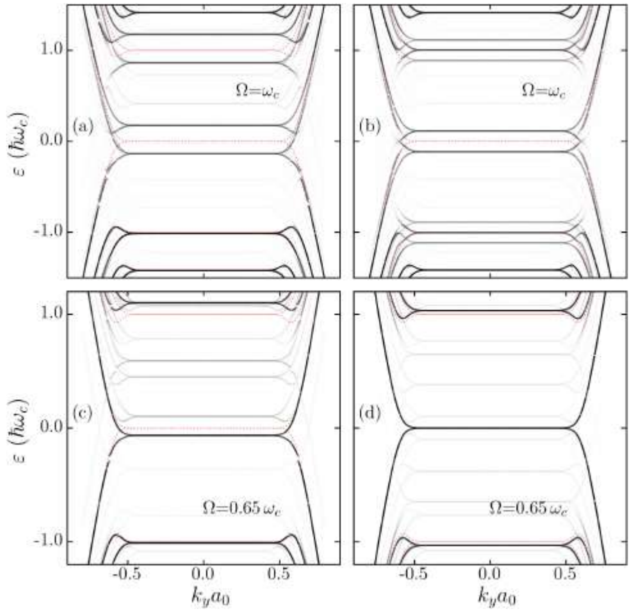


Figure 8. (Color online) Same as previous figure but for an armchair ribbon of width $W = 120\sqrt{3}\,a_0$ (480 atoms in the unit cell).

is then simply $\boldsymbol{R}_{ij} \cdot \boldsymbol{A}(t)$, where $\boldsymbol{R}_{ij} = a_0\,(\cos\theta_{ij}\,\hat{\boldsymbol{x}} + \sin\theta_{ij}\,\hat{\boldsymbol{y}})$ is the vector connecting neighboring sites i and j. For a laser field with positive circular polarization, the Peierls substitution leads to $t_{ij} = t\exp[iz\cos(\Omega t - \theta_{ij})]$, while for the linearly polarized case we have $t_{ij} = t\,\exp[iz\cos(\theta_{ij} - \alpha)\cos\Omega t]$. Here we have introduced the dimensionless quantity $z = ea_0A_0/\hbar c = (\Omega a_0/v_F)\,\eta$ that measures the intensity of the perturbation. It is clear then that the time-dependent TB Hamiltonian for the irradiated ribbon is periodic in time but not harmonic. In the Floquet formulation, the Floquet matrix elements now couple replicas with $\Delta m \neq \pm 1$. The Fourier components of the hamiltonian can be calculated using the well-known Jacobi-Anger identity: $e^{ir\cos\theta} = \sum_{m=-\infty}^{+\infty} i^m J_m(r)\,e^{im\theta}$, where $J_m(r)$ are the Bessel functions of the first kind of integer order. $\mathcal{H}_F$, which is no longer block tridiagonal, is truncated to a finite number of Floquet channels for its numerical diagonalization. In the following we retain five Floquet replicas $-2 \leq m \leq 2$ (unless otherwise stated) and calculate $\rho_0(\varepsilon, k)$ by means of Eq. (10). This number of replicas guarantees that, for the value of the parameters we use, the most relevant features in $\rho_0(\varepsilon, k)$ are well described.

The Landau-Floquet spectral density for a zigzag ribbon of width $W = 224\,a_0$ is shown in Fig. 7: (a) and (c) correspond to the circularly polarized case, while (b) and (d) to a linearly polarized laser, with $\alpha = \pi/2$—the direction of polarization is perpendicular to the edges of the ribbon—which guarantees a maximum size in the gap opening, see Eq. (11). The photon frequency is $\Omega = \omega_c$ (resonant) for (a) and (b) and $\Omega = 0.65\,\omega_c$ (off-resonant) for (c) and (d). We use a dimensionless flux $\zeta = 0.003$ and $z = 0.025$. The bands of the static system (red dashed lines) are also shown for comparison.

All features described in the previous section using the Dirac approximation are observed here. In particular, there are dynamical gaps at around $\varepsilon = \pm\hbar\Omega/2$ with a magnitude in agreement with Eq. (11). In Fig. 7(c) the degeneracy between the two flat modes at $\varepsilon = 0$ is removed, one remains at $\varepsilon = 0$ while the other shifts downwards. We identify the latter with the $n = 0$ bulk Landau level, which obeys the selection rules Eq. (15) and whose shifting is in good agreement with these. The other state corresponds to the edge state solution Eq. (A9) given in Appendix A. Being an edge state, it does not couple to the bulk states and hence it is pinned at $\varepsilon = 0$. A similar analysis applies to Fig. 7(a) except that the bulk $n = 0$ Landau level here is split instead of shifted due to the resonance condition.

The linearly polarized case present similar features. However, in this case, the spectrum is electron-hole symmetric and hence the $n = 0$ Landau level can only split. For an off-resonant photon energy [Fig. 7(d)] the only effects of the laser field are the opening of the dynamical gaps at $\pm\hbar\Omega/2$ and the shifting of the first non-zero static Landau levels. In resonance, Fig. 7(b), there is a splitting in a neighborhood of $\varepsilon = 0$. The states with $\varepsilon \neq 0$ are truly bulk states, coming from the mixing of the $n = 0$ Landau level in the replica $m = 0$, and the $n = \pm 1$ Landau level from the replica $m = \mp 1$. The state that remains at $\varepsilon = 0$ is the zigzag edge state mentioned above.

Figure 8 shows the results for an armchair ribbon. The main feature predicted by the Dirac model is quite apparent: irrespective of the laser polarization, there is not a (first order) gap at $\pm\hbar\Omega/2$, the origin of such absence being the selection rule between Landau edge states, Eq. (14). Moreover, as it was mentioned in the preceding section, the bands around $k_y = 0$ (roughly $|k_y a_0| \leq 0.4$ in Fig. 8), are basically bulk bands, and as such their splitting follows the selection rules obtained in Sec. II D.

IV. TWO-TERMINAL CONDUCTANCE

We now discuss the transport properties of an illuminated ribbon in the QH regime in a two-terminal setup. For that, we take the magnetic field to be present throughout the entire sample (including the semi-infinite leads), whereas the laser field is switched on smoothly (over a length scale λ_1), kept constant for a distance $2\lambda_2$ and finally switched off, as schematically shown in Fig. 9. This defines the scattering region. If we take the coordinate x to be directed along the ribbon, then the scattering region is defined by $|x| \leq \lambda_1 + \lambda_2$, while the local laser field intensity parameter $z(x) = ea_0 A_0(x)/\hbar c$ is taken to be

$$z(x) = \begin{cases} z\ , & |x| \leq \lambda_2 \\ \frac{z}{2}\left[1+\cos\left(\frac{\pi(|x|-\lambda_2)}{\lambda_1}\right)\right]\ , & \lambda_2 \leq |x| \leq \lambda_1+\lambda_2\ . \end{cases} \tag{17}$$

Here z is the maximum value reached by the laser intensity. We use this symmetric profile for the laser field to preserve the left/right symmetry of the ribbon—the intensity is homogeneous along the transverse direction.

The current is computed within a scattering approach [42, 43]. In the non-interacting limit this is equivalent to the Keldysh formalism [49, 52]. This has been used for a variety of systems including laser illuminated graphene [41, 53]. The time-average current, $\bar{I} = \frac{1}{T}\int_0^T dt\, I(t)$, is calculated according to

$$\bar{I} = \frac{2e}{h}\sum_n \int \left[T_{RL}^{(n)}\, f_L(\varepsilon) - T_{LR}^{(n)}\, f_R(\varepsilon)\right] d\varepsilon\,, \tag{18}$$

where $T_{RL}^{(n)}(\varepsilon)$ is the transmission probability for an electron with energy ε from lead L to lead R emitting (absorbing) $n > 0$ ($n < 0$) photons and $f_\alpha(\varepsilon)$ is the Fermi function of the lead α. Defining the quantities $T(\varepsilon) = \sum_n (T_{LR}^{(n)}(\varepsilon) + T_{RL}^{(n)}(\varepsilon))/2$ and $\delta T(\varepsilon) = \sum_n (T_{LR}^{(n)}(\varepsilon) - T_{RL}^{(n)}(\varepsilon))/2$, the average current $\bar{I}$ can be written as the sum of two terms

$$\bar{I} = \frac{2e}{h}\int [T(\varepsilon)(f_L(\varepsilon) - f_R(\varepsilon)) - \delta T(\varepsilon)(f_L(\varepsilon) + f_R(\varepsilon))]\ d\varepsilon\,. \tag{19}$$

At zero temperature, and up to first order in the bias difference δV, it reduces to

$$\bar{I} = \frac{2e^2}{h} T(\varepsilon_F)\,\delta V - \frac{4e}{h}\int_{-\infty}^{\varepsilon_F} \delta T(\varepsilon)\, d\varepsilon\,. \tag{20}$$

Here ε_F is the Fermi energy. The bias independent contribution in Eq. (20) is the so-called *pumped* current. The inversion symmetry of our geometry guarantees that $\delta T(\varepsilon) = 0$. We can then define the linear dc two-terminal conductance $G_{2T}(\varepsilon_F) = \bar{I}/\delta V = (2e^2/h)\, T(\varepsilon_F)$ in terms of the transmittance at the Fermi energy. The latter is calculated using the Green function recursion technique within the Floquet formalism.

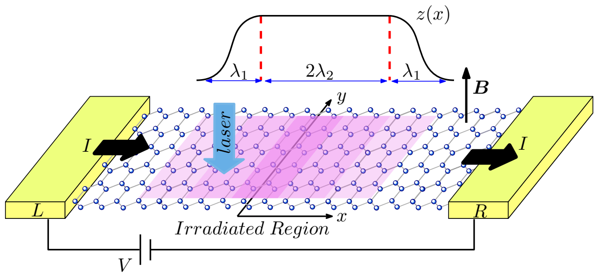


Figure 9. (Color online) Setup for the calculation of the two-terminal conductance G_{2T} using the Landauer-Büttiker approach for Floquet systems. The laser field, defined by the local function $z(x)$, is applied along a central region and its intensity vanishes smoothly towards the leads.

A. Zigzag ribbons

We consider in this section a 300 atoms wide zigzag ribbon ($W = 224\,a_0$) and take $\lambda_1 = 2\lambda_2 = 800 a_0\sqrt{3}$. This value is large enough as to minimize the backscattering of electrons at the interface where the laser is on. The other parameters are $\zeta = 0.003$ ($\hbar\omega_c \simeq 0.46\,\mathrm{eV}$ and $\ell_B \simeq 13\,a_0$) and $z = 0.025$. Unless otherwise mentioned, in all transport calculation we use only three Floquet replicas ($-1 \leq m \leq 1$), with the aim of describing the most important features while reducing the computational cost. First let us analyze the case of a linearly polarized laser ($\alpha = \pi/2$, $\varphi = 0$). The results are shown in Fig. 10 for the off-resonant (top panels) and resonant (bottom panels) situations. For each case, the conductance G_{2T} [(b) and (d)] is shown by the side of the corresponding Landau-Floquet spectral density (projected on the $m = 0$ replica) [(a) and (c)]. Here, electron-hole symmetry guarantees that $G_{2T}(-\varepsilon) = G_{2T}(\varepsilon)$.

For an off-resonant photon frequency [$\Omega = 0.65\,\omega_c$, Figs. 10(a) and 10(b)], two changes appear in the conductance (as compared with that of a non illuminated ribbon shown in red dashed lines): the two dynamical gaps centered at $\pm\hbar\Omega/2$, and a small energy shift in the $n = 1$ (electron and hole) Landau levels where a transition from $G_{2T} = 2e^2/h$ to $G_{2T} = 6e^2/h$ takes place. In the quasienergy region corresponding to the dynamical gaps the conductance is almost completely suppressed in a very sharp way. This is due to the fact that transport is carried out entirely by edge states, which are completely reflected by the laser spot in that particular energy range owed to the appearance of a gap—a related effect was discussed in the case of a driven transition-metal dichalcogenide ribbon in Ref. [54].

When the photon energy is in resonance with the first non zero Landau level [$\Omega = \omega_c$, Figs. 10(c) and 10(d)], the conductance exhibits, in addition to the two dips at $\pm\hbar\Omega/2$,

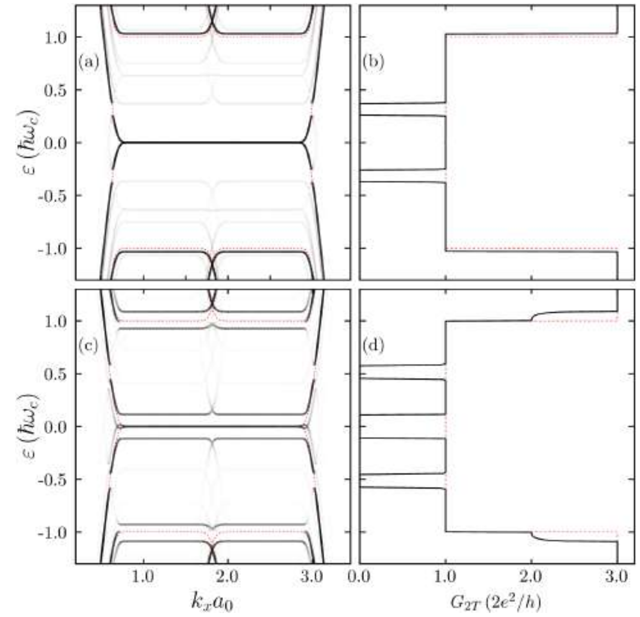


Figure 10. (Color online) Landau-Floquet bands [(a) and (c)] and linear conductance $G_{2T}(\varepsilon)$ [(b) and (d)] of a zigzag ribbon of width W=224 a_0 and ζ=0.003. The laser is linearly polarized ($\alpha = \pi/2$, $\phi = 0$) and we have used z=0.025. Sub-plots (a) and (b) correspond to Ω=0.65 ω_c, whereas in (c) and (d) we have Ω=ω_c (resonance). The results of the static system are also included in red dashed lines for comparison.

a strong suppression coming from the low energy gap created in the Floquet spectrum—it is interesting to note that the faint dispersive states near $\varepsilon = 0$ have a negligible contribution to the conductance. Additionally, whereas in a non irradiated sample the conductance jumps from $2e^2/h$ to $6e^2/h$ when the Fermi energy crosses $\hbar\omega_c$ as the result of the change of the number of available edge states with a given chirality from one to three, here an intermediate quasi plateau at $G_{2T} \approx 4e^2/h$ appears. As mentioned in the discussion of Fig. 2, this effect is related to the fact that there is a range of quasienergies were the number of effective edge modes is reduced by virtue of the upward energy shift of the Floquet edge mode of one of the valleys (the K' valley in the case of Fig. 2). The origin of that shift is the level repulsion between the flat (dispersionless) edge state in the $m = 1$ replica and the first dispersive (ν_2) edge mode of the $m = 0$ replica, which has always a higher quasienergy (this is not the case in the other valley). As such, this happens for opposite valleys in opposite sides of the sample. This explains why in Fig. 10(c) there are no edge modes for $\varepsilon_F \sim \hbar\omega_c$ near $k_x a_0 \sim 1.8$ for the chosen ribbon's width. Due to the electron-hole symmetry, a similar argument holds for $\varepsilon_F = -\hbar\omega_c$.

The results for a circularly polarized laser ($\varphi = 2\alpha = \pi/2$) are shown in Fig. 11 for the same photon frequencies. All the prominent features described for the linear case are also observed here, but with the addition of several important new ones:

(i) In contrast to Figs. 10, here $G_{2T}(-\varepsilon) \neq G_{2T}(\varepsilon)$;

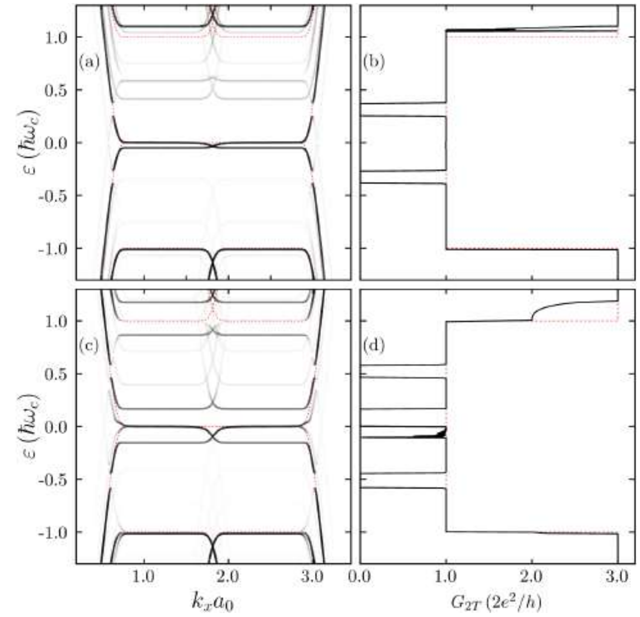


Figure 11. (Color online) Same as Fig. 10 but with a right-handed circularly polarized laser ($\varphi = 2\alpha = \pi/2$). The other parameters remain the same.

(ii) When $\Omega = 0.65\,\omega_c$ the conductance is, apart from the two square dips around $\pm\,\hbar\Omega/2$, quite similar to that of the static system. Notice that bending of the states near $\varepsilon_F = 0$ induced by the laser provides the channel that leads to $G_{2T} = 2e^2/h$ near the Dirac point;

(iii) When $\Omega = \omega_c$, the emergence of the new dispersive edge modes near $\varepsilon_F \sim 0$ (on the negative side for our choice of polarization) is fully developed. These edge modes are the ones described in Fig. 2 when discussing the properties of the edge modes of the K' valley [and the same as those described above in (ii)]. They lead to a quantized conductance that partially fills the gap near zero energy—the narrow dip in the conductance that is observed in that energy region is due to a high-order anticrossing, which cannot be fully appreciated in the Landau-Floquet spectral density [Fig. 11(c)]. The quasi plateau in Fig. 11(d) slightly above $\varepsilon_F = \hbar\omega_c$ has the same origin as those seen in Fig. 10(d), and can be explained in an analogous manner. The lack of a similar feature at $\varepsilon_F = -\hbar\omega_c$ comes from the selection rules Eq. (15).

B. Armchair ribbons: the role of *adiabaticity*

We now consider an armchair ribbon of width $W = 120\sqrt{3}\,a_0$ and keep the same parameters, $\zeta = 0.003$ and $z = 0.025$. Figure 12 shows the results for a linearly polarized laser ($\alpha = \varphi = 0$, in-plane electric field normal to ribbon's edges). Let us point out first some general considerations. As always with this type of polarization, the

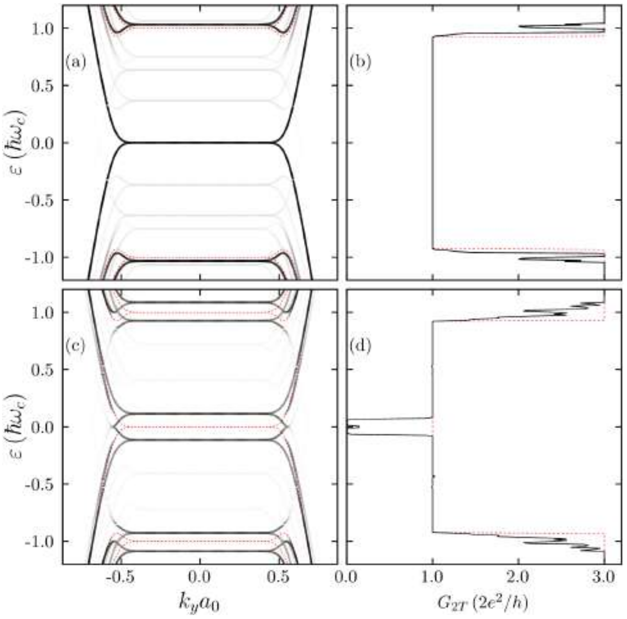


Figure 12. (Color online) Landau-Floquet spectral density and two-terminal conductance for an armchair ribbon ($W = 120\sqrt{3}\,a_0$) under illumination with a linearly polarized ($\alpha = \phi = 0$) laser. The parameters used are $\zeta = 0.003$ and $z = 0.025$. The photon frequency is $\Omega = 0.65\,\omega_c$ [(a) and (b)] and $\Omega = \omega_c$ [(c) and (d)].

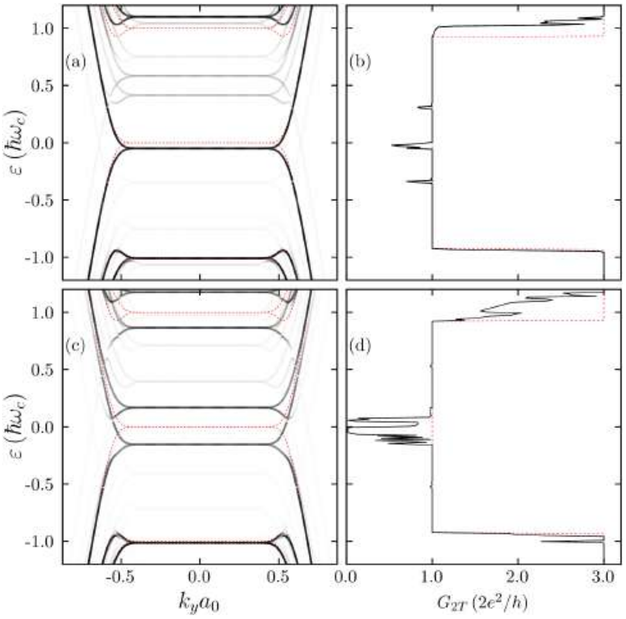


Figure 13. (Color online) Same as Fig. 12 but for a circularly polarized laser field ($\varphi = 2\alpha = \pi/2$).

Landau-Floquet spectrum presents electron-hole symmetry and thus $G_{2T}(\varepsilon_F) = G_{2T}(-\varepsilon_F)$. Moreover, as the Landau-Floquet bands lack the dynamical gaps at $\pm\hbar\Omega/2$ (to first order in z), G_{2T} does not show the typical strong suppression around these points—at most some minor very narrow features can be observed, corresponding to higher order photon processes. Additionally, for ε_F slightly below (above) $\hbar\omega_c$ ($-\hbar\omega_c$), where the static conductance changes from $G_{2T} = 2e^2/h$ to $G_{2T} = 6e^2/h$, the conductance is rather oscillating, a behavior reminiscent of the $4e^2/h$ feature found in the zigzag case—note that, in the latter case, the features appear exactly at $\varepsilon_F = \pm\,\hbar\omega_c$.

For a non-resonant photon frequency $\Omega = 0.65\,\omega_c$, Figs. 12(a) and (b), there are not special features around $\varepsilon_F = 0$. The linear conductance shows a flat profile as a function of ε_F, nearly the same as the static system, in agreement with the small changes induced by the laser field on the spectral density.

The situation changes when the laser photon frequency is in resonance, $\Omega = \omega_c$, Figs. 12(c) and (d). On the one side, there is a sharp dip around $\varepsilon_F = 0$. Its limits are roughly defined by the small avoided crossings, coming from the coupling between the replicas $|\boldsymbol{\chi}^c_{\nu_1 k}, 0\rangle$ and $|\boldsymbol{\chi}^v_{\nu_2 k}, 1\rangle$ ($\varepsilon_F > 0$), and between $|\boldsymbol{\chi}^v_{\nu_1 k}, 0\rangle$ and $|\boldsymbol{\chi}^c_{\nu_2 k}, -1\rangle$ ($\varepsilon_F < 0$)—selection rules Eq. (13) state that the matrix element between these pairs of Floquet states is zero, so these gaps originate from higher order processes, which explains their smallness. This dip is not the product of an evanescent penetration inside the scattering region, since from Fig. 12(c) it is clear that there are conducting states there. However, inside the region defined by the avoided crossings mentioned above, the Landau-Floquet states well inside the scattering region, and in a given edge of the ribbon, have the opposite sign of the velocity as compared with the incoming electrons. Therefore, the only way for these electrons to go through the central illuminated region is to move across the width of the ribbon until reach the opposite edge, where available states with a favorable velocity exist. If the laser's spatial profile (see Fig. 9), is sufficiently smooth or *adiabatic*, as it is in our calculations, this motion of charge between edges is hindered by the presence of the small gaps introduced above. In this scenario, the fraction of incoming electrons reaching the opposite edge is negligible, and instead most of them simply backscatter into the Floquet channels $m = 1$ or $m = -1$, depending on the character (conduction or valence) of the incident electrons. For electrons with energies just above the avoided crossing, the mismatch is still present, although in this case the incident electrons can reach the other edge and transmit into the other lead. This form of transmission is inherently inefficient, although it can be improved by further smoothing out the turning on-off of the laser (not shown). The renormalized (shifted) value of bulk part of $|\boldsymbol{\chi}^c_{\nu_2 k}, -1\rangle$ marks the onset of a constant conductance $G_{2T} = 2e^2/h$. Above this value there are two available Landau-Floquet channels, although with only one incident channel the transmittance reduces to $T(\varepsilon_F) = 1$. Due to electron-hole symmetry, this analysis can be extended to negative energies.

On the other hand, near $\varepsilon_F = \hbar\omega_c$ (similar at $\varepsilon_F = -\hbar\omega_c$), instead of the well defined change from $T(\varepsilon_F) = 1$ to $T(\varepsilon_F) = 3$ in the non driven system, there is a progressive increase from 1 to 3, something that resembles the zigzag case (where an intermediate step with $T(\varepsilon_F) \simeq 2$ was found).

The circularly polarized case (Fig. 13, $\varphi = 2\alpha =$

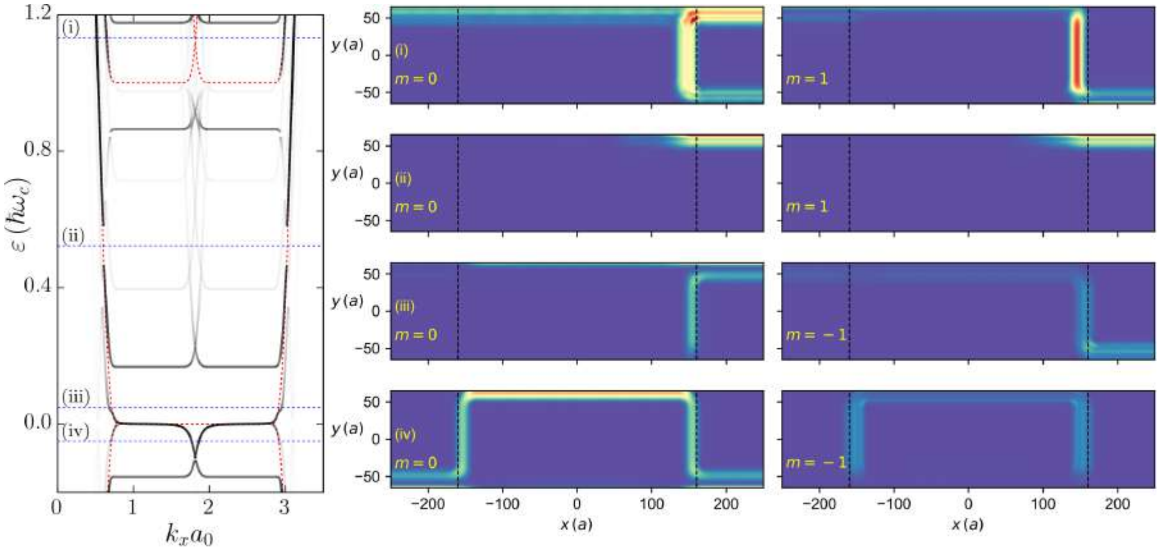


Figure 14. (Color online) Scattering states as a function of the position in a zigzag ribbon for four selected values of the quasienergy. The width of the ribbon is $W = 130\,a$. The parameters of the irradiated region (see Fig. 9) are $\lambda_1 = 130\,a$ and $\lambda_2 = 30\,a$, where a is the lattice parameter, $a = a_0\sqrt{3}$. The vertical dashed lines determine the central irradiated region. We use parameters $\zeta = 0.003$, $z = 0.025$, $\Omega = \omega_c$ and the polarization is circular and right-handed. Three Floquet replicas are used ($-1 \le m \le 1$). We show the scattering states projected over the m=0 replica (central column), and those projections over a replica different from zero carrying most of the weight. In all cases the electrons come from the right lead with velocities to the left and in the replica m=0, in the upper or lower edge depending on the quasienergy.

$\pi/2$) presents similar features, with the expected lack of electron-hole symmetry. It is worth mentioning that the mismatch problem that leads to the suppression of the conductance for $\varepsilon_F \simeq 0$ are already apparent in the non resonant case [Figs. 13(a) and (b)], as well as higher order narrow features. What is more, in resonance [Figs. 13(c) and (d), $\Omega = \omega_c$] the low energy conductance gap contains some fine structure inside it, which can be understood using the same arguments we introduced when dealing with the linearly polarized case. However, a better and more appealing way to analyse this is to look at the scattering states in Floquet space, as we do in the next section.

C. Scattering wave functions in Floquet space

In the previous sections we described some of the features of the two-terminal dc conductance in terms of the appearance of gaps in the projected Landau-Floquet spectral density, the emergence of light induced edge states with different chirality, or the mismatch between the leads' and the system's wavefunction. All these effects can be made clearer by looking at the Floquet scattering states. Namely, we calculate the (squared) amplitude of the scattered wavefunction corresponding to an incident wave coming from the right lead, for instance, with a given energy in the $m = 0$ Floquet replica. The results are shown in Figs. 14 and 15 for the zigzag and the armchair edge terminations, respectively. We have selected four particular values of the incident waves for illustration purposes.

In Fig. 14 (zigzag ribbon, $W = 130\,a$) the four selected scattered waves correspond to an incoming wave function with energy: (i) $\varepsilon/\hbar\omega_c = 1.15$, in which case there are three propagating edge modes in the leads and hence the corresponding equilibrium conductance is $6e^2/h$. Of those edge modes, only two can propagate through the irradiated region, as it can be clearly observed in the projected Floquet spectral density shown in Fig. 14(i), while the other is fully reflected at the interface between the two regions (mostly in the $m = 0$ replica on the other side of the sample). This leads to the quasi plateau of $4e^2/h$ shown in Fig. 11(d); (ii) $\varepsilon = \hbar\Omega/2$, well inside the dynamical gap. Here the incident wave penetrates the scattering region as an evanescent wave before being reflected, on the same side of the sample, into the $m = 1$ replica. Hence the conductance is strongly suppressed; (iii) $\varepsilon/\hbar\omega_c = 0.05$, inside the low energy Floquet gap but above the Dirac point. The incident wave is fully reflected on the $m = -1$ replica but on the other side of the sample; (iv) $\varepsilon/\hbar\omega_c = -0.05$. Here the incoming wave, now on the bottom edge of the sample due to its valence band character, is mostly fully transmitted through the $m = 0$ channel (with some participation of the $m = -1$ replica). However, quite remarkably, this requires the edge mode to *switch* edges inside the irradiated region. That is, the Floquet edge mode inside the sample presents the opposite chirality as compared to the one it has on the (non irradiated) leads. As we show in the next section, this leads to a change in the sign of the Hall conductance.

Figure 15 shows the scattering states for an armchair ribbon of width $W = 120\,a$. Here we concentrate on a small quasienergy region, $\varepsilon \in [-0.3\,\hbar\omega_c, 0.4\,\hbar\omega_c]$, where important departures from the static conductance appear [see Fig. 13(c)

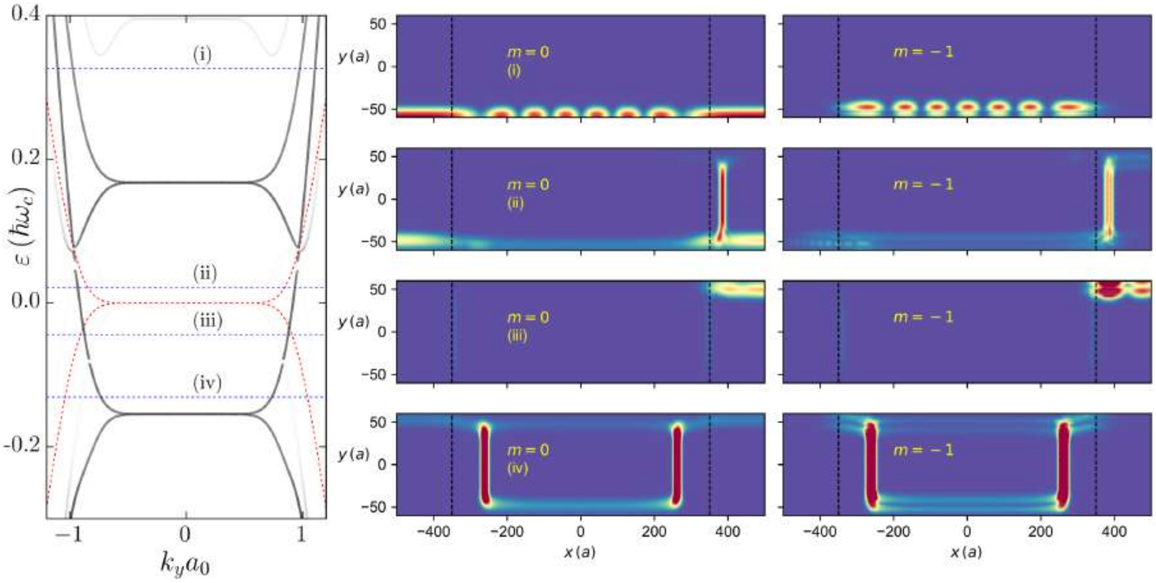


Figure 15. (Color online) Same as the previous figure but for an armchair ribbon of width $W = 120\,a$, $\lambda_1 \approx 450\,a$ and $2\lambda_2 \approx 200\,a$ ($a = \sqrt{3}a_0$). The parameters ζ and z, as well as the polarization of the laser, remain the same. The values of the quasienergies are: (i) $\varepsilon = 0.32\,\hbar\omega_c = 0.15\,\text{eV}$, (ii) $\varepsilon = 0.022\,\hbar\omega_c = 0.01\,\text{eV}$, (iii) $\varepsilon = -0.043\,\hbar\omega_c = -0.02\,\text{eV}$ and (iv) $\varepsilon = -0.11\,\hbar\omega_c = -0.05\,\text{eV}$. For purposes of clarity, the saturation in the cases (iii) and (iv) has been enhanced.

and (d)]. For all the quasienergies selected, the static system presents a conductance G_{2T} equal to $2e^2/h$, with only one propagating channel coming from the right lead. In every case, we see that most of the incident flux is scattered through the channel $m = -1$, with a negligible component going into the channel $m = 1$ (not shown here). This asymmetry in the roles of the Floquet replicas is a consequence of the time-reversal symmetry breaking imposed by the circularly polarized light and the selection rules Eq. (15) and (13). Mainly due to the non monotonous bands of the second branch (see Appendix), this case shows some special peculiarities not present in zigzag ribbons.

As before, we consider four relevant cases with different energies: (i) $\varepsilon/\hbar\omega_c = 0.33$, there are two available channels inside the irradiated region, as the result of the superposition and coupling of two replicas $m = 0$ and $m = -1$, for only one incident conduction channel. This leads to the appearance of oscillations in the probability density on each channel and to a displacement of the center of the orbits inside the illuminated region. The period of these oscillations is roughly equal to $2\pi/\delta k$, where δk is the difference between the wave vectors k_y, corresponding to the incident energy ε, of the two Floquet channels inside the sample. The transmittance in this region is almost perfect so $G_{2T} \simeq 2e^2/h$; (ii) $\varepsilon/\hbar\omega_c = 0.022$, the sign of the velocity of the incoming channel matches that of the (only) Floquet states in the irradiated region, and thus transport is possible, although imperfect due to the different spatial profile of both of them (incoming states are more centered towards the bulk while the Landau-Floquet states are closer the edge). This results in a not very well developed peak of conductance [$T(\varepsilon_F) \sim 0.9$], as it is shown in Fig. 13(d). In (iii) and (iv) the incoming channels move in the opposite direction to those available Landau-Floquet states inside the scattering region and on the same edge. Then, the only way for them to reach the other lead is to scatter into the other edge, where states with the same velocity are available for transport. Because of the adiabatic matching between the wavefunctions (as a consequence of the smooth turning on of the laser field), as it was already discussed in Sec. IV B, there is a certain energy threshold for this to happen set by the presence of a the small gap between (iii) and (iv) as shown in the Floquet spectral density (see Fig. 15). Then we have the following: (iii) $\varepsilon/\hbar\omega_c = -0.044$, here the incoming state cannot go through the avoided crossing and it is mostly reflected backward in the Floquet channel $m = -1$; (iv) $\varepsilon/\hbar\omega_c = -0.11$, the the electrons arriving from the right lead can partially go through the avoided crossing in the Floquet spectrum, reaching the other edge of the ribbon. Characteristic of this regimen is the noisy behavior of the conductance as a function of the Fermi level, leading to an incomplete transmission. Below $\varepsilon = -0.07\,\hbar\omega_c$, the velocities match and the transmission is perfect ($G_{2T} = 2e^2/h$).

V. HALL CONDUCTANCE

In order to measure the Hall conductance one requires at least four terminals. However, the usual (90°) four-terminal Hall bar configuration necessarily breaks the symmetry between leads, since if one pair of leads have zigzag edges, the other pair must have armchair edges. To avoid this, and to maintain the symmetry among all leads, we use the setup shown in Fig. 16. This six-terminal arrangement has also the advantage that is does not generate any pumped current when the laser is turned on in a symmetrical way, and hence

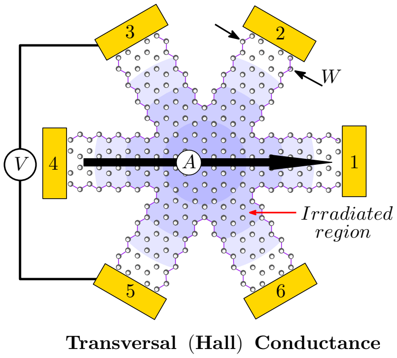


Figure 16. (Color online) Zigzag six-terminal Hall bar to measure the Hall conductance avoiding any pumped current. A similar setup can be established now with armchair leads.

it is simpler to obtain the conductance. Notwithstanding, the computational effort in calculating the scattering matrix [55] (and from this the Hall conductance) in this setup is more involved when compared to the two terminal setup. Following Ref. [41], the time averaged current is now written as a generalization of Eq. (18). If α and β label the terminals (leads), the average current through the α lead is given by

$$\bar{I}_\alpha = \frac{2e}{h} \sum_{\beta\neq\alpha} \sum_n \int \left[T^{(n)}_{\beta\alpha}(\varepsilon) f_\alpha(\varepsilon) - T^{(n)}_{\alpha\beta}(\varepsilon) f_\beta(\varepsilon) \right] d\varepsilon \,. \tag{21}$$

The transmittances $T^{(n)}_{\alpha\beta}$ are the multi-terminal generalizations of those in Eq. (18). Our six-terminal setup guarantees that $T^{(n)}_{\alpha\beta}$ is the same for any pair of adjacent terminals, ruling out the presence of any pumped current in the absence of a voltage bias. In what follows, we assume that the chemical potential μ_α at the lead α is not very different from its equilibrium value ε_F, that is, $\mu_\alpha = \varepsilon_F + \delta\mu_\alpha$ with $\delta\mu_\alpha$ small. Taking the low temperature limit and expanding Eq. (21) up to linear terms in $\delta\mu_\alpha = -eV_\alpha$ we get to

$$\bar{I}_\alpha = \frac{2e^2}{h} \sum_{\beta\neq\alpha} \left[T_{\beta\alpha}(\varepsilon_F)\, V_\alpha - T_{\alpha\beta}(\varepsilon_F)\, V_\beta \right] . \tag{22}$$

Our goal is to study the Hall conductance, as indicated in Fig. 16. To this we let a charge current flow from leads 4 to 1, $\bar{I}_4 = -\bar{I}_1 = I$, and impose $\bar{I}_\alpha = 0$ on the remaining leads. From this we define the Hall conductance as $G_H = I/(V_3 - V_5)$. Furthermore, hereon we shall consider only the circularly polarized resonant case, as it is the one with the most interesting features.

Figure 17 shows the Hall conductance for a six-terminal system with zigzag leads (width $W = 130\,a$), as a function of the Fermi level ε_F for both chiralities of the circular polarization: G_H^{ccw} counterclockwise (ccw, ↺) and G_H^{cw} clockwise (cw, ↻)—the Hall conductance in the absence of a laser field is plotted in black dashed lines. The parameters are the same we have been using so far: $\zeta = 0.003$ and $z = 0.025$. Clearly, $G_H^{\mathrm{ccw}}(\varepsilon_F) = -G_H^{\mathrm{cw}}(-\varepsilon_F)$, as expected. Several features are apparent in the figure, which are directly related to those of the Floquet edge states discussed in the previous sections: (i) the suppression of the Hall signal around the dynamical gaps, $\varepsilon_F \simeq \pm\hbar\Omega/2$, where the scattering states only penetrate as evanescent waves inside the scattering region and hence the incoming states are almost fully backscattered; (ii) the presence of an intermediate plateau with $G_H \approx 4e^2/h$ near $\varepsilon_F = \hbar\omega_c$; (iii) an additional suppression for energies $\varepsilon_F > 0$ ($\varepsilon_F < 0$) near the Dirac point for the ccw (cw) case; and (iv) a switch of the sign the Hall conductance for energies right below (above) the gap mentioned in (c) for the ccw (cw) case.

It is worth emphasizing that, quite remarkably, the switch of the Hall signal depends on the polarization of the laser field. Furthermore, if we set ε_F to lay inside of the low energy gap [item (iii) above] for a given polarization, the Hall conductance can be turned on by simply changing the chirality of the polarization. This is depicted in the insets (b) and (c) of Fig. 17 where we show the scattering states for three selected energies. For the ccw polarization most of the incoming states are backscattered into the same lead, which results in $G_H^{\mathrm{ccw}} \simeq 0$. On the contrary, for the cw polarization, the incoming states cross over the width of the ribbon and reaches the other edge, from where it propagates to lead 6, resulting in a negative Hall signal.

When the leads have armchair edges, G_H changes as shown in Fig. 18. Here, the parameters ζ and z remain the same as in the zigzag case, while the width of the leads is now $W = 120\,a$. First we notice that even in the absence of illumination (dashed black line), the armchair six-terminal geometry is accompanied by a certain degree of back scattering. This effect takes place only in the ε_F-range between the local minimum of the bands belonging to the second branch (see Appendix A), that in Fig. 18 are $|\boldsymbol{\chi}^{c}_{\nu_2 k}\rangle$, and the corresponding bulk level, where counter propagating edge states do exist on the same side of the sample. This leads to the appearance of a shoulder-like structure just below $\varepsilon_F = \hbar\omega_c$, that is in the transition from $G_H = 2e^2/h$ to $6e^2/h$. (The same is true near $\varepsilon_F = -\hbar\omega_c$.) This effect becomes less noticeable as the absolute value of the Landau bulk index $|n|$ increases.

When the laser is applied, the departures from the static Hall conductance (black dashed line) take place in a small vicinity around the Dirac point, $\varepsilon_F = 0$, and around $\pm\hbar\omega_c$, depending on the polarization (ccw or cw). Let us describe what happens in a vicinity of $\varepsilon_F = 0$ —as before, the scattering states for three selected energies, are shown as insets. Firstly, for positive values of ε_F, $G_H^{\mathrm{ccw}}(\varepsilon_F)$ exhibits a somewhat wide dip, indicating the presence of an avoided crossing in the spectrum. Apart from this, the Hall signal roughly follows its non driven value, as it can also be inferred from inset (b). On the other hand, $G_H^{\mathrm{ccw}}(\varepsilon_F)$ vanishes just below $\varepsilon_F = 0$, and changes sign for ε_F further below. The former behavior is analogous to that exhibit by the single ribbon, Sec. IV B: electrons hitting the illuminated region from lead 1 (see inset on the far left), encounter counter propagating states there, but at the same time are unable to reach the opposite edge of the lead, where propagation would be possible, and simply backscatter. Inset (a) graphically describes the situation. In the situation where the Hall signal is reversed, the incident states reach the other side of the

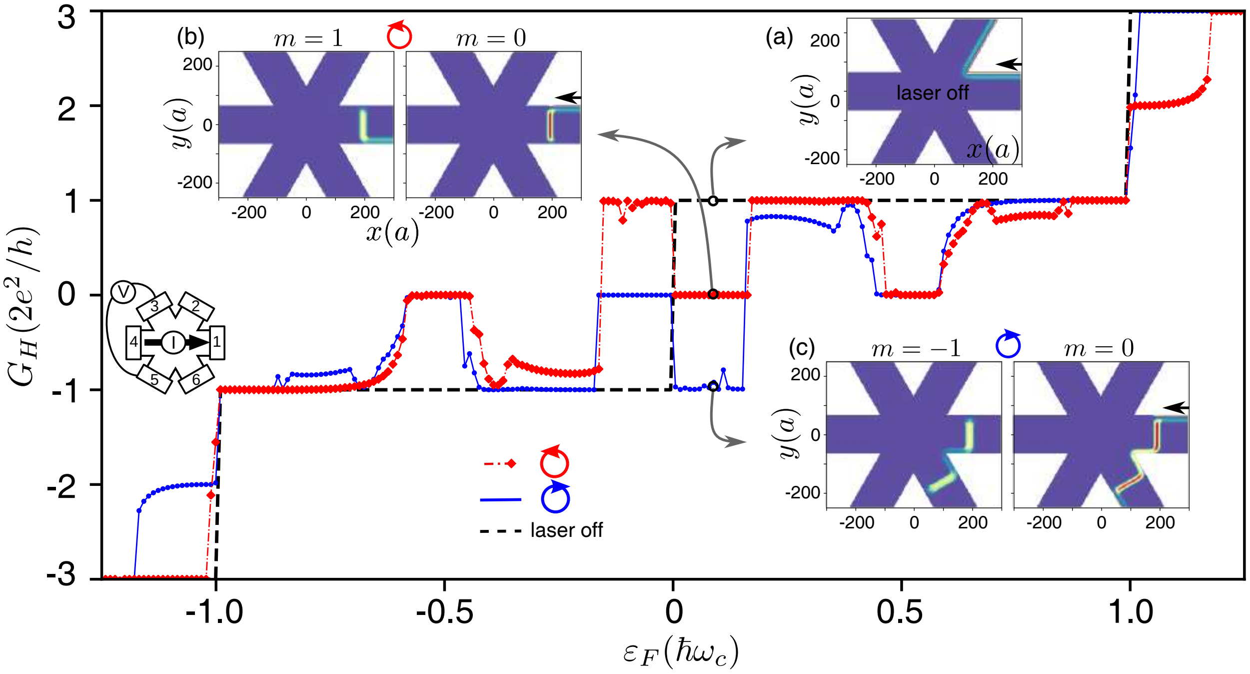


Figure 17. (Color online) Hall conductance G_H in a six-terminal setup (see inset) as a function of the Fermi level ε_F. The width of the leads is $W = 224\,a$ and the magnetic field corresponds to $\zeta = 0.003$. The laser field is circularly polarized as indicated. The different insets shows the scattering states for the selected ε_F indicated in the main plot. In all cases the incoming state is in the channel $m = 0$ (black arrow). Most of the scattering is into the $m = 0$ channel and the $m = 1$ ($m = -1$) for the ccw (cw) polarization. Both the switch-off of the conductance and the change of chirality are clearly seen for the ccw and cw polarization, respectively.

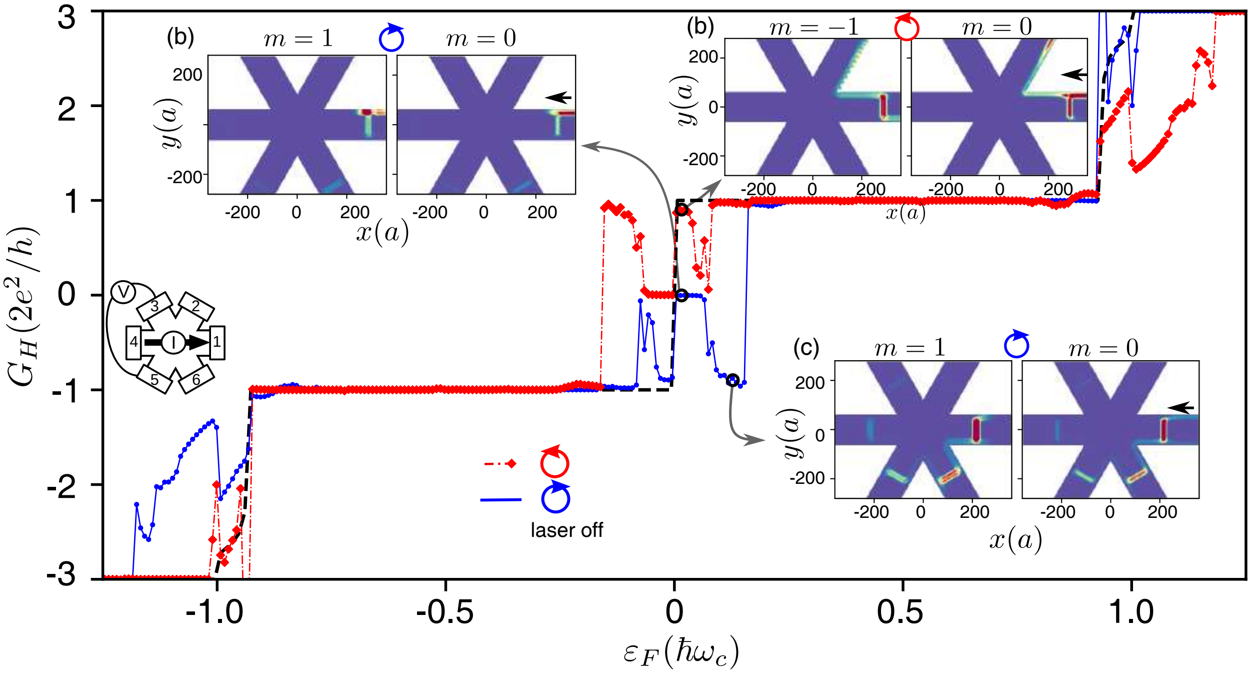


Figure 18. (Color online) Same as in the previous figure but with armchair terminations (leads) of width $W = 120\sqrt{3}\,a_0$. The insets correspond to quasienergies: (a) $\varepsilon_F = 0.01$ eV, cw; (b) $\varepsilon_F = 0.01$ eV, ccw; (c) $\varepsilon_F = 0.06$ eV, cw. The saturation has been enhanced for a better visualization of the incoming states, particularly in (c).

ribbon and spill into lead 6, and the Hall signal passes from $-2e^2/h$ to $2e^2/h$. In the inset (c) the situation is shown for $G_H^{\rm cw}(\varepsilon_F)$ and for a positive energy, although due to the symmetry $G_H^{\rm ccw}(\varepsilon_F) = -G_H^{\rm cw}(-\varepsilon_F)$ the result is entirely equivalent.

VI. SUMMARY

The quantum Hall effect is a paradigmatic topological phase, one of the most robust available phases, a fact that has proved useful for many applications. These developments came hand in hand with concrete measurements in different device setups. Here we present a study of the effect of strong laser illumination on graphene in the Quantum Hall regime and the changes in the Hall response in a multi-terminal setup that is accounted for by an atomistic description. Our results show that the quantum Hall plateaus can be disrupted or tailored by tuning the frequency, polarization and intensity of the laser field. This includes the switching on-off and a chirality inversion of the Hall signal, by tuning the handedness of the circular polarization.

On one hand, the driving laser field induces the appearance of dynamical gaps in the Floquet spectrum that manifest, for instance, in the suppression of the Hall conductance near $\varepsilon_F = \pm\hbar\Omega/2$—the fact that there is essentially a full suppression is related to the nature of the equilibrium transport in the QH regime which only occurs at the edges of the sample [54]. This quench of the topologically protected QH transport arises from the resonant coupling between counter-propagating electron and hole edge states. Quite interestingly, the $\pm\hbar\Omega/2$ gaps are absent in samples with armchair edges and hence no suppression of the conductance is observed in such a case. This is a very particular aspect of the armchair termination, where the symmetry of the Hall edge modes leads to the cancellation of the dominant matrix elements of the time dependent perturbation. We have verified (but have not shown here) that the cove termination leads to similar results as the ones reported here for the zigzag case, and so we expect the suppression of the Hall signal at $\varepsilon_F = \pm\hbar\Omega/2$ to represent the general situation. It is also worth mentioning that, under resonant conditions, and more clearly in the zigzag ribbon, the driving produces a novel $4e^2/h$ feature in the Hall conductance due to the reduction of the available edge channels near $\varepsilon_F = \pm\hbar\omega_c$.

On the other hand, near the Dirac point, the driving becomes a non-equilibrium knob that allows to turn on and off the Hall signal or changing its sign. Quite remarkably, both effects depend on the polarization of the laser field. The possibility to turn transport on and off relies on the appearance, under resonant conditions, of an effective transport gap right above or below the Dirac point, depending on the polarization's handedness. The change in sign of the Hall signal have a less anticipated and more striking origin. It results from a change of the chirality of the propagating modes inside the irradiated region that makes the propagating edge channel to cross from one side of the sample (ribbon or lead) to the other while inside the driving area. In the zigzag case, for instance, the physical origin for this is the fact that the zero energy flat edge modes acquire a polarization-dependent dispersion that is always opposite to the one of the regular QH edge states. We expect this to be a generic feature for all terminations where the K and K' valleys are separate and not fold onto each other. In this sense, the armchair case appears as the only exception.

We hope that the effects described in the present work will help in stimulating new experiments and theory on the interplay between chiral transport and periodic driving as this might open the door to new ways to control it. This includes systems in a QH phase as discussed here, but also topological insulators where spin-orbit coupling may add further intricacies [56].

ACKNOWLEDGMENTS

We acknowledge financial support from ANPCyT (grants PICTs 2016-0791 and 2018-01509), from CONICET (grant PIP 11220150100506), from SeCyT-UNCuyo (grant 2019 06/C603), FondeCyT (Chile) under grant number 1170917 and the EU Horizon 2020 research and innovation program under the Marie-Sklodowska-Curie Grant Agreement No. 873028 (HYDROTRONICS Project). GU and LEFFT acknowledge support from the ICTP associateship program and thank the Simons Foundation. LEFFT acknowledges discussions with Nikolai Kalugin, Paola Barbara and James McIver.

Appendix A: Landau levels in graphene

For the sake of completeness we present here a brief description of the Landau levels in graphene in the low energy approximation, which are obtained from Eq. (1) after the substitution $\boldsymbol{p} \to \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}$.

Before starting, it is useful to point out the following symmetries of the hamiltonian $\hat{\mathcal{H}}_0$,

$$\tau_x \otimes \sigma_y \, \hat{\mathcal{H}}_0 \, \tau_x \otimes \sigma_y = \hat{\mathcal{H}}_0 \,, \qquad \tau_x \otimes \sigma_x \, \hat{\mathcal{H}}_0 \, \tau_x \otimes \sigma_x = -\hat{\mathcal{H}}_0 \,, \tag{A1}$$

that allow us to relate the eigenfunctions belonging to different valleys (before imposing the boundary condition) when the problem can be solved on each of them separately (see bellow).

1. Bulk states

In the Landau gauge, $\boldsymbol{A}(y) = -By\,\hat{\boldsymbol{x}}$, the Hamiltonian Eq. (1) is invariant under translations in x direction so that the eigenfunctions (for the K valley) can be written as $\Psi_n(x,y) = L_x^{-\frac{1}{2}}\, e^{-ikx} \boldsymbol{\chi}_{nk}^K(y)$. Here L_x is the samples's length along the x direction and

$$\boldsymbol{\chi}_{nk}^K(y) = \frac{1}{\sqrt{\ell_B\,(2-\delta_{n0})}} \begin{pmatrix} \phi_{|n|}(\tilde{y}) \\ \mathrm{sgn}(n)\phi_{|n|-1}(\tilde{y}) \end{pmatrix}, \tag{A2}$$

where $\tilde{y} = y/\ell_B - k\ell_B$, $\ell_B = \sqrt{\hbar c/eB}$ is the *magnetic length*, and $\phi_n(x)$ is the normalized harmonic oscillator eigenfunction, $\phi_n(x) = (\sqrt{\pi}\, 2^n n!)^{-1/2}\, e^{-x^2/2} H_n(x)$ with $H_n(x)$ the Hermite polynomial. The corresponding eigenvalues are $E_n{=}\mathrm{sgn}(n)\hbar\omega_c\sqrt{|n|}$, where $n \in \mathbb{Z}$, sgn(n) is the sign function and $\omega_c = \sqrt{2}v_F/\ell_B$. Positive or negative values of n correspond to electrons and holes respectively. Solutions for the K' valley can be obtained by applying the operator σ_y to Eq. (A2), by virtue of Eqs. (A1).

2. Zigzag edge

The Landau levels corresponding to a semi-infinite plane with zigzag edges ending on a 'B' site can be obtained by requiring that the component 'A' of the wavefunction be zero at the edge, which for the case depicted in Fig. 1 corresponds to $y = 0$—notice that (with the gauge $\boldsymbol{A}(y) = -By\,\hat{\boldsymbol{x}}$) translational symmetry along the x direction is preserved. Since this edge termination does not mix the two valleys K and K', the solutions can be found for each valley separately. These solutions still have the form $\Psi(x,y) = L_x^{-\frac{1}{2}} \mathrm{e}^{-ikx} \boldsymbol{\chi}_k(y)$ but the components of $\boldsymbol{\chi}_k(y)$ are not the harmonic oscillator eigenfunctions but the general solutions of the corresponding differential equations, and that are well-behaved in the limit $y \to \infty$. These are the Parabolic Cylinder functions $D_\nu(x)$ with $\nu \in \mathbb{R}$ and $\nu \geq 0$. Hence we have [57]

$$\boldsymbol{\chi}^K_{\nu k}(y) = \frac{1}{\sqrt{C_{\nu k}}} \begin{pmatrix} D_\nu(\xi) \\ \frac{\varepsilon}{\varepsilon_0} D_{\nu-1}(\xi) \end{pmatrix}, \tag{A3}$$

$$\boldsymbol{\chi}^{K'}_{\nu k}(y) = \frac{1}{\sqrt{C_{\nu k}}} \begin{pmatrix} \frac{\varepsilon}{\varepsilon_0} D_{\nu-1}(\xi) \\ -D_\nu(\xi) \end{pmatrix}, \tag{A4}$$

where $\xi{=}\sqrt{2}(y/\ell_B - k\ell_B)$ and $C_{\nu k}$ is a normalization constant. The quantity kl_B can be considered the *center* of the cyclotron orbit of the electrons, and thus the larger kl_B, the deeper into the bulk the electrons reside. The quantity $\varepsilon = s\,\hbar\omega_c\sqrt{\nu}$ is the energy of the state ($s = \pm 1$ refers to the electron and hole bands respectively), with $\omega_c{=}\sqrt{2}v_F/\ell_B$.

The index $\nu(k)$, and thus the energy dispersion, is determined by the boundary condition that the upper component (sublattice A) of the spinor wavefunction vanish at position $y = 0$. That is

$$D_\nu(-\sqrt{2}k\ell_B) = 0\,, \qquad \sqrt{\nu}\, D_{\nu-1}(-\sqrt{2}k\ell_B) = 0, \tag{A5}$$

for the K and K' valleys, respectively. This gives a discrete set of eigenenergies $\varepsilon^K_l(k)$ and $\varepsilon^{K'}_l(k)$ for each value of k, where the integer l labels the solutions of Eqs. (A5). It is straightforward to verify that because for the same value of ν_l the two spinors in the conduction and valence bands (opposite s) must be orthonormal, one has that

$$\int_{-\sqrt{2}k\ell_B}^{+\infty} d\xi\, D^2_{\nu_l}(\xi) = \int_{-\sqrt{2}k\ell_B}^{+\infty} d\xi\, \nu_l D^2_{\nu_l-1}(\xi)\,, \tag{A6}$$

and hence the normalization constant becomes

$$C_{\nu_l k} = \sqrt{2}\ell_B \int_{-\sqrt{2}k\ell_B}^{\infty} d\xi\, D^2_{\nu_l}(\xi)\,. \tag{A7}$$

Figures 19(a) and (b) show the energy dispersion for both valleys where dispersive edge states are apparent. Far from the boundary ($k\ell_B \gg 1$) one has that $\nu_l \to n \in \mathbb{N}^0$, and so the bulk Landau levels $\pm\,\hbar\omega_c\sqrt{n}$ are recovered. Note that the K' valley possesses dispersionless states pinned at the Dirac point, the corresponding spinor being the same as for the zero energy bulk Landau level.

The edge states in the K valley close to zero energy deserve a separate and careful treatment. From Fig. 19(a), it is clear that the lowest conduction and the highest valence bands converge into each other and into $\varepsilon = 0$ as $k\ell_B$ becomes large. At first sight, it might appear, since in this case $\nu \to 0$, that the two spinor states themselves go into the same spinor proportional to $[D_0(\xi), 0]^T$, with a vanishing second (B) component. This state clearly could be identified with the bulk $n = 0$ Landau level [cf. Eq. (A2) with $n = 0$ and $\phi_0(z) = D_0(z)$]. However, Eq. (A6) shows that even in this limit ($\nu \approx 0$), the B-component does not vanish as the left side of the equation converges to $\sqrt{2\pi}$ when $k\ell_B \to \infty$. To properly account for this one needs to take into account the precise way in which $\nu \to 0$ as $k\ell_B \to \infty$. After expanding the equation $D_\nu(-\sqrt{2}k\ell_B) = 0$ for $k\ell_B \gg 1$, the solution for vanishing ν is given by $\nu \simeq \pi^{-\frac{1}{2}} k\ell_B \exp\left[-(k\ell_B)^2\right]$ and therefore the B-component of the spinor wavefunction at the edge goes as $\sqrt{\nu}D_{\nu-1}(\xi) \simeq \pi^{\frac{1}{4}}(2k\ell_B)^{\frac{1}{2}} e^{-ky}$. That is, the B-component gets sharper as $k\ell_B$ grows and so it is more localized at the edge (recall that the condition Eq. (A6) must hold). In summary, for $k\ell_B \gg 1$ the two orthonormal eigenstates have the A-component located mostly in bulk, and the B-component at the edge. The fact that ν is almost zero allow us to treat them as nearly degenerate, and thus change the basis to a symmetrical $\boldsymbol{\chi}_{\mathrm{bulk}}$ and antisymmetrical $\boldsymbol{\chi}_{\mathrm{edge}}$ solutions,

$$\boldsymbol{\chi}_{\mathrm{bulk}}(y) = \sqrt{\frac{2}{C_{\nu k}}} \begin{pmatrix} D_\nu(\xi) \\ 0 \end{pmatrix}, \tag{A8}$$

$$\boldsymbol{\chi}_{\mathrm{edge}}(y) = \sqrt{\frac{2}{C_{\nu k}}} \begin{pmatrix} 0 \\ \sqrt{\nu} D_{\nu-1}(\xi) \end{pmatrix}. \tag{A9}$$

In this representation the spatial profile of each spinor is evident. As it was already pointed out, $\boldsymbol{\chi}_{\mathrm{bulk}}$ is to be identified with the $n = 0$ bulk Landau level, whereas $\boldsymbol{\chi}_{\mathrm{edge}}$ is the well known non-dispersive edge state in a zigzag graphene ribbon [57, 58].

3. Armchair edge

To be consistent with the geometry shown in Figs. 1 and 5, we take the armchair edge to be along the $x = 0$ line and therefore, to make use of the translation symmetry along that direction, we use a different but completely physically equivalent gauge, namely $\boldsymbol{A}(x){=}Bx\,\hat{\boldsymbol{y}}$. The wavefunction

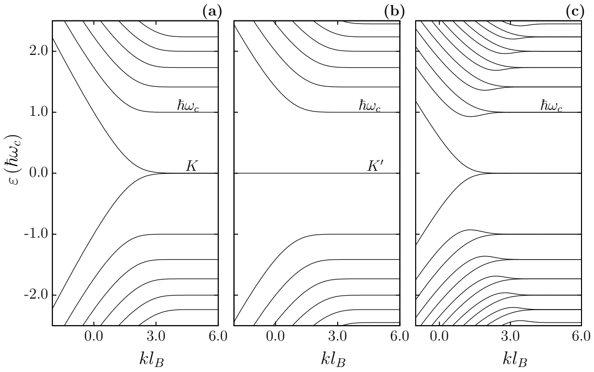


Figure 19. Energy dispersion (in units of $\hbar\omega_c$) for the K [(a)] and K' [(b)] Dirac valleys in a zigzag edge. Note that the K'-valley have a dispersionless mode with $\varepsilon(k) = 0$ while the K-valley possesses two states that converge to the $n = 0$ bulk Landau level. In [(c)] we show the energy dispersion for an armchair ribbon. No distinction between valleys has place here. In every case $\zeta = 0.003$. The first non zero bulk Landau level $E_1 = \hbar\omega_c$ is indicated.

on each valley can then be written as $L_y^{-1/2}\, e^{iky} \boldsymbol{\chi}_{\nu k}^{K(K')}(x)$, where the spinor components are given by

$$\boldsymbol{\chi}_{\nu k}^{K}(x) = \frac{1}{\sqrt{\tilde{C}_{\nu k}}} \begin{pmatrix} D_\nu(\tilde{\xi}) \\ -i\frac{\varepsilon}{\varepsilon_0} D_{\nu-1}(\tilde{\xi}) \end{pmatrix}, \tag{A10}$$

$$\boldsymbol{\chi}_{\nu k}^{K'}(x) = \frac{1}{\sqrt{\tilde{C}_{\nu k}}} \begin{pmatrix} i\frac{\varepsilon}{\varepsilon_0} D_{\nu-1}(\tilde{\xi}) \\ D_\nu(\tilde{\xi}) \end{pmatrix}. \tag{A11}$$

Here $\tilde{\xi} = \sqrt{2}(x/\ell_B - k\ell_B)$ and $\varepsilon = s\hbar\omega_c\sqrt{\nu}$. If we denote the complete wavefunction of the armchair edge as $L_y^{-1/2}\, e^{iky} \boldsymbol{\chi}_{\nu k}(x)$, then the boundary condition in the armchair edge corresponds to make $\boldsymbol{\chi}_{\nu k}(0) = 0$. This cannot be made separately in each valley (the boundary mixes the valley index) and so the solution needs to be constructed using a linear combinations of both valleys [57, 59]

$$\boldsymbol{\chi}_{\nu k}(x) = \alpha\, e^{iKx}\, \boldsymbol{\chi}_{\nu k}^{K}(x) + \beta\, e^{-iKx}\, \boldsymbol{\chi}_{\nu k}^{K'}(x)\,, \tag{A12}$$

where we have added a phase factor $e^{i\boldsymbol{K}\cdot\boldsymbol{r}}$ ($e^{i\boldsymbol{K}'\cdot\boldsymbol{r}}$) to the K (K') component, noticing that in our geometry (see Fig. 5) we have $\boldsymbol{K}' = -\boldsymbol{K} = -K\hat{\boldsymbol{x}}$, with $K = 4\pi/(3\sqrt{3}a)$. The condition $\boldsymbol{\chi}_{\nu k}(0) = 0$ results in the following equation for $\nu(k)$

$$D_\nu^2(-\sqrt{2}k\ell_B) - \nu\, D_{\nu-1}^2(-\sqrt{2}k\ell_B) = 0\,. \tag{A13}$$

This equation yields a discrete infinite set of solutions for each k, that in ascending value we denote as $\nu_n(k)$, with $n = 1, 2, \cdots$. It is interesting to note that these solutions can be arranged in a way that simplifies the construction of their corresponding eigenstates. To this, we first note that Eq. (A13) can be written as follows (all D_ν functions are evaluated at $-\sqrt{2}k\ell_B$)

$$(D_\nu - \sqrt{\nu}\, D_{\nu-1})(D_\nu + \sqrt{\nu}\, D_{\nu-1}) = 0\,. \tag{A14}$$

It can be shown that solutions ν_n with n odd satisfy $D_\nu - \sqrt{\nu}\, D_{\nu-1} = 0$, while those with n even are solutions of $D_\nu + \sqrt{\nu}\, D_{\nu-1} = 0$. To make things more compact, we introduce an index $\tau_n = (-1)^{n+1}$. We then say that the ν_n solutions exist in two *branches* defined by the equation $D_{\nu_n}(-\sqrt{2}k\ell_B) = \tau_n\sqrt{\nu_n}\, D_{\nu_n-1}(-\sqrt{2}k\ell_B)$.

The energy spectrum $\varepsilon_n^s(k) = s\,\hbar\omega_c\sqrt{\nu_n(k)}$ ($s = \pm 1$) is shown in Fig. 19(c). The bands with n odd (the *first branch*), are monotonous decreasing functions of $k\ell_B$, very similar to those of a zigzag edge at the K point. For n even (the *second branch*), the bands are decreasing functions for very negative values of $k\ell_B$, until they reach a local minimum and start to grow as $k\ell_B$ increases. For $k\ell_B$ large enough ($k\ell_B > 3$ if we consider the first six bands), we see that adjacent pairs of bands converge to each other and into the bulk states: $\{\nu_n, \nu_{n+1}\} \to \sqrt{n/2}$, with $n = 2, 4, 6, \cdots$.

We now examine the specific structure of the corresponding eigenfunctions as it plays a crucial role in determining the value of the gaps in the Floquet spectrum. To this, we first notice that the coefficients α and β in Eq. (A12) satisfy $\alpha_n = -i\beta_n s\tau_n$. With this result it is useful to write down these eigenfunctions in the four-component spinor notation,

$$\boldsymbol{\chi}_{\nu_n k}^{s}(x) = \frac{1}{\sqrt{C_{\nu_n k}}} \begin{bmatrix} -is\tau_n D_{\nu_n}(\tilde{\xi})e^{iKx} \\ -\tau_n\sqrt{\nu_n}D_{\nu_n-1}(\tilde{\xi})e^{iKx} \\ is\sqrt{\nu_n}D_{\nu_n-1}(\tilde{\xi})e^{-iKx} \\ D_{\nu_n}(\tilde{\xi})e^{-iKx} \end{bmatrix}, \tag{A15}$$

where the s parameter indicates electron/conduction (s=1) or hole/valence (s=-1) character, which in the main text will be often denoted by superscripts c and v, respectively. Since the probability density is obtained by summing up sub-lattice (A and B) components, it can be shown that the spatial profile of $\boldsymbol{\chi}_{\nu_n k}^{s}(x)$ across the ribbon shows oscillations of the form $\cos 2Kx$, with a period equal to $1/2K$ (the inverse of the distance between the two non equivalent valleys [60]). The normalization constant $C_{\nu k}$ is given by

$$C_{\nu_n k} = \sqrt{2}\,\ell_B \int_{-\sqrt{2}kl_B}^{+\infty} \mathrm{d}\xi\, [D_{\nu_n}^2(\xi) + \nu_n D_{\nu_n-1}^2(\xi) - \tau_n\sqrt{\nu_n}D_{\nu_n}(\xi)D_{\nu-1}(\xi)\cos(2Kx)]. \tag{A16}$$

From the general solutions in Eq. (A15), the matrix elements of the time dependent perturbation $\mathcal{V}$, Eq. (13), can be evaluated.

[1] K. von Klitzing, G. Dorda, and M. Pepper, "New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance," Phys. Rev. Lett. **45**, 494 (1980).

[2] K. von Klitzing, T. Chakraborty, P. Kim, V. Madhavan, X. Dai, J. McIver, Y. Tokura, L. Savary, D. Smirnova, A. M. Rey, C. Felser, J. Gooth, and X. Qi, "40 years of the quantum Hall effect," Nature Reviews Physics , 1 (2020).

[3] R. B. Laughlin, "Quantized Hall conductivity in two dimensions," Physical Review B **23**, 5632 (1981).

[4] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized Hall Conductance in a Two-Dimensional Periodic Potential," Phys. Rev. Lett. **49**, 405 (1982).

[5] M. Z. Hasan and C. L. Kane, "*Colloquium* : Topological insulators," Rev. Mod. Phys. **82**, 3045 (2010).

[6] Y. Ando, "Topological Insulator Materials," Journal of the Physical Society of Japan **82**, 102001 (2013).

[7] F. D. M. Haldane, "Nobel Lecture: Topological quantum matter," Reviews of Modern Physics **89**, 040502 (2017).

[8] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, "Quantum Spin Hall Insulator State in HgTe Quantum Wells," Science **318**, 766 (2007).

[9] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, "A topological Dirac insulator in a quantum spin Hall phase," Nature **452**, 970 (2008).

[10] S.-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C.-C. Lee, S.-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, "Discovery of a Weyl Fermion semimetal and topological Fermi arcs," Science , aaa9297 (2015).

[11] K. von Klitzing, "Essay: Quantum Hall Effect and the New International System of Units," Physical Review Letters **122**, 200001 (2019).

[12] T. Oka and H. Aoki, "Photovoltaic Hall effect in graphene," Phys. Rev. B **79**, 081406 (2009).

[13] N. H. Lindner, G. Refael, and V. Galitski, "Floquet topological insulator in semiconductor quantum wells," Nat Phys **7**, 490 (2011).

[14] T. Kitagawa, T. Oka, A. Brataas, L. Fu, and E. Demler, "Transport properties of nonequilibrium systems under the application of light: Photoinduced quantum Hall insulators without Landau levels," Phys. Rev. B **84**, 235108 (2011).

[15] M. S. Rudner, N. H. Lindner, E. Berg, and M. Levin, "Anomalous edge states and the bulk-edge correspondence for periodically-driven two dimensional systems," Phys. Rev. X **3**, 031005 (2013).

[16] G. Usaj, P. M. Perez-Piskunow, L. E. F. Foa Torres, and C. A. Balseiro, "Irradiated graphene as a tunable Floquet topological insulator," Phys. Rev. B **90**, 115423 (2014).

[17] M. S. Rudner and N. H. Lindner, "Band structure engineering and non-equilibrium dynamics in floquet topological insulators," Nature Reviews Physics **2**, 229 (2020).

[18] "Quantum phases on demand," Nature Physics **16**, 1 (2020).

[19] F. Giustino, M. Bibes, J. H. Lee, F. Trier, R. Valentí, S. M. Winter, Y.-W. Son, L. Taillefer, C. Heil, A. I. Figueroa, B. Plaçais, Q. Wu, O. V. Yazyev, E. P. A. M. Bakkers, J. Nygård, P. Forn-Díaz, S. de Franceschi, L. E. F. Foa Torres, J. McIver, A. Kumar, T. Low, R. Galceran, S. O. Valenzuela, M. V. Costache, A. Manchon, E.-A. Kim, G. R. Schleder, A. Fazzio, and S. Roche, "The 2020 Quantum Materials Roadmap," Journal of Physics: Materials (2020), 10.1088/2515-7639/abb74e.

[20] Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik, "Observation of Floquet-Bloch States on the Surface of a Topological Insulator," Science **342**, 453 (2013).

[21] F. Mahmood, C.-K. Chan, Z. Alpichshev, D. Gardner, Y. Lee, P. A. Lee, and N. Gedik, "Selective scattering between Floquet-Bloch and Volkov states in a topological insulator," Nature Physics **12**, 306 (2016).

[22] J. W. McIver, B. Schulte, F.-U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, "Light-induced anomalous Hall effect in graphene," Nature Physics **16**, 38 (2020).

[23] K. Wintersperger, C. Braun, F. N. Ünal, A. Eckardt, M. D. Liberto, N. Goldman, I. Bloch, and M. Aidelsburger, "Realization of an anomalous Floquet topological system with ultracold atoms," Nature Physics , 1 (2020).

[24] A. Eckardt, "Colloquium: Atomic quantum gases in periodically driven optical lattices," Reviews of Modern Physics **89**, 011004 (2017).

[25] N. Cooper, J. Dalibard, and I. Spielman, "Topological bands for ultracold atoms," Reviews of Modern Physics **91**, 015005 (2019).

[26] H. L. Calvo, H. M. Pastawski, S. Roche, and L. E. F. Foa Torres, "Tuning laser-induced band gaps in graphene," Appl. Phys. Lett. **98**, 232103 (2011).

[27] P. M. Perez-Piskunow, G. Usaj, C. A. Balseiro, and L. E. F. Foa Torres, "Floquet chiral edge states in graphene," Phys. Rev. B **89**, 121401(R) (2014).

[28] S. H. Kooi, A. Quelle, W. Beugeling, and C. Morais Smith, "Genesis of the Floquet Hofstadter butterfly," Physical Review B **98**, 115124 (2018).

[29] L. Du, Q. Chen, A. D. Barr, A. R. Barr, and G. A. Fiete, "Floquet Hofstadter butterfly on the kagome and triangular lattices," Physical Review B **98**, 245145 (2018).

[30] M. Wackerl, P. Wenk, and J. Schliemann, "Driven Hofstadter butterflies and related topological invariants," Physical Review B **100**, 165411 (2019).

[31] M. Zhao, Q. Chen, X.-D. Tian, and L. Du, "Floquet engineering the Hofstadter butterfly in the square lattice and its effective Hamiltonian," arXiv:2007.01071 [cond-mat] (2020), arXiv: 2007.01071.

[32] A. López, A. Di Teodoro, J. Schliemann, B. Berche, and B. Santos, "Laser-induced modulation of the Landau level structure in single-layer graphene," Physical Review B **92**, 235411 (2015).

[33] K.-H. Ding, L.-K. Lim, G. Su, and Z.-Y. Weng, "Quantum Hall effect in ac driven graphene: From the half-integer to the integer case," Physical Review B **97**, 035123 (2018).

[34] H. Dehghani, T. Oka, and A. Mitra, "Out-of-equilibrium electrons and the Hall conductance of a Floquet topological insulator," Physical Review B **91**, 155422 (2015).

[35] H. Dehghani, T. Oka, and A. Mitra, "Dissipative Floquet topological systems," Physical Review B **90**, 195429 (2014).

[36] K. I. Seetharam, C.-E. Bardyn, N. H. Lindner, M. S. Rudner, and G. Refael, "Controlled Population of Floquet-Bloch States via Coupling to Bose and Fermi Baths," Physical Review X **5**, 041050 (2015).

[37] T. Iadecola, T. Neupert, and C. Chamon, "Occupation of topological Floquet bands in open systems," Physical Review B **91**, 235133 (2015).

[38] H. Dehghani and A. Mitra, "Occupation probabilities and current densities of bulk and edge states of a Floquet topological insulator," Physical Review B **93**, 205437 (2016).

[39] L. Peralta Gavensky, G. Usaj, and C. A. Balseiro, "Time-resolved Hall conductivity of pulse-driven topological quantum systems," Physical Review B **98**, 165414 (2018).

[40] M. Schüler, U. De Giovannini, H. Hübener, A. Rubio, M. A. Sentef, T. P. Devereaux, and P. Werner, "How Circular Dichroism in Time- and Angle-Resolved Photoemission Can Be Used to Spectroscopically Detect Transient Topological States in Graphene," Physical Review X **10**, 041013 (2020).

[41] L. E. F. Foa Torres, P. M. Perez-Piskunow, C. A. Balseiro, and G. Usaj, "Multiterminal Conductance of a Floquet Topological Insulator," Phys. Rev. Lett. **113**, 266801 (2014).

[42] M. Moskalets and M. Büttiker, "Floquet scattering theory of quantum pumps," Physical Review B **66**, 205320 (2002).

[43] S. Camalet, J. Lehmann, S. Kohler, and P. Hänggi, "Current Noise in ac-Driven Nanoscale Conductors," Phys. Rev. Lett. **90**, 210602 (2003).

[44] This Floquet scattering picture has been demonstrated to be equivalent to the non-equilibium Green's function formalism in the appropriate limit [49] and a dictionary between the two has been pointed out in Ref. [52].

[45] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene," Rev. Mod. Phys. **81**, 109 (2009).

[46] C. Bena and G. Montambaux, "Remarks on the tight-binding model of graphene," New Journal of Physics **11**, 095003 (2009).

[47] M. Grifoni and P. Hänggi, "Driven quantum tunneling," Physics Reports **304**, 229 (1998).

[48] G. Platero and R. Aguado, "Photon-assisted transport in semiconductor nanostructures," Physics Reports **395**, 1 (2004).

[49] S. Kohler, J. Lehmann, and P. Hänggi, "Driven quantum transport on the nanoscale," Physics Reports **406**, 379 (2005).

[50] J. H. Shirley, "Solution of the Schrödinger Equation with a Hamiltonian Periodic in Time," Phys. Rev. **138**, B979 (1965).

[51] T. Oka and H. Aoki, "Photovoltaic hall effect in graphene," Phys. Rev. B **79**, 081406 (2009).

[52] L. Arrachea and M. Moskalets, "Relation between scattering-matrix and Keldysh formalisms for quantum transport driven by time-periodic fields," Phys. Rev. B **74**, 245322 (2006).

[53] U. Bajpai, M. J. H. Ku, and B. K. Nikolić, "Robustness of quantized transport through edge states of finite length: Imaging current density in floquet topological versus quantum spin and anomalous hall insulators," **2**, 033438.

[54] A. Huamán and G. Usaj, "Floquet spectrum and two-terminal conductance of a transition-metal dichalcogenide ribbon under a circularly polarized laser field," Phys. Rev. B **99**, 075423 (2019).

[55] C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal, "Kwant: a software package for quantum transport," New Journal of Physics **16**, 063065 (2014).

[56] M. Berdakin, E. A. Rodríguez-Mena, and L. E. F. Foa Torres, "Spin-polarized tunable photocurrents," 2010.11883.

[57] V. P. Gusynin, V. A. Miransky, S. G. Sharapov, and I. A. Shovkovy, "Edge states in quantum hall effect in graphene (review article)," Low Temperature Physics **34**, 778 (2008).

[58] K. Nakada, M. Fujita, G. Dresselhaus, and M. Dresselhaus, "Edge state in graphene ribbons: Nanometer size effect and edge shape dependence," Phys. Rev. B **54**, 17954 (1996).

[59] L. Brey and H. Fertig, "Electronic states of graphene nanoribbons studied with the dirac equation," Phys. Rev. B **73**, 235411 (2006).

[60] L. Brey and H. A. Fertig, "Edge states and the quantized hall effect in graphene," Phys. Rev. B **73** (2006), 10.1103/physrevb.73.195408.

arXiv:2010.11883v2 [cond-mat.mes-hall] 31 Jan 2021

Spin-polarized tunable photocurrents

Matías Berdakin,[1, 2, *] Esteban A. Rodríguez-Mena,[3, *] and L. E. F. Foa Torres[3]
[1]*INFIQC (CONICET-UNC), Ciudad Universitaria, Pabellón Argentina, 5000 Córdoba, Argentina.*
[2]*Departamento de Química Teórica y Computacional, Fac. de Ciencias Químicas,*
Universidad Nacional de Córdoba, Ciudad Universitaria, Pabellón Argentina, X5000HUA Córdoba, Argentina.
[3]*Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile*

(Dated: February 2, 2021)

Harnessing the unique features of topological materials for the development of a new generation of topological based devices is a challenge of paramount importance. Using Floquet scattering theory combined with atomistic models we study the interplay between laser illumination, spin and topology in a two-dimensional material with spin-orbit coupling. Starting from a topological phase, we show how laser illumination can selectively disrupt the topological edge states depending on their spin. This is manifested by the generation of pure spin photocurrents and spin-polarized charge photocurrents under linearly and circularly polarized laser-illumination, respectively. Our results open a path for the generation and control of spin-polarized photocurrents.

Introduction.– The early theoretical proposals [1–4] and the subsequent experimental realization of topological insulators [5, 6] have lined up the relentless scientific efforts of an ever growing community in Physics, Materials Science and Chemistry [7, 8]. Besides interesting features such as spin-momentum locking [9], topologically protected states are attractive because, unlike the usual electronic states in solids, they enjoy an intrinsic robustness to perturbations and disorder. But this lack of fragility opens up new challenges for their manipulation. Typical schemes such as surface functionalization [8] are quite ineffective when applied to topological insulators. The difficulty to cleanly turn-off conduction of charge and spin has motivated proposals for a topological field effect transistor [10–13] where the electric field drives a topological transition to a trivial insulating phase, a concept that has been experimentally realized recently [14].

Another stream of research has been looking to exploit light-matter interaction in materials to control their electrical properties. This includes generating effects such as dichroism [15, 16], a situation where electrons at different valleys absorb left and right-handed photons differently, which is of much relevance in the context of two-dimensional materials [17–19]. A different approach is aimed at using intense laser illumination to actually change the properties of the material [20]. Indeed, strong illumination has been demonstrated to produce hybrid electron-photon states [21, 22] (Floquet-Bloch states) which may present new topological properties [23–25] (see also Ref. [26]) and even exhibit a light-induced Hall effect [27].

Here, we study laser-illumination on graphene with spin-orbit coupling and a sublattice-symmetry breaking potential. The parameters are fixed so that, in absence of radiation, the system is in a topologically insulating phase with counter-propagating spin-polarized states protected by time-reversal symmetry. Previous related studies have focused on the rich phase diagram of Floquet topological phases under strong high-frequency radiation ($\hbar\Omega$ larger than the bandwidth) [28] and also considering resonant processes [29]. In contrast to those studies, here we focus on using light to gently disrupt the native topological states. In this regime one might expect an interesting interplay between symmetry breaking (inversion symmetry, or time-reversal symmetry which can be broken or preserved by circular or linearly polarized light), spin-orbit coupling which also intertwined the valley and spin degrees of freedom, and photon-induced processes. Specifically, we show that laser-illumination leads to: (i) pure spin currents under linearly polarized light, and (ii) spin polarized charge currents under circular polarization. In both cases the spin (i) and charge (ii) currents flow even at zero-bias voltage. We rationalize these pumping currents in terms of a selective hybridization of electron-photon states which is enriched by valley and spin-selective selection rules under circular polarization.

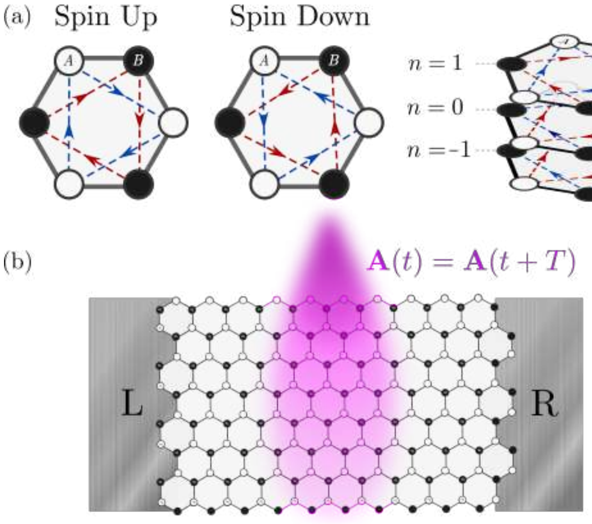


Figure 1. The irradiated Kane-Mele model. The system consists of two decoupled copies, each one representing spin up and spin down, and therefore they are time-reversal partners (a, left). Under the light spot, the system develops the replica scheme unfolding itself into several copies which represent photon dressed processes (a, right). In (b) a schematic representation of the device we will consider in the transport setup. Under particular conditions the transport of one spin might be suppressed while the remaining is *perfectly* unaffected.

In the following we present the lattice model we use to test our ideas, the basics of Floquet theory and the generalized Landauer-Büttiker formalism. Later on we present our results for the spectral and transport properties, rationalizing them in terms of a few main ingredients. Finally, we discuss possible realizations and potential application of these ideas as well as drawbacks that may arise on the realization of these concepts.

Hamiltonian model for Floquet-Kane-Mele system.– Let us now consider the Hamiltonian for graphene with a staggering potential and intrinsic spin-orbit (ISO) interaction [1]:

$$\mathcal{H}_0 = \sum_{i,s_z} E_i \, c^\dagger_{i,s_z} c_{i,s_z} - \gamma_0 \sum_{\langle i,j\rangle ,s_z} c^\dagger_{i,s_z} c_{j,s_z} - \\ -\mathrm{i}\gamma_{SO} \sum_{\langle\langle i,j\rangle\rangle ,s_z} \nu_{i,j} s_z c^\dagger_{i,s_z} c_{j,s_z}, \tag{1}$$

where $c^\dagger_{i,s_z}$ and c_{i,s_z} are the creation and annihilation operators for electrons at the π-orbital on site i with spin up $s_z = 1$ or spin down $s_z = -1$. γ_0 is the nearest-neighbors matrix element and γ_{SO} is the intrinsic spin-orbit coupling. We set $\gamma_0 = 1$ as our energy scale. $\nu_{i,j}$ is $+1$ (-1) if the path from j to i is clockwise (anticlockwise), as shown in Fig. 1a (right for spin up, left for down). The on-site energies E_i are chosen equal to Δ $(-\Delta)$ for the sites on the A (B) sublattice. The single and double brackets denote that the summation is over first or next nearest-neighbors. Although the spin-orbit coupling in bare graphene is too small, the same physics can be realized in other two-dimensional materials such as silicene and germanene [30] where this coupling is stronger.

The effect of laser illumination is captured through the Peierls' substitution [31, 32], a time dependent phase in the nearest-neighbors and next-nearest-neighbors matrix elements:

$$\gamma_{i,j}(t) = \gamma^{(0)}_{i,j} \exp\left[\mathrm{i}\,\frac{2\pi}{\Phi_0}\int_{\mathbf{r}_j}^{\mathbf{r}_i} \mathbf{A}(t)\cdot\, d\mathbf{r}\right]. \tag{2}$$

where $\gamma^{(0)}_{i,j}$ are the unperturbed matrix elements as given in Eq. 1, Φ_0 is the magnetic flux quantum and $\mathbf{A}(t)$ is the vector potential. For a monochromatic plane wave in the z-direction (perpendicular to the graphene sheet) we consider $\mathbf{A}(t) = A_0\cos(\Omega t)\hat{x} + A_0\sin(\Omega t + \phi)\hat{y}$, where Ω is the radiation frequency, A_0 determines the driving amplitude and $\phi = 0, \pm\pi/2$ controls the polarization linear or left/right hand polarized, respectively. Right hand polarization is considered whenever we mention circular polarization. The laser strength can be characterized by the dimensionless parameter $z = A_0 a 2\pi/\Phi_0$.

Similar systems were considered before with a few differences: Ref. [33] studied laser-illuminated transition metal dichalcogenide without spin-orbit coupling considered here, and Refs. [29, 34] studied germanene and silicene in the high-frequency regime while here we focus on frequencies smaller than the bandwidth. Other studies using Floquet theory focused on the topological states induced by light [35–37], rather than the modification of native topological states considered here.

Floquet theory for the spectral and transport properties.– Floquet theory allows for a non-perturbative and non-adiabatic solution of problems involving a time-periodic Hamiltonian such as ours satisfying $\mathcal{H}(t) = \mathcal{H}(t+T)$ with $T = 2\pi/\Omega$. The Floquet theorem assures that there is a complete set of solutions of the form $|\psi_\alpha(t)\rangle = \exp(-i\varepsilon_\alpha t)\,|\phi_\alpha(t)\rangle$, where ε_α are the quasienergies and $|\phi_\alpha(t+T)\rangle = |\phi_\alpha(t)\rangle$ are the Floquet states obeying:

$$\mathcal{H}_F\,|\phi_\alpha(t)\rangle = \varepsilon_\alpha\,|\phi_\alpha(t)\rangle\,, \tag{3}$$

where $\mathcal{H}_F \equiv \mathcal{H} - i\partial/\partial t$ is the Floquet Hamiltonian. Thus, one gets an eigenvalue problem in the direct product space (Floquet or Sambe space [38]) $\mathcal{R}\otimes\mathcal{T}$ where $\mathcal{R}$ is the usual Hilbert space and $\mathcal{T}$ the space of $T-$periodic functions spanned by $\langle t\,|\,n\rangle = \exp\left(in\Omega t\right)$. The index n is often called the *replica* index and can be associated to different photon channels. In this picture, an electron entering lead α in a given reference replica n_0 can scatter to lead β in replica m, thus exchanging $n = m - n_0$ quanta of the time-dependent modulation. Since asymptotically the different replicas are decoupled (as the time-modulation is limited to the sample), the total transmission probability from α to β, $\mathcal{T}_{\beta,\alpha}(\varepsilon)$, can be obtained by summing the probabilities associated to each of these processes (denoted with $\mathcal{T}^{(n)}_{\beta,\alpha}(\varepsilon)$):

$$\mathcal{T}_{\beta,\alpha}(\varepsilon) = \sum_n \mathcal{T}^{(n)}_{\beta,\alpha}(\varepsilon), \tag{4}$$

These probabilities can be computed using standard Green's functions techniques [32, 39, 40]. In a two-terminal setup in the lineal response regime (small bias voltage), the time-averaged current is given by [41]:

$$\bar{\mathcal{I}} \simeq \frac{2e^2}{h}\mathcal{T}_+(\varepsilon_F)V + \frac{4e}{h}\int \mathcal{T}_-(\varepsilon)f(\varepsilon)d\varepsilon. \tag{5}$$

where $\mathcal{T}_\pm = (\mathcal{T}_{R,L} \pm \mathcal{T}_{L,R})/2$, $f = f_L + f_R$ is the sum of the Fermi distribution functions at each lead (L or R) and V is the bias voltage. The second term on the right-hand side corresponds to a *pumping current* arising because of the asymmetry of the transmission coefficients. This current remains even at zero bias voltage (and it may even flow against an applied voltage). Such currents have been studied extensively in the past [42–45] and more recently have applied to the context of the shift photocurrents problem [46]. As we will see later on, in our case these pumping currents can be tuned by the interplay between the native topological states of the system, the laser polarization and photon assisted processes.

Quasi-energy dispersion. Let us start our discussion by analyzing the dispersion relations for a ribbon of laser-illuminated graphene with spin-orbit coupling and a staggering potential. This is shown in Fig. 2 for linear (c-d) and circular (e-f) polarization and also without radiation (a-b). Without

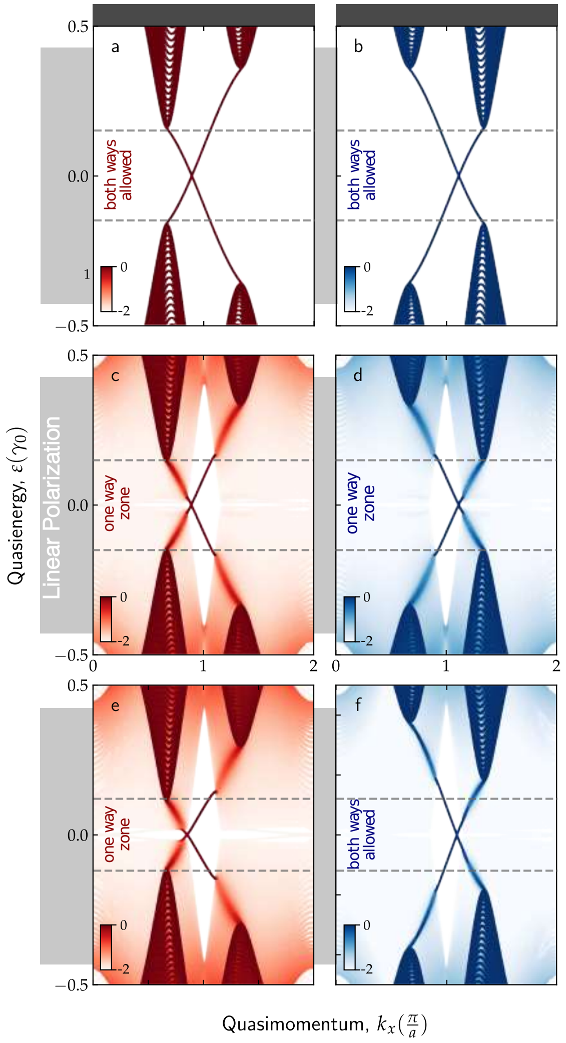


Figure 2. Spin resolved bandstructure of a honeycomb ribbon with intrinsic spin-orbit coupling. Computation performed over a zigzag nanoribbon of width $W = 100a$ ($\sim$ 25 nm). Panels (a-b) without irradiation. Panels (c-d) and (e-f) for linear and circular polarized irradiation, respectively. Red(blue) denote spin up(down). We consider $\Delta = 0.1$, $\gamma_{\rm SO} = 0.05$, $\hbar\Omega = 1.5$ and $z_x = z_y = 0.15$. One or both ways allowed transport is highlighted for edge states bridge the energy gap. The color scale in the bottom shows the time-averaged density of states in log-scale.

radiation, when the spin-orbit term dominates over the staggering one has the expected topological states bridging the gap. The staggering is responsible for the asymmetry between the valleys, while the overall time-reversal symmetry enforces the mirror symmetry between the plots (when exchanging k by $-k$) for the different spin-components. The color scale encodes the contribution of each state to the time-averaged density of states [23] which is given by the weight of each state on the reference replica ($n = 0$), which is uniform and equal to unity in absence of radiation. For linear and circular polarization, the lighter tones (notice the log scale) show the regions with states due to the other replicas (each shifted by $\hbar\Omega$). Radiation will introduce a coupling between the replicas (or, in other words, a coupling between a state with a given k at energy ε and other states at the same k with energy $\varepsilon + n\hbar\Omega$). The effect of such coupling is the hybridization of the native topological states of this system with the continuum provided by the replicas. In the figure this is evidenced as regions where the lines bridging the gap become blurred (the log scale emphasizes these regions which in normal scale will be hardly visible). Later on, we will see how transport is disrupted due to this hybridization.

Notice that the hybridization with the continuum appearing here is different from that studied in Refs. [47–49] where the continuum is provided by the states of a second layer in bilayer graphene. In contrast, here this is due to coupling with the continuum in other replicas through photon-assisted processes. Furthermore, spin plays a crucial role in the selection rules as we will highlight later on.

When comparing the results for linear and circular polarization in Fig. 2 we find an interesting asymmetry: while with linearly polarized light time reversal symmetry is preserved, circular polarization breaks it. The panels for spin up and down in Fig. 2e-f do not mirror each other as when TRS is preserved (panels c-d) and the response is thus expected to be spin-selective. As we will see next, this leads to a deeper selection rule tied to a spin-dependent circular-dichroism effect.

Transport properties.– Let us now turn to the transport properties. We consider a two-terminal setup where a central region is being illuminated while the leads remain in equilibrium. All the Hamiltonian parameters of the scattering zone and leads are equal. By using Floquet scattering theory as mentioned earlier, we compute the total transmission probabilities as a function of the energy of the incident electrons (ε). Furthermore, we can resolve the contributions of both spin components as shown in Fig. 3a-d (readers can find a detailed comparison between Floquet bandstructure and transport signatures in the supporting information). While without laser illumination one would expect a perfect and reciprocal transmission equal to unity for energies within the bulk gap, here we see a different picture. First, the left-to-right and right-to-left transmission probabilities differ, as is usual in driven systems with broken symmetries. But interestingly, the response is also highly sensitive to the spin component for circularly polarized light: the results show that one spin component gets stronger scattering (deviations from unity) while the other is less affected. This startling difference in the response for different spins (this is, the difference between Figs. 3c-d) begs for an explanation.

The results for the current, which we can resolve in its spin components are shown in Fig. 3 e (linear polarization and f (circular polarization). A first fact advanced earlier is that because of the non-reciprocity, there is a photo-generated current that appears even at zero bias voltage. This type of pumped current [42] or photocurrent in this case [46] is intertwined with the symmetry breaking induced by the different terms in the Hamiltonian. While inversion symmetry is broken in all cases (due to the staggering term), for linear polarization TRS

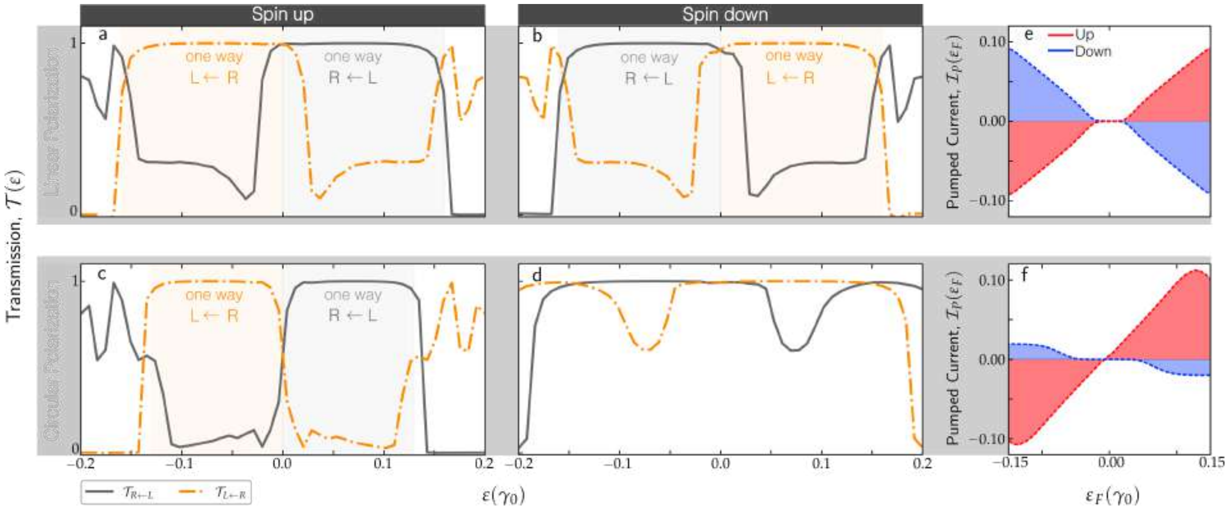


Figure 3. Effective transport properties for the scattering process averaged over one irradiation period. Parameters are the same than in Figure 2. Panels (a-d) show the transmission probabilities within the native gap. One-way charge transport is achieved in (a), (b) and (c) while in (d) no one-way effect is witnessed. Remarkably, the difference between both polarizations can be understood in terms of the circular dichroism on each copy of Haldane model. Panels (e-f) present the spin-resolved pump currents obtained for linear and circular polarization. Due to the presence of pumping currents in Floquet context, this behavior translates into a zero charge pumping with linear polarization yielding pure spin currents (e) while with circular polarizations (f) one can achieve spin polarized charge currents. The former can be tweaked by changing from right-hand to left-hand circular polarization.

is preserved and hence no net charge current can flow in this spinful case. The current per spin component is non-vanishing as shown in Fig. 3e, and together both components give a pure spin current.

In contrast, for circular polarization there is a non-vanishing current which turns out to be spin-polarized (see Fig. 3f). The polarization depends on the Fermi energy, being almost perfect close to the charge neutrality point and of about 83% at higher/lower energies. The spin-selective non-reciprocity (Figs. 3c-d) under circular polarization together with the spin-polarized photocurrents (Fig. 3f) are the main numerical results of this paper. Notice that the spin polarization can be inverted by changing the handedness of the laser polarization.

To rationalize the transport results and the quasienergy dispersions we now discuss several points that altogether explain our findings. But first, we need to dig deeper by presenting the different contributions to the total transmissions shown in Fig. 3. Indeed, an electron entering the illuminated sample with energy ε can exit elastically (without emitting or absorbing a net number of photons) or inelastically. The partial transmissions $\mathcal{T}_{i,j}^{(n)}$ for linear and circular polarization are shown in Fig. 4 and a discussion of the role of the inelastic back scattering can be found in the supplementary information. The insets of Fig. 4 show the transition matrix elements between the unperturbed initial and final states. These insets confirm our previous observation that the propagating state with spin down traversing the device is much less disturbed by circularly polarized light than the other.

The following general points explain the observed numerical features:

1. *Generalized symmetry in Floquet space.* The following relation among transmission probabilities is verified in our case:

$$\mathcal{T}_{\beta,\alpha}^{(n)}(\varepsilon) = \mathcal{T}_{\alpha,\beta}^{(-n)}(-\varepsilon). \tag{6}$$

This is enforced by an underlying symmetry in Floquet-space:

$$\Gamma\mathcal{H}(\mathbf{k})\Gamma^\dagger = -\mathcal{H}(\mathbf{k})^*, \tag{7}$$

where $\Gamma = \sigma_y K$, K being the complex conjugation operator. This generalized symmetry thus inverts the energy sign while mirroring the space and replica coordinates. This symmetry, which is fulfilled in our device setup, therefore links the transmission probabilities in the different panels of Figs. 3 and 4, which also serves as a numerical test of our results. We note that the similar relations have also been used in a different context in Floquet systems [50].

2. *Spin-selective dichroism effect: a selection rule linking Chern number, circular polarization handedness and spin.* Under illumination with circularly polarized light we observe a marked transport asymmetry between spin components, and also within the same spin subspace when the handedness of the laser polarization changes. The latter is commonly referred to as circular dichroism. The existence of circular dichroism in the presence of both a complex next-nearest-neighbor coupling and a staggering potential for the bulk states has been discussed in Ref. [51]. In that reference, Ghalamkari and coauthors find that there is a selection rule which ties the Chern number of the topological phases found

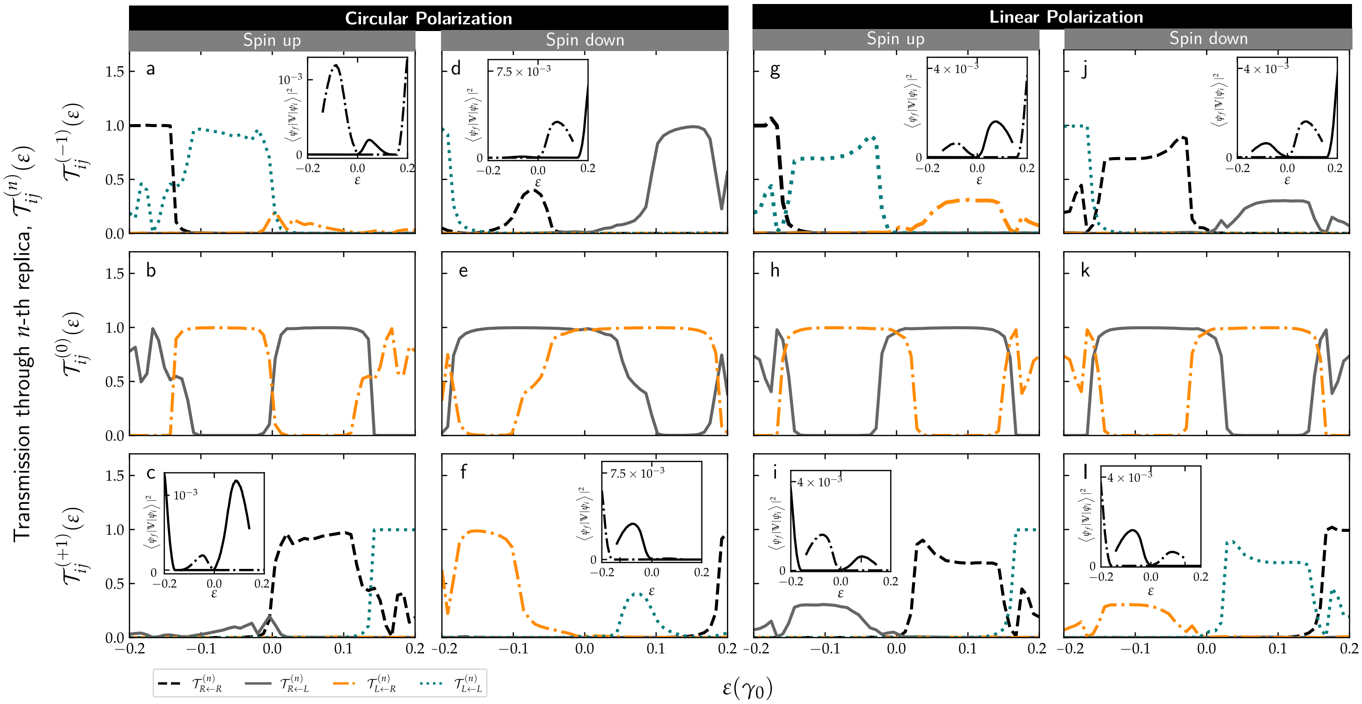


Figure 4. Detailed transmission probabilities averaged in time. Here, $\mathcal{T}_{ij}^{(n)}(\varepsilon)$ stands for the transmission through a channel mediated by an exchange of n photons with an electron incident with quasienergy ϵ. From panels (a) to (f) circular polarization transport is shown, describing the whole process for each spin channel independently. From panel (g) to (l) the same information is depicted for linear polarization. Elastic channels prove to be the main source of transmission, while reflection processes, leading to a one-way transport in regions within the bulk gap, are completely mediated by photon-dressed processes, a distinctive signature of hybridization of the states with the continuum. Insets on each panel quantify the degree of coupling of a native topological state and the bulk bands states induced by higher order replicas. In the insets, solid and dashed lines represent opposite edge chiral states.

in the Haldane model [52] with a distinctive response to left and right circular polarization.

In our case, when looking at each spin component separately, our numerical results show that this selection rule persists for a finite system, when the transitions include an edge state and a state in the bulk spectrum. Furthermore, the fact that both spin components are related by time-reversal in $\mathcal{H}_0$ produces an inversion of the circular dichroism when passing from spin up to spin down, since it follows the sign inversion of the spin-resolved Chern number. Here we observe that this inversion of the circular dichroism is also fulfilled for the finite system, a fact which one might intuitively tie in with the bulk-boundary correspondence.

3. *Hybridization of edge states with the continuum provided by a different Floquet replica.* The selection rule stated in point 2 plays a crucial role in establishing the possible light-induced transitions among the electronic states. For Fermi energies within the bulk gap of the sample, thanks to photon-assisted processes, the topological edge states at ε can now transition towards the continuum of states at $\varepsilon + n\hbar\Omega$ (where the system is ungapped). Based on point 2, this hybridization with the states of a different replica is expected to be insensitive to the spin for linear polarization but not for circular polarization. This is verified by numerically computing the modulus squared of the matrix element of the perturbation among the initial and final states, see insets of Fig. 4.

For our numerical simulations we employed a general model with staggering potential and spin-orbit coupling compatible with Germanene and Stanene, but we notice that the interplay between the staggering strength, the spin-orbit coupling and the laser frequency allows for a broad range of materials where the predicted photocurrents could be observed. Indeed, we require a system with with broken inversion symmetry hosting topological states. Within the Kane-Mele model this means that $2\Delta/\gamma_{SO} < 3\sqrt{3}$ [1, 52]. Furthermore, for the hybridization of the topological states with the continuum of the Floquet replicas to occur, the photon energy needs to be not smaller than the bulk gap of the non-irradiated system and not so large so that there are no continuum states $\hbar\Omega$ above a given energy. Fortunately, these conditions do not impose a restriction within the experimentally relevant regime of laser frequencies spanning from the mid-infrared to the visible range (see supplementary information). On the other hand, the temperature needed for the experimental realization and the fine-tuning of the irradiation condition will depend on the energy gap of the unperturbed material. The required laser intensities are smaller than those required to observe Floquet-Bloch states, as here we need sufficient coupling with a continuum of states. From our numerics, for typical mid-infrared wavelengths (~ 160 meV) we estimate that intensities in the range of $1-10\ \mathrm{mW}/\mu\mathrm{m}^2$ would suffice.

Let us now discuss the influence of Rashba spin-orbit coupling. A Rashba term introduces spin-flip processes and it is a legitimate source of concern. Our numerical results evidence robustness of the photocurrents against this term (see supplementary information). This is because the mechanism relies on the existence of topologically protected states in the non-irradiated material, which are originated by intrinsic SOC and which extend to a region of parameters with moderate values of Rashba SOC. Indeed, the topological states are robust against a moderately strong Rashba SOC [1]: for $\gamma_R < 2\sqrt{e}\gamma_{SO}$ (γ_R and γ_{SO} being the strengths of the Rashba and intrinsic spin-orbit coupling terms) the resulting phase diagram is adiabatically connected to the quantum spin-hall phase of the Kane-Mele model. This topological protection is expressed in the fact that the matrix element between two topological counter-propagating states at one edge of any perturbation that preserves TRS is zero. In our case, circular polarization does not preserve TRS but rather than introducing a matrix element between counter-propagating edge states, leads to a selective hybridization of the edge states with the continuum at an energy differing by the photon energy from them. This is why our results are robust against spin non-conserving terms over a broad range of parameters.

Final remarks.– Using Floquet scattering theory we show how laser illumination can selectively disrupt the edge states of a two-dimensional topological insulator depending on their spin. This selectivity, which stems from the interplay between a spin-selective selection rule together with the hybridization of the edge states with the continuum of another Floquet replica, manifests by the generation of pure spin currents and spin-polarized charge photocurrents under linearly and circularly polarized laser-illumination, respectively. We emphasize that, in both cases, the spin and charge currents flow even at zero-bias voltage. Furthermore, the direction and spin polarization of these currents can be tuned by changing the incident electronic energy and the handedness of light polarization, thereby providing an experimental handle to control photocurrents. In this sense, given the generality of our model, we expect for the photocurrents predicted here to be experimentally accessible in two-dimensional materials by using laser illumination in the mid-infrared.

E.A.R.-M. acknowledges support by the ANID (Chile)PFCHA, DOCTORADO NACIONAL/2017, under ContractNo. 21171229. We thank the support of FondeCyT (Chile) under grant number 1170917, and by the EU Horizon 2020 research and innovation program under the Marie-Sklodowska-Curie Grant Agreement No. 873028 (HYDROTRONICS Project). L. E. F. F. T. also acknowledges the support of The Abdus Salam International Centre for Theoretical Physics and the Simons Foundation.

---

[*] These two authors contributed equally.

[1] C. L. Kane and E. J. Mele, "Quantum Spin Hall Effect in Graphene," Phys. Rev. Lett. **95**, 226801 (2005).

[2] C. L. Kane and E. J. Mele, "${Z}_{2}$ Topological Order and the Quantum Spin Hall Effect," Physical Review Letters **95**, 146802 (2005).

[3] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, "Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells," Science **314**, 1757 (2006).

[4] L. Fu and C. L. Kane, "Topological insulators with inversion symmetry," Phys. Rev. B **76**, 045302 (2007).

[5] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, "Quantum Spin Hall Insulator State in HgTe Quantum Wells," Science **318**, 766 (2007).

[6] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, "A topological Dirac insulator in a quantum spin Hall phase," Nature **452**, 970 (2008).

[7] Y. Ren, Z. Qiao, and Q. Niu, "Topological phases in two-dimensional materials: a review," Reports on Progress in Physics **79**, 066501 (2016).

[8] D. Kong and Y. Cui, "Opportunities in chemistry and materials science for topological insulators and their nanostructures," Nature Chemistry **3**, 845 (2011).

[9] F. Ortmann, S. Roche, S. O. Valenzuela, and L. W. Molenkamp, eds., *Topological Insulators: Fundamentals and Perspectives* (Wiley, 2015).

[10] X. Qian, J. Liu, L. Fu, and J. Li, "Quantum spin Hall effect in two-dimensional transition metal dichalcogenides," Science **346**, 1344 (2014).

[11] J. Liu, T. H. Hsieh, P. Wei, W. Duan, J. Moodera, and L. Fu, "Spin-filtered edge states with an electrically tunable gap in a two-dimensional topological crystalline insulator," Nature Materials **13**, 178 (2014).

[12] H. Pan, M. Wu, Y. Liu, and S. A. Yang, "Electric control of topological phase transitions in Dirac semimetal thin films," Scientific Reports **5**, 14639 (2015).

[13] W. G. Vandenberghe and M. V. Fischetti, "Imperfect two-dimensional topological insulator field-effect transistors," Nature Communications **8**, 14184 (2017).

[14] J. L. Collins, A. Tadich, W. Wu, L. C. Gomes, J. N. B. Rodrigues, C. Liu, J. Hellerstedt, H. Ryu, S. Tang, S.-K. Mo, S. Adam, S. A. Yang, M. S. Fuhrer, and M. T. Edmonds, "Electric-field-tuned topological phase transition in ultrathin Na 3 Bi," Nature **564**, 390 (2018).

[15] D. Xiao, W. Yao, and Q. Niu, "Valley-Contrasting Physics in Graphene: Magnetic Moment and Topological Transport," Physical Review Letters **99**, 236809 (2007).

[16] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shi, Q. Niu, P. Tan, E. Wang, B. Liu, and J. Feng, "Valley-selective circular dichroism of monolayer molybdenum disulphide," Nature Communications **3**, 887 (2012).

[17] E. J. Sie, J. W. McIver, Y.-H. Lee, L. Fu, J. Kong, and N. Gedik, "Valley-selective optical Stark effect in monolayer WS_2," Nature Materials **14**, 290 (2015).

[18] H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, "Valley polarization in MoS 2 monolayers by optical pumping," Nature Nanotechnology **7**, 490 (2012).

[19] L. Zhang, K. Gong, J. Chen, L. Liu, Y. Zhu, D. Xiao, and H. Guo, "Generation and transport of valley-polarized current in transition-metal dichalcogenides," Physical Review B **90**,

195428 (2014).

[20] "Quantum phases on demand," Nature Physics **16**, 1 (2020).

[21] Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik, "Observation of Floquet-Bloch States on the Surface of a Topological Insulator," Science **342**, 453 (2013).

[22] F. Mahmood, C.-K. Chan, Z. Alpichshev, D. Gardner, Y. Lee, P. A. Lee, and N. Gedik, "Selective scattering between Floquet-Bloch and Volkov states in a topological insulator," Nature Physics **12**, 306 (2016).

[23] T. Oka and H. Aoki, "Photovoltaic Hall effect in graphene," Phys. Rev. B **79**, 081406 (2009).

[24] N. H. Lindner, G. Refael, and V. Galitski, "Floquet topological insulator in semiconductor quantum wells," Nat Phys **7**, 490 (2011).

[25] M. S. Rudner and N. H. Lindner, "Band structure engineering and non-equilibrium dynamics in Floquet topological insulators," Nature Reviews Physics **2**, 229 (2020).

[26] F. Giustino, M. Bibes, J. H. Lee, F. Trier, R. Valentí, S. M. Winter, Y.-W. Son, L. Taillefer, C. Heil, A. I. Figueroa, B. Plaçais, Q. Wu, O. V. Yazyev, E. P. A. M. Bakkers, J. Nygård, P. Forn-Díaz, S. de Franceschi, L. E. F. Foa Torres, J. McIver, A. Kumar, T. Low, R. Galceran, S. O. Valenzuela, M. V. Costache, A. Manchon, E.-A. Kim, G. R. Schleder, A. Fazzio, and S. Roche, "The 2020 Quantum Materials Roadmap," Journal of Physics: Materials (2020), 10.1088/2515-7639/abb74e.

[27] J. W. McIver, B. Schulte, F.-U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, "Light-induced anomalous Hall effect in graphene," Nature Physics **16**, 38 (2020).

[28] M. Ezawa, "Photoinduced Topological Phase Transition and a Single Dirac-Cone State in Silicene," Physical Review Letters **110**, 026603 (2013).

[29] A. López, A. Scholz, B. Santos, and J. Schliemann, "Photoinduced pseudospin effects in silicene beyond the off-resonant condition," Phys. Rev. B **91**, 125105 (2015).

[30] M. Ezawa, "Monolayer Topological Insulators: Silicene, Germanene, and Stanene," Journal of the Physical Society of Japan **84**, 121003 (2015).

[31] H. L. Calvo, H. M. Pastawski, S. Roche, and L. E. F. Foa Torres, "Tuning laser-induced band gaps in graphene," Appl. Phys. Lett. **98**, 232103 (2011).

[32] H. L. Calvo, P. M. Perez-Piskunow, H. M. Pastawski, S. Roche, and L. E. F. Foa Torres, "Non-perturbative effects of laser illumination on the electrical properties of graphene nanoribbons," Journal of Physics: Condensed Matter **25**, 144202 (2013).

[33] A. Huamán and G. Usaj, "Floquet spectrum and two-terminal conductance of a transition-metal dichalcogenide ribbon under a circularly polarized laser field," Physical Review B **99**, 075423 (2019).

[34] M. Tahir, Q. Y. Zhang, and U. Schwingenschlögl, "Floquet edge states in germanene nanoribbons," Scientific Reports **6**, 31821 (2016).

[35] U. Bajpai, M. J. H. Ku, and B. K. Nikolic, "How robust is quantized transport through edge states of finite length: Imaging current density in Floquet topological vs. quantum spin and anomalous Hall insulators," (2020).

[36] A. Farrell and T. Pereg-Barnea, "Photon-Inhibited Topological Transport in Quantum Well Heterostructures," Physical Review Letters **115**, 106403 (2015).

[37] S. A. Sato, P. Tang, M. A. Sentef, U. De Giovannini, H. Hübener, and A. Rubio, "Light-induced anomalous Hall effect in massless Dirac fermion systems and topological insulators with dissipation," arXiv:1905.12981 [cond-mat, physics:physics] (2019), arXiv: 1905.12981.

[38] H. Sambe, "Steady States and Quasienergies of a Quantum-Mechanical System in an Oscillating Field," Phys. Rev. A **7**, 2203 (1973).

[39] G. Stefanucci, S. Kurth, A. Rubio, and E. K. U. Gross, "Time-dependent approach to electron pumping in open quantum systems," Phys. Rev. B **77**, 075339 (2008).

[40] L. Arrachea and M. Moskalets, "Relation between scattering-matrix and Keldysh formalisms for quantum transport driven by time-periodic fields," Phys. Rev. B **74**, 245322 (2006).

[41] L. E. F. Foa Torres, P. M. Perez-Piskunow, C. A. Balseiro, and G. Usaj, "Multiterminal Conductance of a Floquet Topological Insulator," Phys. Rev. Lett. **113**, 266801 (2014).

[42] M. Moskalets and M. Büttiker, "Floquet scattering theory of quantum pumps," Phys. Rev. B **66**, 205320 (2002).

[43] B. L. Altshuler and L. I. Glazman, "Pumping Electrons," Science **283**, 1864 (1999).

[44] F. Giazotto, P. Spathis, S. Roddaro, S. Biswas, F. Taddei, M. Governale, and L. Sorba, "A Josephson quantum electron pump," Nature Physics **7**, 857 (2011).

[45] B. Kaestner and V. Kashcheyevs, "Non-adiabatic quantized charge pumping with tunable-barrier quantum dots: a review of current progress," Reports on Progress in Physics **78**, 103901 (2015).

[46] U. Bajpai, B. S. Popescu, P. Plecháč, B. K. Nikolić, L. E. F. F. Torres, H. Ishizuka, and N. Nagaosa, "Spatio-temporal dynamics of shift current quantum pumping by femtosecond light pulse," Journal of Physics: Materials **2**, 025004 (2019).

[47] L. E. F. Foa Torres, V. Dal Lago, and E. Suárez Morell, "Crafting zero-bias one-way transport of charge and spin," Physical Review B **93**, 075438 (2016).

[48] V. Dal Lago, E. Suárez Morell, and L. E. F. Foa Torres, "One-way transport in laser-illuminated bilayer graphene: A Floquet isolator," Physical Review B **96**, 235409 (2017).

[49] M. Berdakin, J. E. B. Vargas, and L. E. F. Foa Torres, "Directional control of charge and valley currents in a graphene-based device," Physical Chemistry Chemical Physics **20**, 28720 (2018).

[50] O. Balabanov and H. Johannesson, "Transport signatures of symmetry protection in 1D Floquet topological insulators," J. Condens. Matter Phys. **32**, 015503 (2019),.

[51] K. Ghalamkari, Y. Tatsumi, and R. Saito, "Perfect Circular Dichroism in the Haldane Model," Journal of the Physical Society of Japan **87**, 063708 (2018).

[52] F. D. M. Haldane, "Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly"," Phys. Rev. Lett. **61**, 2015 (1988).

Control of electron-electron interaction in graphene by proximity screening

M. Kim[1], S. G. Xu[1,2], A. I. Berdyugin[1], A. Principi[1], S. Slizovskiy[1,2,3], N. Xin[1,2], P. Kumaravadivel[1,2], W. Kuang[1], M. Hamer[2], R. Krishna Kumar[1], R. V. Gorbachev[2], K. Watanabe[4], T. Taniguchi[4], I. V. Grigorieva[1], V. I. Fal'ko[1,2], M. Polini[5,1,6], A. K. Geim[1,2]

[1]School of Physics and Astronomy, University of Manchester, Manchester, M13 9PL, United Kingdom
[2]National Graphene Institute, University of Manchester, Manchester, M13 9PL, United Kingdom
[3]Saint-Petersburg INP, Gatchina 188300, Russia
[4]National Institute for Materials Science, Tsukuba, 305-0044, Japan
[5]Dipartimento di Fisica dell'Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy
[6]Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, 16163 Genova, Italy

Electron-electron interactions play a critical role in many condensed matter phenomena[1-4], and it is tempting to find a way to control them by changing the interactions' strength. One possible approach is to place a studied electronic system in proximity of a metal, which induces additional screening and hence suppresses electron interactions. Here, using devices with atomically-thin gate dielectrics and atomically-flat metallic gates, we measure the electron-electron scattering length ℓ_{ee} in graphene at different concentrations n and temperatures. The proximity screening is found to enhance ℓ_{ee} and change qualitatively its n dependence. Counterintuitively, the screening becomes important only at gate dielectric thicknesses of a few nm, much smaller than the average separation $D \approx 1/\sqrt{n}$ between electrons. The critical thickness is given by $\sim 0.03\ \varepsilon D$, where ε is the gate dielectric's permittivity, and the theoretical expression agrees well with our experiment. The work shows that, using van der Waals heterostructures with ultra-thin dielectrics, it is possible to modify many-body phenomena in adjacent electronic systems.

Elementary electrostatics tells us that the electron charge e placed at the distance d from a bulk metal leads to a dipole potential evolving as $2ed^2/r^3$ at large in-plane distances $r \gg d$, which is much weaker than the original, unscreened Coulomb potential, e/r. Accordingly, a metallic gate placed sufficiently close to another electronic system can alter its electron-electron (e-e) interactions. Electrostatic screening by metallic gates has previously been employed to suppress charge inhomogeneity in graphene[5,6], alter its plasmon spectra[7,8] and renormalize an electronic structure of monolayer semiconductors[9]. In principle, proximity-gate screening may also affect e-e interactions. They can be parametrized by ℓ_{ee} and, a priori, it is unclear how close a metallic gate should be to change this parameter appreciably. From the above electrostatic considerations, one can infer that what matters most is the ratio d/D. For a two-dimensional (2D) electron system with typical $n = 10^{12}\ \mathrm{cm}^{-2}$, $D \approx 10\ \mathrm{nm}$ and, therefore, the inferred gate separation $d \approx D$ is relatively easy to achieve experimentally. However, as shown below, the naïve expectations fail because of a small numerical factor δ such that e-e interactions for massless Dirac fermions are altered only if $d \leq \delta D \approx 0.03\ \varepsilon D$. For typical gate dielectrics with $\varepsilon < 5$, the required separation falls into a $1\ \mathrm{nm}$ range. For massive charge carriers such as those in bilayer graphene and 2D semiconductors, even smaller (atomic-scale) d are necessary for efficient screening (Methods). It seems impossible to realize such small d because of inevitable surface roughness of the metal and insulating films used for gating and electrical leakage through dielectrics of nanometer thickness. In this report, we achieve the extremely challenging conditions for proximity-gate screening by using van der Waals heterostructures with atomically-thin dielectric layers and atomically-flat gates.

Our devices were graphene monolayers encapsulated between hexagonal boron nitride (hBN) crystals whereas graphite monocrystals served as a bottom gate (Fig. 1). These heterostructures were fabricated using the standard dry-transfer procedures[5] described in Methods. Multiterminal Hall bar devices with several point contacts and closely placed voltage probes (Fig. 1a) were then defined by electron-beam lithography and plasma etching. An extra metal gate was deposited on top of the heterostructures, which allowed us to vary n without applying voltages to the bottom screening gate. This was particularly important for our case of ultra-thin dielectrics to avoid their accidental breakdown and electrical leakage. The minimum thickness d for the gate dielectric (Fig. 1b) was limited to 4 hBN layers (i.e. $\sim 1.3\ \mathrm{nm}$) because thinner crystals exhibited notable electron tunneling[10]. The devices typically had low-temperature (T) mobility μ of about $10^6\ \mathrm{cm^2\ V^{-1} s^{-1}}$ and highly reproducible characteristics such that, at finite T, their longitudinal resistivity ρ was practically independent of d (Supplementary Fig. 1). This ensured that the reported behavior of ℓ_{ee} was due to changes in d rather than transport characteristics. Because graphite is a semimetal with a relatively low carrier density of $\sim 10^{19}\ \mathrm{cm^{-3}}$, we also crosschecked that our conclusions were independent of the gate material using screening gates made from other layered metals such as $Bi_2Sr_2CaCu_2O_{8+x}$ and TaS_2 (Methods; Supplementary Fig. 2).

To demonstrate that e-e interactions can be tuned by proximity-gate screening, a reliable diagnostic tool is essential. Many quantum transport characteristics are known to be affected by the strength of e-e interactions. For example, the phase breaking length depends on it and can be measured in quantum interference experiments[11] (other possibilities are discussed in ref. 12). In principle, it should be possible to use such 'mesoscopic physics' tools to probe e-e interactions in graphene but, because of its ballistic transport at micrometer-scale distances, the approach is not easy to implement in practice and its results could be difficult to interpret. On the other hand, recent experiments have shown that graphene at finite T and away from the charge neutrality point (NP) exhibits pronounced hydrodynamic effects[13-17], which allowed measurements of the kinematic electron viscosity ν_0, and the extracted values of $\ell_{\mathrm{ee}} = 4\nu_0/v_{\mathrm{F}}$ were in quantitative agreement with theory (v_{F} is the Fermi velocity). The viscosity measurements can be carried out using three complementary approaches: vicinity resistance[14,15], point contact geometry[16,18] and the viscous Hall effect[17]. Below we use all three to show that ℓ_{ee} changes with d. In another approach, we demonstrate that umklapp e-e scattering in graphene superlattices[19] is also affected by proximity-gate screening.

First, let us demonstrate the screening effect qualitatively. Figure 1c shows that the vicinity resistance R_{V} is notably affected if a thin gate dielectric is employed. Vicinity measurements are discussed in detail in ref. 14 but, briefly, an electric current is injected through a narrow contact into a wide graphene channel. The negative voltage drop arising locally from a viscous electron flow is detected using a vicinity contact at a short distance L from the current-injecting contact (Fig. 1a). One can see from Fig. 1c that, as T increases, R_{V} first decreases and then becomes negative. This indicates a transition from the ballistic transport regime (positive R_{V}) into a regime where ballistics is strongly affected by e-e scattering[15]. The minimum in $R_{\mathrm{V}}(T)$ corresponds to the condition $\ell_{\mathrm{ee}} \approx L$ and indicates an onset of hydrodynamic behavior[15]. As ℓ_{ee} decreases further with increasing T, R_{V}

becomes less negative and eventually positive, being dominated by currents caused by electron-phonon scattering[14,15]. The dependences $R_{\mathrm{V}}(T)$ shown in Fig. 1c were measured for two similar devices at the same L. One had $d \approx 300\ \mathrm{nm}$ (conventional Si back gate) whereas the other was made using 4-layer hBN as the gate dielectric. Despite the similar behavior of $R_{\mathrm{V}}(T)$, the curve for $d \approx 1.3\ \mathrm{nm}$ is clearly shifted to higher T. The shift direction indicates that the nearby gate caused an increase in ℓ_{ee}, which is equivalent to a reduction in electron temperature by $\sim 30\ \mathrm{K}$. Note that, for T above $100\ \mathrm{K}$ where the hydrodynamic regime develops, electron transport in high-quality graphene is universal and insensitive to experimental details.

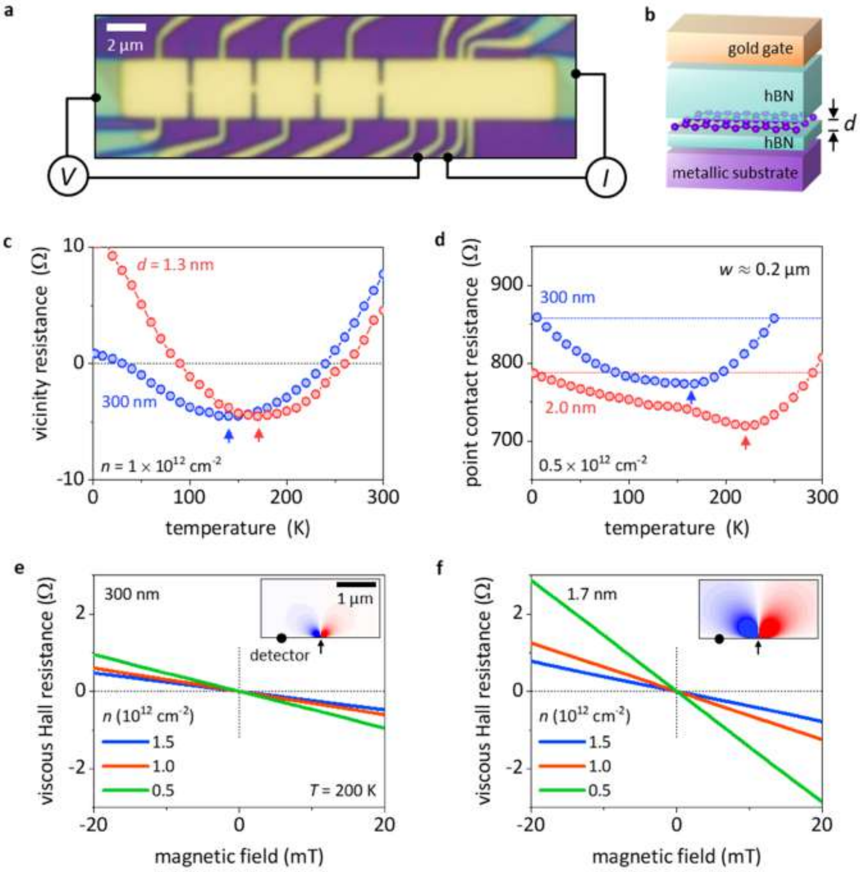


Figure 1 | Graphene devices with proximity gating and its effect on electron hydrodynamics. **a**, Optical micrograph of one of our devices with 4 sub-μm constrictions used for point-contact measurements and several closely spaced contacts for vicinity measurements. The wiring schematic illustrates current and voltage configurations for the latter measurements. **b**, Schematic side view of our heterostructures. **c**, R_{V} as a function of T for representative devices with a close graphite gate ($d \approx 1.3\ \mathrm{nm}$, red) and in the reference geometry ($d = 300\ \mathrm{nm}$, blue). The devices had similar geometry and μ; same $L = 0.5\ \mu\mathrm{m}$. **d**, $R_{\mathrm{PC}}(T)$ for screened and reference constrictions of the same width $w \approx 0.2\ \mu\mathrm{m}$ (same color coding as in **c**). Dashed lines in **d** denote the resistance in the ballistic limit. Arrows in **c** and **d** indicate minima in R_{V} and R_{PC}. **e** and **f**, Viscous Hall effect for reference and close-gate devices ($d = 300$ and $1.7\ \mathrm{nm}$, respectively). The color-coded curves correspond to different n; all measurement conditions and geometries were same, including $L = 1\ \mu\mathrm{m}$ and $T = 200\ \mathrm{K}$. The insets illustrate electric potentials that appear due to a viscous electron flow (the arrow and circle indicate positions of current and voltage contacts, respectively). The calculations[20] were carried out for the experimentally determined $\ell_{\mathrm{ee}} \approx 0.3$ and $0.8\ \mu\mathrm{m}$ for panels **e** and **f**, respectively; $B = 10\ \mathrm{mT}$. Blue-to-red color scale is arbitrary but same for both panels.

Similar phenomenology was observed in the point contact geometry (Fig. 1a). Again, the T-dependence of the point contact resistance R_{PC} exhibits a clear minimum due to a viscous flow[16]. The shift to higher T for the device with a proximity gate (Fig. 1d) indicates an increase in ℓ_{ee} for a given T. Such influence of the proximity gating was consistently observed in all our experiments. The R_{V} and R_{PC} dependences could also be used to extract $\ell_{\mathrm{ee}}(T)$ following the recipe reported in refs. 14,16. Unfortunately, we found that, for atomically-thin gate dielectrics, detailed behavior of R_{V} and, to some extent, R_{PC} notably varied between different devices with nominally the same d. Those variations can be traced back to the fact that R_{V} is sensitive to current injector's geometry[14] whereas a viscous contribution to R_{PC} becomes smaller for close-gate devices as compared to those with thicker gate dielectrics.

In contrast to the vicinity and point-contact measurements, the viscous Hall effect[17] was found to be very robust, yielding quantitatively same results for different devices with same d. Accordingly, for quantitative analysis of how ℓ_{ee} depended on d, we focused on the latter measurements. The Hall viscosity experiments utilize the already discussed vicinity geometry (Fig. 1a) but a non-quantizing magnetic field B is applied perpendicular to graphene[17]. The field leads to an asymmetry in the potential created by the viscous flow around the injection contact (insets of Figs. 1e,f). The viscous contribution asymmetric in B is called the viscous Hall resistance R_{A} and given by[17,20]

$$R_{\mathrm{A}} = \rho\xi\left(\frac{L}{\sqrt{\nu_0\tau}}\right)\frac{B}{B_0}, \qquad (1)$$

where $\xi(x)$ is a dimensionless function[20], τ is the transport scattering time, $B_0 = E_{\mathrm{F}}/(8\,|e|\,\nu_0)$ is a characteristic magnetic field, and E_{F} is the Fermi energy. Because $|\xi(x)|$ is a monotonically decreasing function of its argument for $x > 0$, $|R_{\mathrm{A}}|$ increases with increasing ℓ_{ee} and, accordingly, devices with weaker e-e scattering should exhibit larger $|R_{\mathrm{A}}|$.

To illustrate the effect of proximity-gate screening on Hall viscosity, Figs. 1e,f plot $R_{\mathrm{A}}(B)$ for two representative devices with $d \approx 1.7$ and $300\ \mathrm{nm}$. The curves are taken under exactly the same conditions for several same n. As the two devices exhibited close ρ and τ (Supplementary Fig. 1), the profound difference between Figs. 1e and 1f can only be attributed to different screening. The device with the thin dielectric exhibited much larger Hall viscosity than the reference device, and the effect was most pronounced at low n. This behavior proves again that the proximity screening suppresses e-e scattering, in agreement with the conclusions reached from the vicinity and point-contact measurements.

For the known transport characteristics (ρ and τ), Eq. 1 allows us to convert R_{A} into ℓ_{ee}, as described in detail in ref. 17. Figure 2a shows examples of $\ell_{\mathrm{ee}}(T)$ found for close-gate and reference devices. At all T, the screened device displays ℓ_{ee} approximately twice longer than that in the standard device of the same electronic quality. This agrees well with many-body theory (solid curves in Fig. 2; Supplementary Fig. 3). Importantly, the proximity-gate screening qualitatively changes the dependence $\ell_{\mathrm{ee}}(n)$ so that, away from the NP, ℓ_{ee} decreases with increasing n (Fig. 2b). This contrasts with monotonically increasing $\ell_{\mathrm{ee}}(n)$ for the reference devices, which was also reported previously[16,17]. Figure 2c summarizes our results by showing ℓ_{ee} measured for more than 10 different devices at characteristic n and T where viscous effects become most pronounced in graphene.

Despite the experimental scatter, Fig. 2c clearly shows that ℓ_{ee} can be altered appreciably by using thin gate dielectrics, if d is smaller than a few nm.

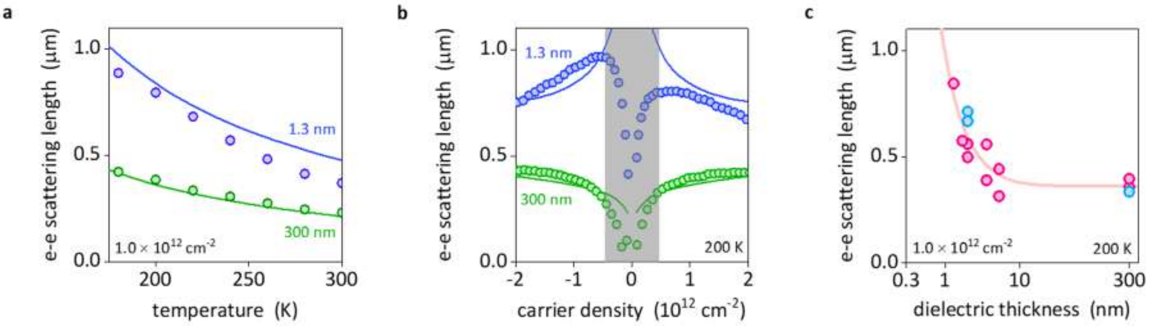


Figure 2 | Dependence of the e-e scattering length on distance to the gate. **a**, $\ell_{\mathrm{ee}}(T)$ extracted from Hall viscosity measurements for the given n. Data for a close-gate device (blue symbols) are compared with a reference (green). **b**, Density dependence of ℓ_{ee} at $200\ \mathrm{K}$ (same color coding as in **a**). The grey-shaded region indicates the regime near the NP where the single-component hydrodynamic theory is not applicable[14,15,21] and, also, the cyclotron diameter became comparable with the width of our devices[17]. **c**, ℓ_{ee} as a function of d for the given n and T. Red and blue symbols: Results from Hall viscosity and point-contact measurements, respectively; shown are the average values for electron and hole doping (see panel **b** for an example of scatter due to electron-hole asymmetry). For all the panels, the solid curves are theoretical results (Supplementary Information).

To explain the observed dependences of ℓ_{ee} on n and d, we carried out numerical calculations in the random phase approximation for the dynamically screened interactions[12,22,23]. The metallic gate was modelled as a perfect conductor, and small departures from this model caused by a finite carrier density were estimated in Supplementary Section 4. The results are shown by the solid curves in Fig. 2. No fitting parameters were used, except for multiplying all the theoretical curves by the same small factor of 1.3 (its non-Fermi-liquid origins are discussed in Supplementary Section 4). However, to gain better insight about the observed behavior, we also derived the following analytical expression

$$\ell_{\mathrm{ee}} \approx \frac{4\hbar v_{\mathrm{F}} E_{\mathrm{F}}}{\pi} \frac{1}{(k_{\mathrm{B}}T)^2 \ln\left(\frac{2E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right)} \left(\frac{1+2dq_{\mathrm{TF}}}{2dq_{\mathrm{TF}}}\right)^2, \qquad (2)$$

where $k_{\mathrm{F}} = \sqrt{\pi n}$ and $q_{\mathrm{TF}} = 4\alpha_{\mathrm{ee}} k_{\mathrm{F}}$ are the Fermi and Thomas-Fermi wavenumbers, respectively. Here, $\alpha_{\mathrm{ee}} \approx 2.2/\varepsilon$ is graphene's coupling constant and k_{B} is the Boltzmann constant (Supplementary Section 4 discusses the case of generally anisotropic ε). The expression is accurate in the Fermi-liquid regime ($k_{\mathrm{B}}T \ll E_{\mathrm{F}}$), where it matches our numerical results (Supplementary Section 4). The last term in Eq. 2 appears due to the gate presence, and the key parameter describing its screening effect is dq_{TF}. In the far-gate regime, $d \gg 1/q_{\mathrm{TF}}$, Eq. 2 reduces to the standard unscreened expression[23]. In the opposite limit, $d \ll 1/q_{\mathrm{TF}}$, e-e scattering is strongly reduced due to screening, and ℓ_{ee} increases with decreasing both d and n, as $1/d^2$ and approximately $1/\sqrt{n}$, respectively, in agreement with our experiment (Fig. 2). The latter dependence is opposite to the unscreened case, where ℓ_{ee} increases as $\sqrt{n}$, in agreement with the results of Fig. 2b. The crossover between the far- and close- gate regimes occurs at a critical distance d_{c} such that $d_{\mathrm{c}} \approx 1/2q_{\mathrm{TF}} = 1/(8\alpha_{\mathrm{ee}} k_{\mathrm{F}})$, which translates into the previously introduced parameter $\delta \approx 0.03\varepsilon$. For hBN with $\varepsilon \approx 3.5$ and at typical $n = 10^{12}\ \mathrm{cm}^{-2}$, we obtain $d_{\mathrm{c}} \approx 1.1\ \mathrm{nm}$, which explains why the gate screening becomes noticeable only for our smallest

d (Fig. 2c). Further information about our theoretical analysis is provided in Supplementary Information.

To check how robust our conclusions are, we have also examined the effect of gate-induced screening on umklapp e-e scattering[19] that dominates resistivity ρ of graphene-on-hBN superlattices at elevated T. We made several superlattice devices with the moiré periodicity $\lambda \approx 15\ \mathrm{nm}$, as confirmed by the periodicity of Brown-Zak oscillations[24] and the appearance of secondary NPs[25-28] at the expected n (Fig. 3a). One of the devices was the standard Hall bar with $d = 300\ \mathrm{nm}$, like those reported previously[19]. The other two were same in design but had a bottom graphite gate placed at short d, as in the above viscosity experiments. Figure 3 shows typical $\rho(n, T)$ measured for these graphene superlattices. For $d = 300\ \mathrm{nm}$, the observed behavior was same as reported previously, and the T dependent part ($\Delta\rho$) of graphene superlattice's resistivity could be described quantitatively by umklapp e-e scattering[19]. It is responsible for the rapid increase of $\Delta\rho \propto T^2$ (Fig. 3b). The proximity-gate screening notably suppressed $\Delta\rho(T)$, by a factor > 2 for $d \approx 1.3\ \mathrm{nm}$. Our theoretical analysis (Supplementary Section 5) shows that $\Delta\rho$ for the close-gate devices should exhibit the same T dependence ($\propto T^2$) but with a reduced absolute value. The umklapp e-e scattering length, $\ell_{\mathrm{ee}}^{\mathrm{U}}$, is governed by distinctive processes with a momentum transfer of $\sim \hbar g$ where $g = \frac{4\pi}{\sqrt{3}}\lambda^{-1}$ is the superlattice reciprocal vector. As shown in Supplementary Section 5, proximity screening for $\ell_{\mathrm{ee}}^{\mathrm{U}}$ becomes important if $d < 0.1\lambda$, which again means that few-nm-thick gate dielectrics are essential to observe the screening effect. It is convenient to quantify this effect by the dimensionless ratio, $\Delta\rho(\infty)/\ \Delta\rho(d) \approx \ell_{\mathrm{ee}}^{\mathrm{U}}(d)/\ell_{\mathrm{ee}}^{\mathrm{U}}(\infty)$. The results are plotted in the inset of Fig. 3b and show good agreement with theory (for details, see Supplementary Section 5).

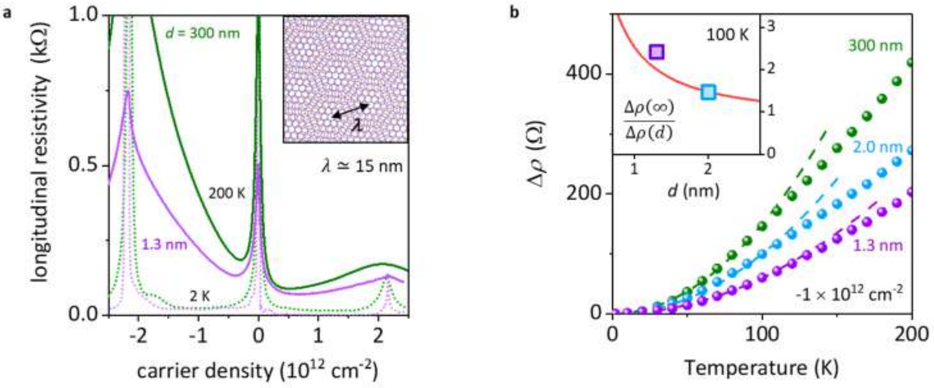


Figure 3 | Suppression of umklapp e-e scattering in graphene superlattices by proximity-gate screening. **a,** $\rho(n)$ of graphene-on-hBN superlattices for $d \approx 1.3$ and $300\ \mathrm{nm}$ (purple and green curves, respectively). Dotted and solid curves: $T = 2$ and $200\ \mathrm{K}$, respectively. Inset: Illustration of the moiré pattern arising from crystallographic alignment between graphene and hBN lattices. **b,** T-dependent part of ρ for superlattice devices with different d (color-coded symbols); $n = -1 \times 10^{12}\ \mathrm{cm}^{-2}$ so that superlattices' first Brillouin zones are approximately half-filled with holes[26-28]. Dashed curves: Best fit to the predicted T^2 dependence[19]. All the devices had $\lambda \approx 15\ \mathrm{nm}$ and close ρ at 2 K. Inset: $\Delta\rho(d)$ for the two close-gate superlattices normalized by $\Delta\rho(\infty)$ measured for the reference (far-gate) superlattice. The color-coded symbols in the inset are taken from the main panel and valid for all $T \leq 120\ \mathrm{K}$ because of the T^2 dependence. Solid curve: Theory.

To conclude, e-e scattering in monolayer graphene at finite n can be strongly suppressed if a metallic gate is placed at d of $\sim 1\ \mathrm{nm}$. This "close-gate" regime has become accessible due to the use of van der Waals assembly that allows atomically sharp interfaces and ultra-thin dielectrics. It is tempting to exploit the outlined strategy to assess interaction phenomena near the NP where low n allow the condition $d \ll 1/\sqrt{n}$ to be satisfied easier but interpretation of some observations had proven difficult. Other interesting candidates are exotic phenomena driven by strong correlations (e.g., various many-body phases in twisted bilayer graphene[29,30]) and, especially, interaction effects governed by lengths longer than ℓ_{ee}. The experimental challenge to reach the close-gate regime can partially be mitigated by using high-ε dielectrics.

Methods

Device fabrication. Our heterostructures were assembled using 'stamps' made from polypropylene carbonate (PPC) as a sacrificial polymer placed on polydimethylsiloxane (PDMS). Such polymer stamps were used to pick up exfoliated thin crystals in the following sequence: top hBN (typically thicker than $30~\mathrm{nm}$), monolayer graphene and thin bottom hBN. The latter served as a gate dielectric in the final device configuration (Fig. 1b), and its thickness was determined by atomic force microscopy. The resulting hBN/graphene/hBN stack was then released onto relatively small graphite crystals with thickness of $3-10~\mathrm{nm}$, which were prepared in advance on an oxidized Si wafer. The stack was large enough to extend outside the bottom graphite region, which allowed us to make quasi-one-dimensional contacts to graphene[31] without electrically contacting the graphite gate. The metallic contacts were defined by electron-beam lithography. We first used a mixture of CHF_3 and O_2 to plasma-etch hBN/graphene and expose the required contact regions. This was followed by deposition of $2~\mathrm{nm}$ Cr/ $60~\mathrm{nm}$ Au to make Ohmic contacts to graphene. A gold top gate was then fabricated using another round of electron-beam lithography and, also, served as an etching mask for the final etching step to define the Hall bar geometry.

The devices with other metallic gates ($Bi_2Sr_2CaCu_2O_{8+x}$ and TaS_2) required fabrication in an oxygen- and moisture- free atmosphere of a glovebox[32] to avoid deterioration of the metal surfaces. Even using glovebox encapsulation, we observed a notable reduction in graphene's quality for the above gate materials, presumably because of electrical charges at the exposed surfaces (for small d, typical μ became $< 10^5~\mathrm{cm}^2~\mathrm{V}^{-1}\mathrm{s}^{-1}$ and charge inhomogeneity near the NP considerably increased). Accordingly, reliable measurements of ℓ_{ee} in this case were only possible at high $n \gtrsim 2.0 \times 10^{12}~\mathrm{cm}^{-2}$ (Supplementary Section 2). We also note that encapsulated graphene devices with the conventional gates made by metal deposition on top of a thin gate dielectric ($d < 2~\mathrm{nm}$) exhibited extremely low μ of only $\sim 10^4~\mathrm{cm}^2~\mathrm{V}^{-1}\mathrm{s}^{-1}$. Such poor electronic quality made it impossible to carry out the ℓ_{ee} measurements described in the main text.

Electrical measurements. The devices were measured in a variable temperature insert that allowed stable T between 2 and $300~\mathrm{K}$. The standard lock-in amplifier techniques were employed using excitation currents of typically $0.1-1~\mu\mathrm{A}$ at a frequency of $30.5~\mathrm{Hz}$. For measurements of Hall viscosity, we used the same vicinity geometry as shown in the schematic of Fig. 1a. The distance between injector and detector contacts was usually between 0.5 and $1.5~\mu\mathrm{m}$. The viscous Hall resistance was determined as an antisymmetric-in-B component of the vicinity resistance in fields below $\pm 30~\mathrm{mT}$. For the point-contact measurements, we employed the quasi-four-probe geometry by driving the current through the wide contacts (on the left and right in Fig. 1a) and using the leads next to the studied constrictions as voltage probes.

Proximity screening for systems with the parabolic spectrum. The close-gate condition depends on the density of states at the Fermi energy of the material one wants to control. We have studied graphene not only because of its electronic quality but also because of the low-density of states provided by its Dirac spectrum. For a 2D system with the conventional parabolic spectrum, the close-gate condition is much more difficult to achieve. In the latter case, a proximity metal gate can provide efficient screening of e-e interactions only for distances d below $\sim \varepsilon\, m_\mathrm{e} a_\mathrm{B}\, /(2\, N_\mathrm{f}\, m_\mathrm{eff})$, where $a_\mathrm{B} \approx 0.5$ Å is the Bohr radius, m_e and m_eff are the free-electron and effective masses, respectively, and N_f is the number of spin/valley flavors. Here, $\varepsilon = \varepsilon_\perp$ is the perpendicular component of the dielectric permittivity of a gate dielectric. For bilayer graphene[33,34] with $N_\mathrm{f} = 4$, $m_\mathrm{eff} \geq 0.03\, m_\mathrm{e}$ and using hBN as a dielectric ($\varepsilon_\perp \approx 3.5$), the close-gate condition requires $d < 7$ Å , which is essentially out of experimental reach.

References

1 Coleman, P. *Introduction to many-body physics*. (Cambridge University Press, 2015).

2 Keimer, B., Kivelson, S. A., Norman, M. R., Uchida, S. & Zaanen, J. From quantum matter to high-temperature superconductivity in copper oxides. *Nature* **518**, 179 (2015).

3 Keimer, B. & Moore, J. E. The physics of quantum materials. *Nat. Phys.* **13**, 1045 (2017).

4 Hartnoll, S. A., Lucas, A. & Sachdev, S. *Holographic quantum matter*. (MIT Press, 2018).

5 Yankowitz, M. *et al*. van der Waals heterostructures combining graphene and hexagonal boron nitride. *Nat. Rev. Phys.* **1**, 112-125 (2019).

6 Ponomarenko, L. A. et al. Tunable metal-insulator transition in double-layer graphene heterostructures. *Nat. Phys.* **7**, 958-961 (2011).

7 Lundeberg, M. B. *et al.* Tuning quantum nonlocal effects in graphene plasmonics. *Science* **357**, 187-191 (2017).

8 Alcaraz Iranzo, D. *et al.* Probing the ultimate plasmon confinement limits with a van der Waals heterostructure. *Science* **360**, 291-295 (2018).

9 Waldecker, L. *et al.* Rigid band shifts in two-dimensional semiconductors through external dielectric screening. *Phys. Rev. Lett.* **123**, 206403 (2019).

10 Britnell, L. *et al.* Electron tunneling through ultrathin boron nitride crystalline barriers. *Nano Lett.* **12**, 1707-1710 (2012).

11 Imry, Y. *Introduction to mesoscopic physics*. (Oxford University Press, Oxford, 1997).

12 Giuliani, G. F. & Vignale, G. Quantum theory of the electron liquid. (Cambridge University Press, Cambridge, 2005).

13 Lucas, A. & Fong, K. C. Hydrodynamics of electrons in graphene. *J. Phys. Condens. Matter* **30**, 053001 (2018).

14 Bandurin, D. A. *et al.* Negative local resistance caused by viscous electron backflow in graphene. *Science* **351**, 1055-1058 (2016).

15 Bandurin, D. A. *et al.* Fluidity onset in graphene. *Nat. Commun.* **9**, 4533 (2018).

16 Krishna Kumar, R. *et al.* Superballistic flow of viscous electron fluid through graphene constrictions. *Nat. Phys.* **13**, 1182-1185 (2017).

17 Berdyugin, A. I. *et al.* Measuring Hall viscosity of graphene's electron fluid. *Science* **364**, 162-165 (2019).

18 Guo, H., Ilseven, E., Falkovich, G. & Levitov, L. S. Higher-than-ballistic conduction of viscous electron flows. *PNAS* **114**, 3068-3073 (2017).

19 Wallbank, J. R. *et al.* Excess resistivity in graphene superlattices caused by umklapp electron–electron scattering. *Nat. Phys.* **15**, 32-36 (2019).

20 Pellegrino, F. M. D., Torre, I. & Polini, M. Nonlocal transport and the Hall viscosity of two-dimensional hydrodynamic electron liquids. *Phys. Rev. B* **96**, 195401 (2017).

21 Torre, I., Tomadin, A., Geim, A. K. & Polini, M. Nonlocal transport and the hydrodynamic shear viscosity in graphene. *Phys. Rev. B* **92**, 165433 (2015).

22 Li, Q. & Das Sarma, S. Finite temperature inelastic mean free path and quasiparticle lifetime in graphene. *Phys. Rev. B* **87**, 085406 (2013).

23 Polini, M. & Vignale, G. in *No-nonsense Physicist: An overview of Gabriele Giuliani's work and life* (eds Marco Polini, Giovanni Vignale, Vittorio Pellegrini, & Jainendra K. Jain) 107-124 (Scuola Normale Superiore, 2016).

24 Krishna Kumar, R. *et al.* High-temperature quantum oscillations caused by recurring Bloch states in graphene superlattices. *Science* **357**, 181-184 (2017).

25 Yankowitz, M. *et al.* Emergence of superlattice Dirac points in graphene on hexagonal boron nitride. *Nat. Phys.* **8**, 382-386 (2012).

26 Ponomarenko, L. A. *et al.* Cloning of Dirac fermions in graphene superlattices. *Nature* **497**, 594 (2013).

27 Dean, C. R. *et al.* Hofstadter's butterfly and the fractal quantum Hall effect in moiré superlattices. *Nature* **497**, 598-602 (2013).

28 Hunt, B. *et al.* Massive Dirac fermions and Hofstadter butterfly in a van der Waals heterostructure. *Science* **340**, 1427-1430 (2013).

29 Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43-50 (2018).

30 Goodwin, Z. A. H. *et al*. Critical role of device geometry for the phase diagram of twisted bilayer graphene. Preprint at http://arXiv.org/abs/1911.08464 (2019).

Methods-only references

31 Wang, L. et al. One-Dimensional Electrical Contact to a Two-Dimensional Material. *Science* **342**, 614-617 (2013).

32 Cao, Y. *et al.* Quality heterostructures from two-dimensional crystals unstable in air by their assembly in inert atmosphere. *Nano Lett.* **15**, 4914-4921 (2015).

33 Hwang, E. H. & Das Sarma, S. Screening, Kohn anomaly, Friedel oscillation, and RKKY interaction in bilayer graphene. *Phys. Rev. Lett.* **101**, 156802 (2008).

34 Borghi, G., Polini, M., Asgari, R. & MacDonald, A. H. Dynamical response functions and collective modes of bilayer graphene. *Phys. Rev. B* **80**, 241402 (2009).

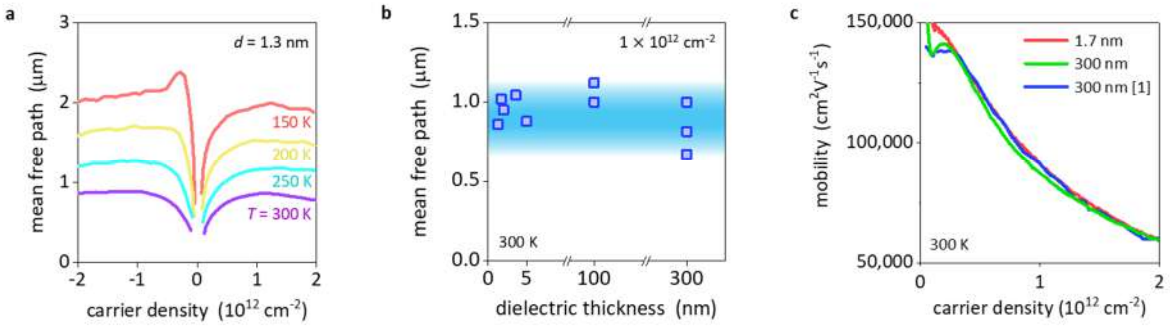
Acknowledgements

This work was supported by the European Research Council, Graphene Flagship, EPSRC Grand Challenges (EP/N010345/1), the Royal Society and Lloyd’s Register Foundation. A.I.B., W.K. were supported by Graphene NowNANO Doctoral Training Programme.

SUPPLEMENTARY INFORMATION

#1. Mean free path and mobility

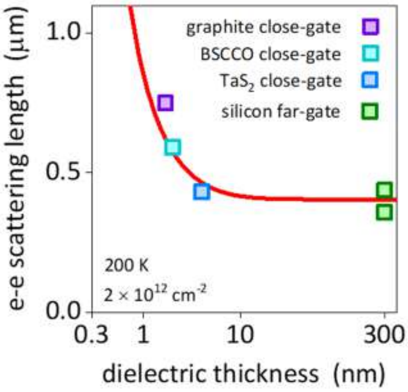
We carefully examined transport characteristics for several monolayer graphene devices with different dielectric thicknesses d. The mean free path ℓ with respect to momentum-non-conserving collisions was determined from the measured longitudinal resistivity ρ by using the Drude formula. The carrier density n was found from Hall measurements. Typical results for ℓ as a function of n are shown in Supplementary Fig. 1a. The mean free path first increases with increasing n and then saturates for $n \gtrsim 1.0 \times 10^{12}$ cm^{-2}. It monotonically decreases with temperature T as expected. Such behavior was observed for all the measured devices independently of their d. This is elucidated by Supplementary Fig. 1b that shows ℓ for different d at the given n at room T. One can see that the measured ℓ varied only slightly, from ~ 0.7 to $1.1\ \mu\text{m}$, depending on graphene device's quality. Similarly, carrier mobilities $\mu(n)$ exhibited little dependence on d (Supplementary Fig. 1c).



Supplementary Figure 1 | Transport characteristics for different thicknesses of the gate dielectric. a, $\ell(n)$ for a graphene device with $d \approx 1.3$ nm at a few representative T. **b,** ℓ for devices with different d at 300 K; $n = 1 \times 10^{12}$ cm^{-2}. **c,** Density dependence $\mu(n)$ at room T. The mobilities measured for devices with different d collapse on a single curve. The red and green curves are for gate dielectrics with ≈ 1.7 and 300 nm, respectively. The blue curve: Data from ref. 1 to indicate the generality of such behavior at elevated T.

#2. Different screening materials

Because graphite is a semimetal[2,3] with a relatively low carrier concentration of the order of 10^{19} cm^{-3}, we have checked the generality of our conclusions using other metallic substrates, namely $Bi_2Sr_2CaCu_2O_{8+x}$ (BSCCO) and TaS_2 which have concentrations of $\sim 10^{22}$ cm^{-3} (ref. 4). To this end, devices similar to those shown in Fig. 1a of the main text were fabricated but, instead of graphite, cleaved BSCCO and TaS_2 crystals served as metallic substrates. To protect them from degradation, fabrication had to be carried out in an argon atmosphere of a glovebox as discussed in Methods. The carrier mobility μ for the latter devices was comparable to that of the devices made with graphite screening gates but only for high $n \gtrsim 2 \times 10^{12}$ cm^{-2}. At lower n, the electronic quality was insufficient to probe electron viscosity because of short ℓ, presumably due to extra charges that appear on the metallic surfaces exposed to the ambient atmosphere. Accordingly, for the alternative screening substrates, we worked in the high n regime to measure the viscous Hall resistance and then extract ℓ_{ee}. Supplementary Fig. 2 shows the resulting ℓ_{ee} for graphene devices using various screening materials. Within our experimental accuracy, no difference in ℓ_{ee} could be noticed, and the experimental data closely followed the theoretical predictions.



Supplementary Figure 2 | Electron-electron scattering in devices with different materials used for proximity screening. Symbols: Measured ℓ_{ee} at 200 K and $2 \times 10^{12}\ \mathrm{cm}^{-2}$ (color coded). Solid curve: Theory.

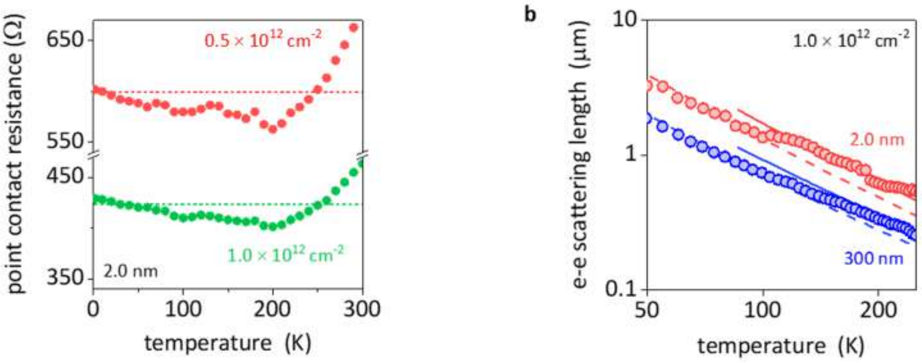
#3. Point contact geometry

For completeness, we also measured ℓ_{ee} using the point-contact geometry[5]. By applying an electric current through a graphene constriction and monitoring a voltage drop at nearby contacts (see Fig. 1a of the main text), the point contact resistance R_{PC} was measured. Supplementary Fig. 3a shows $R_{\mathrm{PC}}(T)$ for a graphene constriction with a geometrical width of $\sim 350\ \mathrm{nm}$ as found by atomic force microscopy. The transport width w of the constriction was somewhat smaller, $\sim 270\ \mathrm{nm}$, as found by fitting $R_{\mathrm{PC}}(n)$ at liquid-helium T by the standard Sharvin formula ($R_{\mathrm{Sh}} = \frac{\pi h}{4e^2}\frac{1}{w\sqrt{n\pi}}$). The smaller width inferred from the fit is expected and presumably caused by edge roughness[5]. R_{PC} exhibited a nonmonotonic T dependence, becoming at intermediate T notably smaller than the ideal value in the ballistic limit (Supplementary Fig. 3a). This "superballistic" behavior is due to e-e scattering as discussed elsewhere[5,6].

To extract ℓ_{ee} from the measurements such as those shown in Supplementary Fig. 3a, we used the expression[5,6]

$$R_{\mathrm{PC}} = (1/R_{\mathrm{Sh}} + G_v)^{-1} + R_{\mathrm{C}}$$

where $R_{\mathrm{C}} = b\rho$ is the contact resistance arising from the wide regions near the point contact. R_{C} can be determined accurately for the known ρ whereas the dimensionless coefficient b is found from numerical simulations[5]. The viscous contribution G_v to the point-contact conductivity is given by[6] $G_v = \frac{\sqrt{|n|\pi}e^2 w^2}{8\hbar\ell_{\mathrm{ee}}}$. Supplementary Fig. 3b shows examples of $\ell_{\mathrm{ee}}(T)$ found using the above analysis. The behavior of ℓ_{ee} agrees well with that found from the Hall viscosity measurements in the main text. For example, ℓ_{ee} is clearly enhanced for devices with close metallic gates. The experimental data also agree with theory whereas relatively small deviations from it at high T are due to non-Fermi-liquid corrections as reported in ref. 5 and, also, explained below.



Supplementary Figure 3 | Electron-electron scattering length found from point-contact measurements. a, Point-contact resistance for a device with a close gate ($d \approx 2.0$ nm) at different n (color coded). Dots: Experimental data. The dashed lines indicate the ideal value expected in the ballistic limit at low T. **b,** $\ell_{\rm ee}(T)$ for $d = 2.0$ (red) and $300\ {\rm nm}$ (blue) for the given n. Symbols: Experiment. Dashed curves: Theoretical predictions with no fitting parameters. Solid curves: Same theory data but multiplied by a numerical coefficient of 1.3.

#4. Microscopic theory of screened electron-electron scattering

In this Section we briefly described our approach to calculate $\ell_{\rm ee} = v_{\rm F}\tau_{\rm ee}$. The mean free time $\tau_{\rm ee}$ for e-e scattering is controlled by the one-body Green's function $G_\lambda(\boldsymbol{k},\omega)$, where $\lambda = \pm 1$ is a band index ($\lambda = +1$ for conduction-band states and $\lambda = -1$ for valence-band states). This quantity satisfies the Dyson equation (setting $\hbar = 1$), $G_\lambda(\boldsymbol{k},\omega) = [\omega - \xi_{\boldsymbol{k},\lambda} - \Sigma_\lambda(\boldsymbol{k},\omega)]^{-1}$, where $\xi_{\boldsymbol{k},\lambda}$ are single-particle band energies measured from the chemical potential μ and $\Sigma_\lambda(\boldsymbol{k},\omega)$ is the retarded self-energy. The latter quantity needs to be approximated. In weakly-correlated materials, a good approximation is the so-called GW approximation[7,8] in which the electron self-energy is expanded to first order in the dynamically screened Coulomb interaction $W(\boldsymbol{q}, i\Omega)$

$$\Sigma_\lambda(\boldsymbol{k}, i\omega_n) = -k_{\rm B}T\sum_{\lambda'=\pm1}\int\frac{d^2\boldsymbol{q}}{(2\pi)^2}\sum_{m=-\infty}^{+\infty}W(\boldsymbol{q}, i\Omega_m)F_{\lambda\lambda'}(\theta_{\boldsymbol{k},\boldsymbol{k}-\boldsymbol{q}})G_{\lambda'}(\boldsymbol{k}-\boldsymbol{q}, i\omega_n + i\Omega_m) \quad \text{(S1)}$$

where $\omega_n = (2n+1)\pi k_B T$ is a fermionic Matsubara frequency, the sum runs over all the bosonic Matsubara frequencies $\Omega_m = 2m\pi k_{\rm B}T$, $\theta_{\boldsymbol{k},\boldsymbol{k}-\boldsymbol{q}}$ is the angle between $\boldsymbol{k}$ and $\boldsymbol{k}-\boldsymbol{q}$, and $F_{\lambda\lambda'}(\varphi) = [1 + \lambda\lambda'\cos(\varphi)]/2$ is the so-called chirality factor[9]. The retarded self-energy can be obtained after analytical continuation $i\omega_n \to \omega + i0^+$. For the sake of concreteness and without loss of generality due to particle-hole symmetry, we focus on electron-doped graphene, i.e. on the case $E_{\rm F} > 0$, where $E_{\rm F} = v_{\rm F}k_{\rm F}$ is the Fermi energy. Here, $v_{\rm F} \sim 10^6\ {\rm m/s}$ ($k_{\rm F} = \sqrt{\pi n}$) is the Fermi velocity (Fermi wave number), with $n > 0$ the electron density.

The Dyson equation combined with the approximate GW expression for the electron self-energy define a *self-consistent* approximation, whose self-energy and Green's function can be calculated based on an iterative procedure. One first calculates the self-energy from the GW expression by using in the right-hand side of the *non-interacting* Green's function $G_{\lambda'}(\boldsymbol{k}-\boldsymbol{q}, i\omega_n + i\Omega_m) \to G^{(0)}_{\lambda'}(\boldsymbol{k}-\boldsymbol{q}, i\omega_n + i\Omega_m) = 1/(i\omega_n - \xi_{\boldsymbol{k}-\boldsymbol{q},\lambda'})$. The obtained result is then replaced in the right-hand side of the Dyson equation, obtaining a new Green's function. The latter is then used to re-calculate the self-energy via the GW equation, until self-consistency is achieved. Now, the key point is that deep in the Fermi liquid regime, i.e. for $|\boldsymbol{k}| \simeq k_{\rm F}$ and $|\omega|/E_{\rm F}, k_{\rm B}T/E_{\rm F} \ll 1$, the self-energy is a small correction to the bare band energy $\xi_{\boldsymbol{k},\lambda}$ and such self-consistency is unnecessary. In this limit indeed, quasiparticles are long lived because of the ineffectiveness of e-e collisions (Pauli blocking) and $\mathrm{Im}[\Sigma_+(\boldsymbol{k},\omega)] \propto (k_{\rm B}T/E_{\rm F})^2 + (\omega/E_{\rm F})^2$, modulo logarithmic corrections. In this regime, it is therefore well justified to replace $G_{\lambda'}(\boldsymbol{k}-\boldsymbol{q}, i\omega_n + i\Omega_m)$ with $G^{(0)}_{\lambda'}(\boldsymbol{k}-\boldsymbol{q}, i\omega_n + i\Omega_m)$ in the right-hand side of the GW equation obtaining the so-called $G^{(0)}W$ approximation[7,8].

Since this is the simplest possible theory, we use the $G^{(0)}W$ approximation also away from the Fermi liquid regime, being aware of the fact, however, that the lack of full self-consistency is expected to lead to inaccuracies. In particular, it is easy to demonstrate that $\ell_{\rm ee}|_{G^{(0)}W} < \ell_{\rm ee}|_{GW}$. Since in weakly correlated materials such as graphene the GW approximation is expected to be quantitatively good (i.e. $\ell_{\rm ee}|_{GW}$ is expected to be close to the experimentally value of $\ell_{\rm ee}$), we do expect the non-self-consistent result $\ell_{\rm ee}|_{G^{(0)}W}$ to systematically *underestimate* the experimentally measured $\ell_{\rm ee}$. Therefore, in the main text, we have compared experimental data with $\ell_{\rm ee}|_{G^{(0)}W}$ after multiplying the latter by a *constant* enhancement factor of 1.3, which is independent of all microscopic parameters (Fig. 2 of the main text).

The quantity $\ell_{\rm ee}|_{G^{(0)}W}$ can be calculated numerically once one specifies the dynamically screened potential $W(\boldsymbol{q}, i\Omega_m)$. In the random phase approximation[7], $W(\boldsymbol{q},\omega) = V_{\boldsymbol{q}}/[1 - V_q\chi^{(0)}_{nn}(\boldsymbol{q},\omega)]$, where $\chi^{(0)}_{nn}(\boldsymbol{q},\omega)$ is the well-known density-density response function of doped graphene[9] and $V_{\boldsymbol{q}}$ is the 2D Fourier transform of the e-e interaction potential, which is sensitive to screening caused by nearby metal gates and gate dielectrics. For our metal/hBN/graphene/hBN/metal heterostructures, electrostatic calculations yield

$$V_{\boldsymbol{q}} = \frac{4\pi e^2}{q\sqrt{\epsilon_x\epsilon_z}} \frac{\sinh\left(qd\sqrt{\frac{\epsilon_x}{\epsilon_z}}\right)\sinh\left(qd'\sqrt{\frac{\epsilon_x}{\epsilon_z}}\right)}{\sinh\left[q(d+d')\sqrt{\frac{\epsilon_x}{\epsilon_z}}\right]} \tag{S2}$$

where d' (d) is the thickness of hBN above (below) graphene, and ϵ_x and ϵ_z are the static in-plane and out-of-plane permittivities of hBN. Two metal gates, modelled as perfect conductors, are placed above and below graphene at distances d' and $d \ll d'$, respectively, and are separated from graphene by hBN. Numerical calculations of $\ell_{\rm ee}|_{G^{(0)}W}$ have been carried out by using this effective screened e-e interaction for sufficiently large $d' \approx 60$ nm and known $\epsilon_x = 6.70$, and $\epsilon_z = 3.56$ (see, for example, ref. 10). Values of d, n, and T were variables in our calculations. Pertinent results are presented in Fig. 2 of the main text.

For a qualitative understanding of the role of screening, it is useful to obtain an approximate expression for $\ell_{\rm ee}|_{G^{(0)}W}$ as a function of all system parameters. To this end, we follow ref. 8 and derive a formula for $\ell_{\rm ee}|_{G^{(0)}W}$ which is exact in the Fermi-liquid regime, $k_{\rm B}T \ll E_{\rm F}$. The calculations follow essentially the same steps as in ref. 8, modulo minor differences, which stem from the regularity of $V_{\boldsymbol{q}}$

in the long-wavelength $q \to 0$ limit and will be discussed elsewhere. Indeed, $\lim_{q \to 0} V_{\boldsymbol{q}} = 4\pi e^2 d_{\text{eff}}/\epsilon_z \equiv V_0$, where $d_{\text{eff}} = dd'/(d + d')$. This formula allows a simple interpretation. Having the two, top and bottom, gates is like having two capacitors in parallel. Indeed, we can write $V_0 = e^2/C_{\text{eff}}$, where the $C_{\text{eff}} = C_d + C_{d'}$ is the sum of the two relevant geometrical capacitances (per unit area), $C_d = \epsilon_z/(4\pi d)$ and $C_d = \epsilon_z/(4\pi d')$. After restoring $\hbar$, we obtain

$$\lim_{\frac{k_{\text{B}}T}{E_{\text{F}}} \to 0} \ell_{\text{ee}}\big|_{G^{(0)}W} = \frac{4\hbar v_{\text{F}} E_{\text{F}}}{\pi} \frac{1}{(k_{\text{B}}T)^2 \ln\left(\frac{2E_{\text{F}}}{k_{\text{B}}T}\right)} \left(\frac{1+2d_{\text{eff}}q_{\text{TF}}}{2d_{\text{eff}}q_{\text{TF}}}\right)^2 \qquad \text{(S3).}$$

Eq. 2 in the main text is simply obtained from Eq. S3 by taking the limit $d' \to \infty$.

Before concluding this section, let us comment on possible corrections to our model caused by the fact that real gates are not the assumed perfect conductors. The effect of a finite density-of-states in a metallic gate can be estimated using the Thomas-Fermi approximation. It is possible to show that, in this approximation, the previous asymptotic result for $\ell_{\text{ee}}\big|_{G^{(0)}W}$ in the limit $k_{\text{B}}T \ll E_{\text{F}}$ holds if one replaces $d \to d + 1/q_{\text{TF}}$, where q_{TF} is the Thomas-Fermi screening wavenumber in gate's material. As a crude estimate for graphite, we take $q_{\text{TF}}^{(\text{G})}$ to be the same as that of a three-dimensional metal[7], i.e. $q_{\text{TF}}^{(\text{G})} = \sqrt{4e^2 m_{\text{G}} k_{\text{F}}^{(\text{G})}/(\pi\hbar^2\epsilon_{\text{G}})}$, where $m_{\text{G}} \approx 0.2\, m_{\text{e}}$ is the effective mass of charge carriers in graphene, $\epsilon_{\text{G}} \approx 3$, and $k_{\text{F}}^{(\text{G})} = (3\pi^2 n_{\text{G}})^{1/3}$, with $n_{\text{G}} \approx 10^{19}\ \text{cm}^{-3}$. This yields an extra gate spacing of about 9 Å. This is probably an overestimate as undoped graphene layers provide efficient screening at the same distance due to self-doping[11]. For the other metallic substrates, we find $1/q_{\text{TF}} \approx 2$ Å, that corresponds to interatomic distances as expected.

#5. Suppression of umklapp e-e scattering by proximity screening

It has been shown[12] that umklapp e-e scattering (U_{ee}) substantially increases the resistivity of high-quality graphene-on-hBN superlattices (SL) in the range of T between 50 and 200 K. The SL potential is generated by the moiré pattern that has a period $\lambda \approx 15\ \text{nm}$ for a perfectly aligned graphene and hBN crystals. U_{ee} is a process where a crystal lattice (superlattice in our case) provides interacting electrons with an additional momentum kick such that the momentum conservation takes the form $\mathbf{k_3} + \mathbf{k_4} = \mathbf{k_1} + \mathbf{k_2} + \mathbf{g}$, where $\mathbf{k}_{1,2}$ and $\mathbf{k}_{3,4}$ are the initial and final momenta of two electrons near the Fermi level, and $\mathbf{g} = (\mathbf{g}_x, \mathbf{g}_y)$ is a reciprocal vector of the crystal (Supplementary Fig. 4a). Such a process becomes possible only for $4k_{\text{F}} > g$, where $g = |\mathbf{g}| = \frac{4\pi}{\sqrt{3}\lambda}$ is the length of one of the 6 shortest vectors of the reciprocal SL.

The contribution of U_{ee} towards graphene's resistivity ρ is given by[12]

$$\Delta\rho = \frac{\hbar\pi}{e^2 k_{\text{F}}}\, l_{U\text{ee}}^{-1} \ \text{ with } \ l_{U\text{ee}}^{-1} = \frac{(k_{\text{B}}T)^2}{12\,\pi^2 v_{\text{F}}^4 k_{\text{F}}} \sum_{\mathbf{g}} (\mathrm{g}_x)^2 \int \frac{d\theta_{\mathrm{k}_1} d\theta_{\mathrm{k}_3}}{\left|\sin\left(\theta_{\mathrm{k}_2} - \theta_{\mathrm{k}_4}\right)\right|} \left|\sum_{i=\mathrm{I}}^{\mathrm{IV}} \sum_{s'=\pm} M_{ss'}^{(i)}\right|^2 \qquad \text{(S4)}$$

where $\theta_{\mathbf{k}}$ denotes an angle between $\mathbf{k}$ and x-axis, $s = \pm$ stands for the conductance/valence-band states (fixed by doping), and s' marks virtual intermediate states. In Eq. S4, the inverse umklapp scattering length, $l_{U\mathrm{ee}}^{-1}$, is determined by the sum of four Feynman diagrams shown in Supplementary Fig. 4b, each described by the scattering amplitude $M_{ss'}^{(i)}$ $(i = \mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV})$. For example, the first diagram gives a contribution

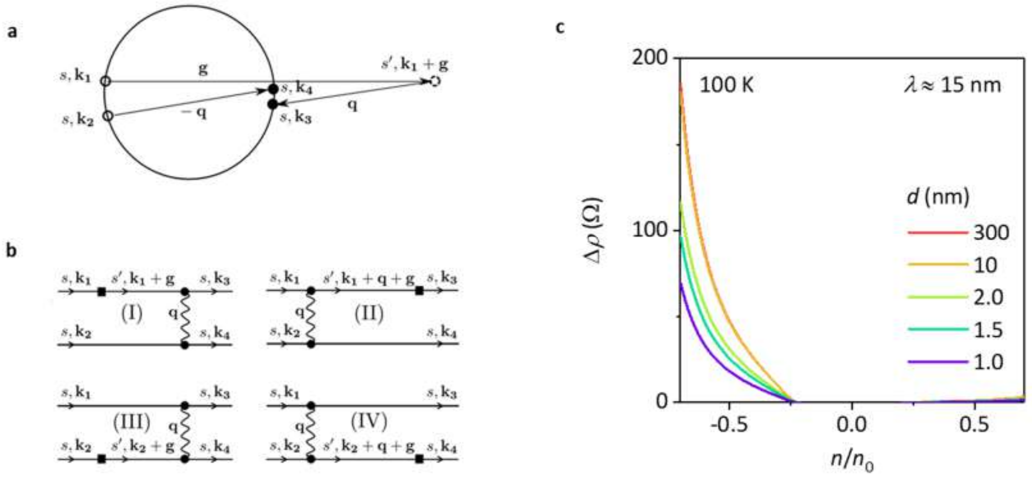
$$M_{ss'}^{(\mathrm{I})} = \frac{W(\mathbf{g})\frac{1+ss'e^{i\theta_{\mathbf{k}_1+\mathbf{g}}-i\theta_{\mathbf{k}_3}}}{2}V(|\mathbf{k}_2-\mathbf{k}_4|)\frac{1+e^{i\theta_{\mathbf{k}_2}-i\theta_{\mathbf{k}_4}}}{2}}{sv|\mathbf{k}_1|-s'v|\mathbf{k}_1+\mathbf{g}|} \tag{S5}$$

where $W(\mathbf{g})$ stands for the scattering amplitude of an electron off the moiré SL[13,14], and

$$V(q) = \frac{V_{\boldsymbol{q}}(q,d,d')}{1+V_{\boldsymbol{q}}(q,d,d')\Pi(q)} \tag{S6}$$

is the Coulomb interaction screened by both gate and the Fermi sea in graphene; $\Pi(q \le 2k_{\mathrm{F}}) = \frac{2k_{\mathrm{F}}}{\hbar\pi v_{\mathrm{F}}}$ is the Thomas-Fermi polarization operator[15-18]. From the form of $V_{\boldsymbol{q}}$ in Eq. S2, it is straightforward to see that, for e-e scattering with the momentum transfer $q \sim g/2$, the gate starts playing a notable screening role only if $d_{\mathrm{eff}} \lesssim \sqrt{\frac{\epsilon_z}{\epsilon_x}\frac{1}{g}} \approx 0.1\lambda$ $(d_{\mathrm{eff}} \approx d \ll d')$. Expressions for the other diagrams in Supplementary Fig. 4b can be obtained by changing input momenta and $\boldsymbol{q}$ in Eq. S5.

The U_{ee} contribution, computed using the same SL parameters as those in refs. 12 and 19, exhibits a significant suppression for $d \lesssim 2$ nm (Supplementary Fig. 4c). In these calculations, the absolute value of $\Delta\rho \propto l_{U\mathrm{ee}}^{-1}$ obviously depends on the moiré potential's strength. To compare the effect of proximity screening on U_{ee}, without relying on a detailed choice of SL parameters, we also plot the ratio $l_{U\mathrm{ee}}^{-1}(\infty)/l_{U\mathrm{ee}}^{-1}(d)$ at $n \approx -\frac{1}{2}n_0$ and compare the theoretical results with the experimentally found ratio $\Delta\rho(\infty)/\Delta\rho(d)$ [see Fig. 3 of the main text].



Supplementary Figure 4 | Screened umklapp e-e scattering in graphene superlattices. a, Kinematics of U_{ee} scattering. **b,** Feynman diagrams for $M_{ss'}^{(i)}$. **c,** Additional resistivity caused by U_{ee} for different distances to the gate (color coded).

Supplementary references

1 Wang, L. *et al.* One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614-617 (2013).

2 Dresselhaus, M. S. & Dresselhaus, G. Intercalation compounds of graphite. *Adv. Phys.* **51**, 1-186 (2002).

3 Yin, J. *et al.* Dimensional reduction, quantum Hall effect and layer parity in graphite films. *Nat. Phys.* **15**, 437-442 (2019).

4 Sterpetti, E., Biscaras, J., Erb, A. & Shukla, A. Comprehensive phase diagram of two-dimensional space charge doped $Bi_2Sr_2CaCu_2O_{8+x}$. *Nat. Commun.* **8**, 2060 (2017).

5 Krishna Kumar, R. *et al.* Superballistic flow of viscous electron fluid through graphene constrictions. *Nat. Phys.* **13**, 1182-1185 (2017).

6 Guo, H., Ilseven, E., Falkovich, G. & Levitov, L. S. Higher-than-ballistic conduction of viscous electron flows. *PNAS* **114**, 3068-3073 (2017).

7 Giuliani, G. F. & Vignale, G. *Quantum theory of the electron liquid*. (Cambridge University Press, Cambridge, 2005).

8 Polini, M. & Vignale, G. in *No-nonsense Physicist: An overview of Gabriele Giuliani's work and life* (eds Marco Polini, Giovanni Vignale, Vittorio Pellegrini, & Jainendra K. Jain) 107-124 (Scuola Normale Superiore, 2016).

9 Kotov, V. N., Uchoa, B., Pereira, V. M., Guinea, F. & Castro Neto, A. H. Electron-electron interactions in graphene: current status and perspectives. *Rev. Mod. Phys.* **84**, 1067-1125 (2012).

10 Woessner, A. *et al.* Highly confined low-loss plasmons in graphene–boron nitride heterostructures. *Nat. Mater.* **14**, 421 (2014).

11 Guinea, F. Charge distribution and screening in layered graphene systems. *Phys. Rev. B* **75**, 235433 (2007).

12 Wallbank, J. R. *et al.* Excess resistivity in graphene superlattices caused by umklapp electron–electron scattering. *Nat. Phys.* **15**, 32-36 (2019).

13 Wallbank, J. R., Patel, A. A., Mucha-Kruczyński, M., Geim, A. K. & Fal'ko, V. I. Generic miniband structure of graphene on a hexagonal substrate. *Phys. Rev. B* **87**, 245408 (2013).

14 Wallbank, J. R., Mucha-Kruczyński, M., Chen, X. & Fal'ko, V. I. Moiré superlattice effects in graphene/boron-nitride van der Waals heterostructures. *Ann. Phys.* **527**, 359-376 (2015).

15 Gorbar, E. V., Gusynin, V. P., Miransky, V. A. & Shovkovy, I. A. Magnetic field driven metal-insulator phase transition in planar systems. *Phys. Rev. B* **66**, 045108 (2002).

16 Ando, T. Screening effect and impurity scattering in monolayer graphene. *J. Phys. Soc. Jpn.* **75**, 074716 (2006).

17 Wunsch, B., Stauber, T., Sols, F. & Guinea, F. Dynamical polarization of graphene at finite doping. *New J. Phys.* **8**, 318-318 (2006).

18 Hwang, E. H. & Das Sarma, S. Dielectric function, screening, and plasmons in two-dimensional graphene. *Phys. Rev. B* **75**, 205418 (2007).

19 Lee, M. *et al.* Ballistic miniband conduction in a graphene superlattice. *Science* **353**, 1526-1529 (2016).

Extreme electron–hole drag and negative mobility in the Dirac plasma of graphene

Leonid A. Ponomarenko[1,2], Alessandro Principi[2], Andy D. Niblett[1], Wendong Wang[3], Roman V. Gorbachev[3], Piranavan Kumaravadivel[3], Alexey I. Berdyugin[2,5], A. V. Ermakov[2], Sergey Slizovskiy[2,3], Kenji Watanabe[4], Takashi Taniguchi[4], Qi Ge[5], Vladimir I. Fal'ko[2,3], Laurence Eaves[6], Mark T. Greenaway[6,7], Andre K. Geim[2,3]

[1]Department of Physics, University of Lancaster, Lancaster LA1 4YW, UK
[2]Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK
[3]National Graphene Institute, University of Manchester, Oxford Rd, Manchester M13 9PL, UK
[4]National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan
[5]Institute for Functional Intelligent Materials, National University of Singapore, 117544 Singapore
[6]School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK
[7]Department of Physics, Loughborough University, Loughborough LE11 3TU, UK

Coulomb drag between adjacent electron and hole gases has attracted considerable attention, being studied in various two–dimensional systems, including semiconductor and graphene heterostructures. Here we report measurements of electron–hole drag in the Planckian plasma that develops in monolayer graphene in the vicinity of its Dirac point above liquid–nitrogen temperatures. The frequent electron–hole scattering forces minority carriers to move against the applied electric field due to the drag induced by majority carriers. This unidirectional transport of electrons and holes results in nominally negative mobility for the minority carriers. The electron–hole drag is found to be strongest near room temperature, despite being notably affected by phonon scattering. Our findings provide better understanding of the transport properties of charge–neutral graphene, reveal limits on its hydrodynamic description and also offer insight into quantum–critical systems in general.

If electron– and hole– doped two–dimensional (2D) conductors are placed in close proximity to each other, Coulomb interactions between charge carriers in adjacent layers lead to electron–hole drag (for review, see refs.[1,2]). The drag was extensively studied using various electronic systems based on GaAs heterostructures and, more recently, graphene[1-12]. The strength of Coulomb interaction rapidly increases with decreasing the distance between 2D systems, and the ultimately strong drag is expected if electrons and holes coexist within the same atomic plane. Graphene near its Dirac or neutrality point (NP) provides the realization of such an electronic system. Indeed, close to the NP, a finite temperature T leads to thermal excitations of electrons and holes, whereas their relative concentrations can be controlled by gate voltage. The resulting electron–hole plasma is strongly interacting and represents a quantum critical system where particle–particle collisions are governed by Planckian dissipation[13-23]. The system is also often referred to as Dirac fluid, assuming inter–carrier scattering dominates other scattering mechanisms. Because the Dirac plasma in graphene is a relatively simple and tunable electronic system, its behavior can be insightful for understanding of electron transport in more complex Planckian systems including "strange metals" and high–temperature superconductors in the normal state[24,25]. There is also an interesting conceptual overlap with relativistic electron–positron plasmas generated in cosmic events, which are difficult to recreate in laboratory experiments[26]. Previous experimental studies of the Dirac plasma reported its hydrodynamic flow[20], the violation of the Wiedemann–Franz law[17], giant linear magnetoresistance[23] and other anomalies indicative of the quantum–critical regime[18-23]. So far, the possibility to probe mutual drag between electron and hole subsystems within the Dirac plasma has escaped attention.

Longitudinal and Hall resistivity of the Dirac plasma

Our devices were multi-terminal Hall bars made from monolayer graphene encapsulated between two crystals of hexagonal boron nitride. Relatively thick graphite placed under the trilayer heterostructures served as a gate electrode. To avoid an obscuring contribution from edge scattering and charge accumulation at device boundaries[12,27], it was essential to make the Hall bars no less than 10 μm in size. This allowed charge carrier mobilities to reach ~10^6 cm^2 V^{-1} s^{-1} at low T (measured at finite carrier densities of a few 10^{11} cm^{-2}). The remnant doping was also low, typically ~$2{\cdot}10^{10}$ cm^{-2}. We studied several such devices, and 3 of them were chosen for detailed analysis of their longitudinal and Hall resistivities near the NP (ρ and R_H, respectively). All the devices exhibited practically identical characteristics so that, for brevity and consistency, we illustrate the observed behavior using the data obtained from one of them. Its optical micrograph is shown in the inset of Fig. 1a.

Near the NP, where both electrons and holes are present, the total charge density in graphene is given by $en = e(n_e - n_h)$ where n_e and n_h are the sheet densities of electrons and holes, respectively, and e is the electron charge. The charge density en can be controlled capacitively by gate voltage (Supplementary note 1). The device's resistivity ρ as a function of n is shown in Fig. 1a (positive and negative n correspond to electrons and holes, respectively). At low T, $\rho(n)$ exhibits a sharp peak at the NP (red curve). It is instructive[18,23] to replot $\rho(n)$ in a logarithmic scale (right inset) which reveals that ρ is weakly density dependent for $n \lesssim 10^{10}$ cm^{-2}. The point at which ρ becomes notably dependent on n is labelled as δn (arrow in the inset of Fig. 1a). The value of δn at low T provides a measure of residual charge inhomogeneity ("electron–hole puddles")[18,23]. Despite its extra–large size (15×30 μm^2), the device exhibited δn of only ~$5{\cdot}10^9$ cm^{-2} at low T. As T increased, the peak in $\rho(n)$ became wider and smaller because of thermally excited electrons and holes (black curves in Figs. 1a,b). At room T, the measured value of δn increased by an order of magnitude with respect to that at liquid–helium T (inset of Fig. 1b).

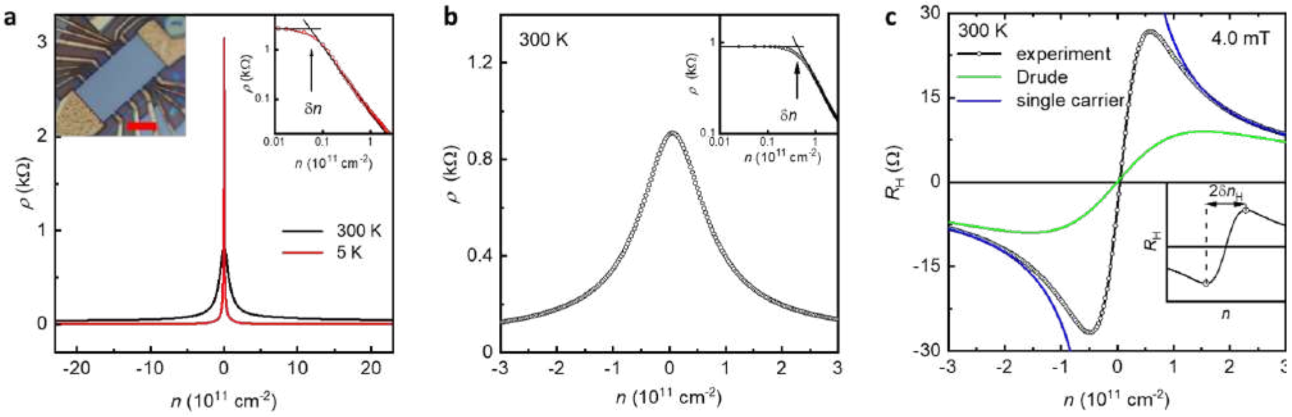


Figure 1| Transport characteristics of monolayer graphene near the neutrality point. (**a**) Resistivity at room and low T in zero magnetic field. Left inset: Optical micrograph of one of the studied devices. Scale bar, 10 μm. Right inset: $\rho(n)$ measured at 5 K is replotted on a log-log scale. The arrow marks δn at which the resistivity starts responding to gate voltage. (**b**) Zooming in on the behavior of ρ in the vicinity of the NP at 300 K (black symbols, same curve as in **a**). Inset: same as the inset of **a** but at room T. (**c**) Room-T Hall resistivity in small B (open symbols). The green curve plots R_H expected from the standard Drude model assuming electron-hole symmetry $\mu_e(n) = \mu_h(n)$ (the curve does not depend on the mobilities' absolute values). Blue curves: $R_H = B/ne$ as expected for a single carrier type. The inset explains how we define δn_H that is analogous to δn in panels **a**, **b**.

The corresponding behavior of Hall resistivity R_H near the NP is shown in Fig. 1c. A small magnetic field B was applied perpendicular to graphene, and its value (4 mT) was carefully chosen to keep electron transport deep in the weak–field limit where R_H remained linear in B (nonlinearities started emerging typically above 10 mT) and, at the same time, to ensure a large enough Hall response to record R_H with high accuracy. Both conditions were essential for our analysis described below. Away from the NP, at densities $|n| > 10^{11}$ cm^{-2}, R_H evolved as B/ne, as expected for transport dominated by one type of charge carriers (Fig. 1c). Near the NP, $R_H(n)$ departed from this dependence due to the presence of both electrons and holes and changed its sign at the NP, indicating a switch from majority hole to majority electron transport. The range of n over which both electrons and holes contributed to the Hall effect can be characterized by δn_H, the distance between the maximum and minimum in R_H (see the inset of Fig. 1c). δn_H did not depend on B (in the discussed limit of weak B) and increased with T as the density of thermally excited charge carriers increased. This is illustrated in Fig. 2a that shows $R_H(n)$ at three different T. As the temperature increased, the extrema in R_H were broadened and moved further apart. Fig. 2b shows δn_H measured over a wide range of T and compares the behavior with $\delta n(T)$ determined from the broadening of the peak in $\rho(n)$. Both δn_H and δn exhibit similar values and a roughly parabolic T dependence. At low T, they tend to a constant value due to residual charge inhomogeneity.

The transport behavior described above and illustrated by Figs. 1 and 2 is archetypical of high–quality graphene. It was previously observed in numerous experiments but not subjected to in–depth analysis. Most often, $\rho(n)$ curves have been used only to evaluate the charge inhomogeneity of a device (as described above) and extract the field–effect mobility defined as $\mu(n) = 1/ne\rho(n)$. The latter expression is valid only in the case of one type of carrier so that, unsurprisingly, μ has been found to diverge near the NP because n goes through zero (blue curve in Fig. 3a). As for the behavior of $R_H(n)$, the region close to the NP has usually been ignored with reference to the presence of electron–hole puddles. This is justified at liquid–He temperatures but, as thermal excitations overpower the effects of charge inhomogeneity with increasing T, electron transport at the NP becomes intrinsic. This high–T regime was overlooked previously and merits better understanding, which is provided below.

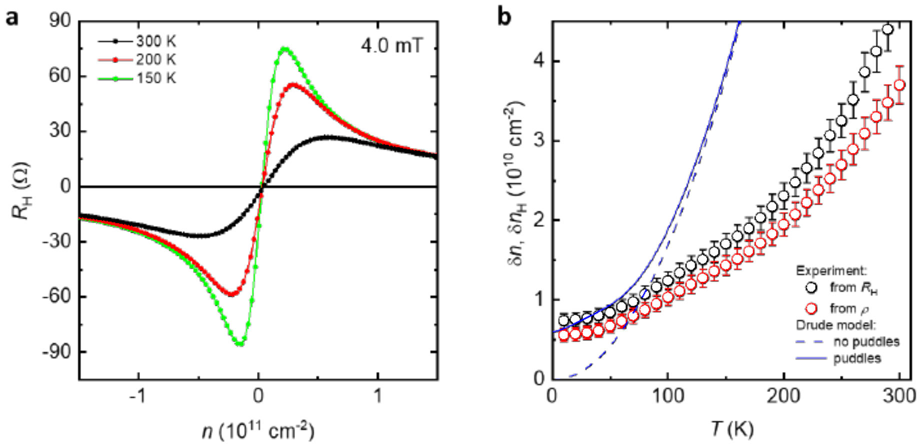


Figure 2| Hall resistivity in low magnetic fields and thermal broadening at the NP. (**a**) Examples of $R_H(n)$ at different T (color coded). (**b**) Characteristic width of the region where both electrons and holes are present. Symbols: δn_H and δn extracted from the Hall and longitudinal resistivities, respectively. Blue curves: $\delta n_H(T)$ expected from the standard Drude model that ignores electron-hole drag.

Two–fluid model for the Dirac plasma

For two types of charge carriers present in graphene near its NP, it is sensible to try to describe the transport characteristics using the standard two carrier Drude model[28]:

$$\rho(n) = \frac{1}{e(\mu_e n_e + \mu_h n_h)} \tag{1}$$

$$R_H(n) = \frac{B}{e}\frac{n_e\mu_e^2 - n_h\mu_h^2}{(n_h\mu_h + n_e\mu_e)^2}, \tag{2}$$

where μ_e and μ_h are the mobilities of electrons and holes, respectively. Their densities are given by $n_{e,h}$ = $\int f(\pm\varepsilon_k,\Theta)DoS(\varepsilon_k)d\varepsilon_k$ where $f(\pm\varepsilon_k,\Theta)$ is the Fermi–Dirac distribution for electrons (+) and holes (-), and $DoS(\varepsilon_k)$ is the density of states. For a given n, the electrochemical potential Θ can be found by solving the integral equation $n = n_e - n_h$ (Supplementary note 2), which in turn allows us to find $n_{e,h}$ as a function of n. At the NP (n and $\Theta = 0$), $n_e = n_h \equiv n_{th} = (2\pi^3/3)(k_B T/hv_F)^2$ where k_B and h are the Boltzmann and Planck constants, respectively, and v_F is graphene's Fermi velocity. At room T, the observed intrinsic broadening $\delta n_H \approx \delta n$ was approximately twice smaller than n_{th} (Fig. 2b). The room–T resistivity of ~0.9 kOhm at the NP (Fig. 1) corresponds to $\mu_e = \mu_h \approx 47{,}000$ cm^2 V^{-1} s^{-1} and yields the scattering rate of ~0.3 ps, in agreement with the Planckian frequency $\tau_P^{-1} \approx Ck_BT/h$ where C is the constant of about unity[13-23].

If the electron and hole subsystems were to respond independently to the electric field E, as the standard Drude model assumes, the electron–hole symmetry of graphene's spectrum would imply equal drift velocities and, therefore, $\mu_e = \mu_h$ (although the mobilities' value may depend on n). Then, eq. 2 simplifies to $R_H(n) = nB/e(n_e + n_h)^2$ which is independent of scattering times. This dependence is shown in Fig. 1c by the green curve that has no adjustable parameters. This curve is profoundly different from those observed experimentally. The extrema of the Drude curve are much shallower and occur further away from the NP than in the experiment. It also yields $\delta n_H \approx 2.07 n_{th}$ (dashed curve in Fig. 2b), which is ~4 times larger than δn_H measured at room T. If a finite charge inhomogeneity is included within the Drude model (solid blue curve in Fig. 2b; Supplementary note 4), we achieve a better match between experimental and theoretical curves for δn_H at low T but obviously this cannot resolve the discrepancy at high T. The failure to explain the sharp transition in $R_H(n)$ near the NP shows that the standard Drude model, assuming non–interacting fluids, is inadequate to describe the Dirac plasma's transport properties.

Next, we relax assumptions and, empirically, allow electron and hole mobilities to be unequal and even negative (the latter contradicts the Drude model's assumptions). Equations 1 and 2 contain two unknown functions $\mu_e(n)$ and $\mu_h(n)$ and, for each n, their values can uniquely be evaluated from the two measured variables, ρ and R_H. Combining eqs. 1 and 2, we obtain the following expression:

$$\mu_{e,h}(n) = \pm\frac{1}{ne\rho}\left[1 - \sqrt{\frac{n_{h,e}}{n_{e,h}}\left(1 - \frac{eR_H}{B}n\right)}\ \right] \qquad (3).$$

The electron and hole mobilities extracted using eq. 3 are plotted in Figs. 3a,b where we limit our analysis to $T \geq 150$ K so that the density of thermally excited carriers dominates over the residual charge inhomogeneity (Fig. 2b).

At the NP, the electron and hole mobilities are found to be equal as required by symmetry. Away from the NP, as the carrier density of either electrons or holes is increased, their mobility also increases, until it saturates at n of several n_{th}, where the charge density is dominated by one type of carriers (see Supplementary Fig. 2). In contrast, the mobility of the minority carriers rapidly decreases away from the NP and becomes negative at $|n| \gtrsim n_{th}$. Near room T, the mobility of minority carriers saturates to an absolute value comparable to that of majority carriers (Fig. 3a). This means that the minority carriers are dragged by majority carriers in the direction opposite to their expected drift direction and, if one type of carriers dominates, the other one is forced to drift along with a similar velocity. We observe this reversal in the drift direction for minority carriers over our entire temperature range (Fig. 3b) and for all devices. The behavior is attributed to strong Coulomb interaction between electrons and holes. For completeness, Fig. 3c shows the T dependence of the extracted mobilities at the NP where $\mu_e \equiv \mu_h$ and at a finite density where one carrier type remains present. For charge–neutral graphene, the mobilities evolve approximately as $\propto 1/T^2$ and start saturating below 150 K (Fig. 3c) where electron–hole puddles can no longer be neglected. Square dependence is expected because the

Planckian scattering time τ_P and the effective mass of the Dirac fermions are both linearly dependent on $1/T$ (ref.[29]).

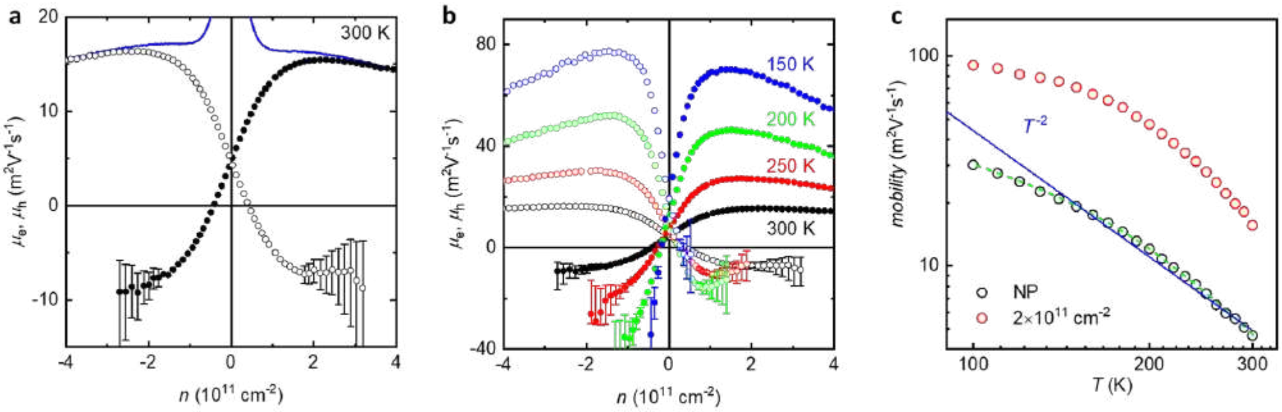


Figure 3| Trade-off between electron and hole mobilities near the Dirac point. (**a**) Room-T mobilities of electrons (solid symbols) and holes (open) extracted from Hall and longitudinal resistivities at low B using the modified Drude model. As the density of majority carriers increases away from the NP, their mobility also increases but the mobility of minority carrier rapidly becomes negative. Blue curve: field-effect mobility extracted under assumption of one type of charge carriers. The error bars arise from noise of ~0.2 Ohm in the measured Hall resistance. (**b**) Same as in **a** but at different T (color coded). (**c**) Mobility at the NP (black symbols) and a density of $2\cdot 10^{11}$ cm^{-2} (red) as a function of T. For self-consistency, the green curve shows the mobility calculated directly from the minimum conductivity rather than using eq. 3. Blue curve, T^2 dependence.

Boltzmann model for the Dirac plasma

Equations 1-3 are inadequate to accurately describe an interacting plasma. The fundamental reason for this is that electron–hole scattering leads to momentum relaxation in the electric field direction but not in the perpendicular Hall field direction[23]. Because of this relaxation anisotropy, electron–hole scattering (described by time $\tau_{\rm eh}$) contributes to the transport coefficients in a different way compared to scattering by phonons and impurities which can be parametrized by another time τ. Accordingly, to describe electron transport in the Dirac plasma at finite B, we have used the linearized Boltzmann model[30] that is presented in Supplementary note 3. In brief, the Boltzmann model yields the following coupled Drude–like equations:

$$\pm\frac{eE}{m_{\rm e,h}}+\frac{u_{\rm e,h}}{\tau}\pm\frac{\rho_{\rm h,e}}{\rho_{\rm e}+\rho_{\rm h}}\left(\frac{u_{\rm e}-u_{\rm h}}{\tau_{\rm eh}}\right)=0 \qquad (4)$$

where $u_{\rm e,h}$ are the drift velocities of electrons and holes, $m_{\rm e,h}$ are their energy–dependent effective masses, and $\rho_{\rm e,h}$ are the mass densities (Supplementary note 3). Both $\rho_{\rm e,h}$ and $m_{\rm e,h}$ are positive and depend on n and T (Supplementary Fig. 3). The first two terms of eq. 4 have exactly the same form as the standard Drude equation describing the force acting on a charge carrier due to the electric field and an opposing “frictional” force proportional to $1/\tau$. The third term corresponds to an additional frictional force caused by electron–hole scattering. This term can attain a value opposite to the electric field term and dominate over it. In the latter case, charge carriers would be dragged in the direction opposite to E. Solving eq. 4, we determine ρ and $R_{\rm H}$ as a function of n and the two scattering times (these bulky but analytical expressions are provided in Supplementary note 3). For each n, we again have only two unknowns ($\tau_{\rm eh}$ and τ) that fully define ρ and $R_{\rm H}$ whereas all the other relevant parameters are determined by graphene’s electronic spectrum. The resulting coupled nonlinear equations can be solved numerically, which has allowed us to obtain both scattering times as shown in Figs. 4a,b.

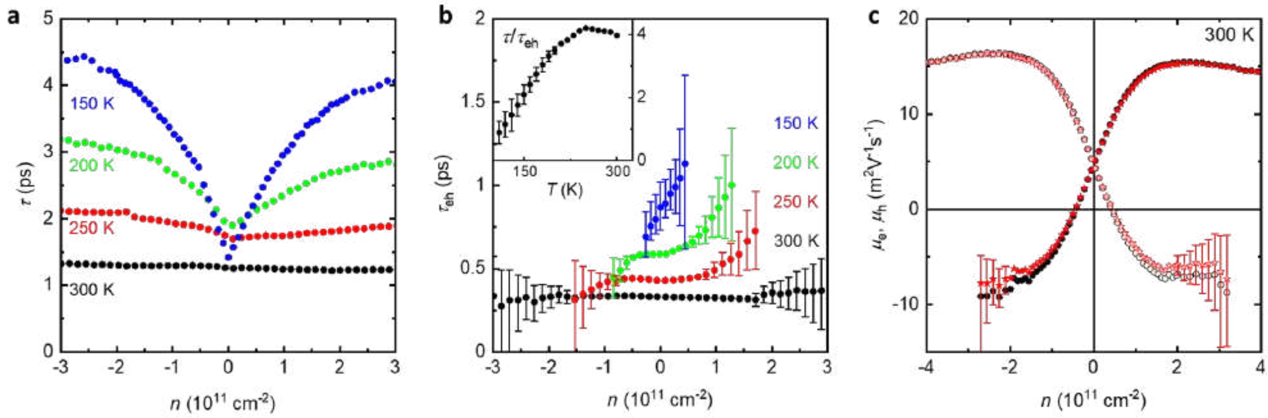


Figure 4| Scattering times in graphene's Dirac plasma and comparison between the Boltzmann and modified Dirac models. (**a**) Extracted $\tau(n)$ caused by impurities and phonons at different T (color coded). (**b**) Similarly for electron-hole scattering. Experimental errors rapidly increase away from the NP because a contribution of τ_{eh} towards the transport coefficients is exponentially small beyond a few n_{th}. The inset shows the temperature dependence of τ/τ_{eh} at the NP. (**c**) Comparison of the mobilities found using the modified Drude model (black curves, same as in Fig. 3a) and calculated from the Boltzmann model (red) using the scattering times from panels **a** and **b**.

Near room T, the extracted τ is practically independent of n. With decreasing T, τ starts exhibiting a dependence close to $\sqrt{n}$ which is expected for charged impurities and other mechanisms sensitive to screening by charge carriers. This square–root dependence yields a mobility independent of carrier density[29], typical of graphene at low T. At the NP, τ depends relatively weakly on T over our entire temperature range. Nonetheless, note that $\tau(n=0)$ first increases with increasing T, presumably due to stronger screening by the increasingly dense Dirac plasma. Then, above 200 K, τ decreases because of phonon scattering (Fig. 4a). As for τ_{eh}, it exhibits relatively weak dependence on doping (note that the prefactor in the third term of eq. 4 account for the n dependence of electron–hole friction caused by the varying mass densities). Some electron–hole asymmetry observed below 200 K (Fig. 4b) originates from subtly asymmetric $\rho(n)$ and $R_{\mathrm{H}}(n)$ found in the experiment, probably because of remnant doping. With increasing T, τ_{eh} evolves as Planckian scattering, that is, $\tau_{\mathrm{eh}} \approx h/Ck_{\mathrm{B}}T$ where $C \approx 0.6$, in good agreement with the coefficient reported previously[23]. Furthermore, the inset of Fig. 4b plots the T dependence of τ/τ_{eh} at the NP. It exhibits relatively little scatter thanks to the fact that $R_{\mathrm{H}}(n)$ depends only on the ratio τ/τ_{eh} rather than the individual times and is very sensitive to its absolute value, which minimizes errors in our numerical analysis (Supplementary note 3).

We emphasize that the ratio τ/τ_{eh} does not exceed 4 at any T, meaning that phonon and impurity scattering significantly affect electron–hole drag in the Dirac plasma, especially below 200 K. This bears ramifications for hydrodynamic description of the Dirac plasma. Indeed, to observe a viscous flow, it is imperative to have particle–particle scattering more frequent than momentum–relaxing scattering. The particle–particle scattering time τ_{v} that defines electron viscosity of the Dirac plasma is generally expected to be comparable to τ_{eh}. This means that, even under the most favorable conditions (close to room T), the ratio τ/τ_{v} near the neutrality point is modest (a factor of several at most), suppressing viscous effects, which agrees with recent observations[31,32]. At lower T where values of τ_{eh} and τ become close, it would be difficult, if not impossible, to observe even remnants of electron hydrodynamics.

Justifying the modified Drude model

It is instructive to calculate electron and hole mobilities from the scattering times found using the Boltzmann model (Supplementary note 3). The results are shown in Fig. 4c for the case of room T where our accuracy was highest because of the largest $\tau/\tau_{\mathrm{eh}} \approx 4$. As expected, the Boltzmann analysis also yields negative mobilities for minority carriers at $|n| > n_{\mathrm{th}}$ and saturating drift velocities in the same direction for both electrons and holes, if doping is larger than a few n_{th}. In the limit $\tau \to \infty$, both electrons and holes are expected to drift with the same velocity (Supplementary note 3). The finite τ/τ_{eh} reduces the drift velocity of minority carriers and, at room T, it is approximately twice smaller than the velocity of majority carriers. Although the modified Drude model does not distinguish between electron–hole and electron–phonon scattering, it agrees surprisingly well with the Boltzmann analysis. Notable deviations occur only for minority carriers and do not exceed ~20% (Fig. 4c). The agreement was found to be similar for all the studied devices at T above 150 K. This shows that, however empirical, the Drude model with different and sign–varying $\mu_{\mathrm{e}}(n)$ and $\mu_{\mathrm{h}}(n)$ can be used for a semi–quantitative description of the Dirac plasma in weak fields (R_{H} should remain linear in B; see Supplementary note 3). Furthermore, both Boltzmann and modified–Drude models describe equally well the measured dependence $\delta n_{\mathrm{H}}(T)$ shown in Fig. 2b (Supplementary Figs. 4-5).

Conclusion

Graphene's transport characteristics near the NP cannot possibly be understood without considering the strong interaction between the electron and hole subsystems within the Dirac plasma because minority carriers are dragged in the same direction as majority carriers. The observed behavior of both the longitudinal and Hall resistivities is accurately described by our Boltzmann analysis, which allows quantitative evaluation of the scattering rates. Inevitable scattering by phonons and impurities reduces the achievable value of the ratio τ/τ_{eh} so that the minority carriers in the Dirac plasma always lag behind majority ones. For high–quality encapsulated graphene, mutual drag is strongest near room T where the minority carriers drift at approximately half the velocity of majority carriers. This shows that impurity and phonon scattering significantly affects the transport properties of graphene's Dirac plasma and, in particular, suppresses its viscous (hydrodynamic) behavior[31-34].

References

1. Rojo, A. G. Electron–drag effects in coupled electron systems. *J. Phys: Cond. Matter* **11**, R31 (1999).
2. Narozhny, B. N. & Levchenko, A. Coulomb drag. *Rev. Mod. Phys.* **88**, 025003 (2016).
3. Sivan, U., Solomon, P. M. & Shtrikman, H. Coupled electron–hole transport. *Phys. Rev. Lett.* **68**, 1196 (1992).
4. Croxall, A. F. *et al*. Anomalous Coulomb drag in electron–hole bilayers. *Phys. Rev. Lett.* **101**, 246801 (2008).
5. Seamons, J. A., Morath, C. P., Reno, J. L. & Lilly, M. P. Coulomb drag in the exciton regime in electron–hole bilayers. *Phys. Rev. Lett.* **102**, 026804 (2009).
6. Kim, S. *et al.* Coulomb drag of massless fermions in graphene. *Phys. Rev. B* **83**, 161401 (2011).
7. Gorbachev, R. V. *et al.* Strong Coulomb drag and broken symmetry in double–layer graphene. *Nat. Phys.* **8**, 896–901 (2012).
8. Li, J. I. A. *et al*. Negative Coulomb drag in double bilayer graphene. *Phys. Rev. Lett.* **117**, 046802 (2016).
9. Lee, K. *et al*. Giant frictional drag in double bilayer graphene heterostructures. *Phys. Rev. Lett.* **117**, 046803 (2016).
10. Liu, X., Watanabe, K., Taniguchi, T., Halperin, B. I. & Kim, P. Quantum Hall drag of exciton condensate in graphene. *Nat. Phys.* **13**, 746–750 (2017).
11. Bandurin, D. A. *et al*. Interlayer electron–hole friction in tunable twisted bilayer graphene semimetal. *Phys. Rev. Lett.* **129**, 206802 (2022).
12. Titov, M. *et al.* Giant Magnetodrag in Graphene at Charge Neutrality. *Phys. Rev. Lett.* **111** (2013).
13. Fritz, L., Schmalian, J., Müller, M. & Sachdev, S. Quantum critical transport in clean graphene. *Phys. Rev. B* **78**, 085416 (2008).
14. Müller, M., Schmalian, J. & Fritz, L. Graphene: A Nearly Perfect Fluid. *Phys. Rev. Lett.* **103**, 025301 (2009).
15. Svintsov, D., Vyurkov, V., Yurchenko, S., Otsuji, T. & Ryzhii, V. Hydrodynamic model for electron–hole plasma in graphene. *J. Appl. Phys.* **111**, 083715 (2012).

16. Lucas, A. & Fong, K. C. Hydrodynamics of electrons in graphene. *J. Phys. Cond. Matt.* **30**, 053001 (2018).
17. Crossno, J. *et al.* Observation of the Dirac fluid and the breakdown of the Wiedemann–Franz law in graphene. *Science* **351**, 1058–1061 (2016).
18. Nam, Y., Ki, D.-K., Soler-Delgado, D. & Morpurgo, A. F. Electron–hole collision limited transport in charge-neutral bilayer graphene. *Nat. Phys.* **13**, 1207–1214 (2017).
19. Gallagher, P. *et al.* Quantum–critical conductivity of the Dirac fluid in graphene. *Science* **364**, 158 (2019).
20. Ku, M. J. H. *et al.* Imaging viscous flow of the Dirac fluid in graphene. *Nature* **583**, 537–541 (2020).
21. Block, A. *et al.* Observation of giant and tunable thermal diffusivity of a Dirac fluid at room temperature. *Nat. Nanotechnol.* **16**, 1195–1200 (2021).
22. Tan, C. *et al.* Dissipation–enabled hydrodynamic conductivity in a tunable bandgap semiconductor. *Sci. Adv.* **8**, eabi8481 (2022).
23. Xin, N. *et al.* Giant magnetoresistance of Dirac plasma in high–mobility graphene. *Nature* **616**, 270 (2023).
24. Hartnoll, S. A. & Mackenzie, A. P. Planckian dissipation in metals. *Rev. Mod. Phys.* **94**, 041002 (2022).
25. Phillips, P. W., Hussey, N. E. & Abbamonte, P. Stranger than metals. *Science* **377**, eabh4273 (2022).
26. Sarri, G. *et al.* Generation of neutral and high–density electron–positron pair plasmas in the laboratory. *Nat. Commun.* **6**, 6747 (2015).
27. Alekseev, P. S. *et al.* Magnetoresistance of compensated semimetals in confined geometries. *Phys. Rev. B* **95**, 20 (2017)
28. Ziman, J. M. *Electrons and phonons: the theory of transport phenomena in solids*. (Clarendon Press, 2007).
29. Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–162 (2009).
30. Poncé, S., Li, W., Reichardt, S. & Giustino, F. First–principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two–dimensional materials. *Rep. Prog. Phys.* **83**, 036501 (2020).
31. Jenkins, A. *et al*. Imaging the breakdown of ohmic transport in graphene. *Phys. Rev. Lett.* **129**, 087701 (2022).
32. Palm, M. L. *et al*. Observation of current whirlpools in graphene at room temperature. *Science* **384**, 465 (2024).
33. Talanov, A. *et al*. Observation of electronic viscous dissipation in graphene magneto-thermal transport. arXiv:2406.13799.
34. Zeng, Y. *et al*. Quantitative measurement of viscosity in two-dimensional electron fluids. arXiv:2407.05026.

Gilbert damping in two-dimensional metallic anti-ferromagnets

R. J. Sokolewicz,[1, 2] M. Baglai,[3] I. A. Ado,[1] M. I. Katsnelson,[1] and M. Titov[1]
[1] *Radboud University, Institute for Molecules and Materials, 6525 AJ Nijmegen, the Netherlands*
[2] *Qblox, Delftechpark 22, 2628 XH Delft, the Netherlands*
[3] *Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20, Uppsala, Sweden*
(Dated: March 29, 2024)

A finite spin life-time of conduction electrons may dominate Gilbert damping of two-dimensional metallic anti-ferromagnets or anti-ferromagnet/metal heterostructures. We investigate the Gilbert damping tensor for a typical low-energy model of a metallic anti-ferromagnet system with honeycomb magnetic lattice and Rashba spin-orbit coupling for conduction electrons. We distinguish three regimes of spin relaxation: exchange-dominated relaxation for weak spin-orbit coupling strength, Elliot-Yafet relaxation for moderate spin-orbit coupling, and Dyakonov-Perel relaxation for strong spin-orbit coupling. We show, however, that the latter regime takes place only for the in-plane Gilbert damping component. We also show that anisotropy of Gilbert damping persists for any finite spin-orbit interaction strength provided we consider no spatial variation of the Néel vector. Isotropic Gilbert damping is restored only if the electron spin-orbit length is larger than the magnon wavelength. Our theory applies to $MnPS_3$ monolayer on Pt or to similar systems.

I. INTRODUCTION

Magnetization dynamics in anti-ferromagnets continue to attract a lot of attention in the context of possible applications[1–4]. Various proposals utilize the possibility of THz frequency switching of anti-ferromagnetic domains for ultrafast information storage and computation[5,6]. The rise of van der Waals magnets has had a further impact on the field due to the possibility of creating tunable heterostructures that involve anti-ferromagnet and semiconducting layers[7].

Understanding relaxation of both the Néel vector and non-equilibrium magnetization in anti-ferromagnets is recognized to be of great importance for the functionality of spintronic devices[8–13]. On one hand, low Gilbert damping must generally lead to better electric control of magnetic order via domain wall motion or ultrafast domain switching[14–16]. On the other hand, an efficient control of magnetic domains must generally require a strong coupling between charge and spin degrees of freedom due to a strong spin-orbit interaction, that is widely thought to be equivalent to strong Gilbert damping.

In this paper, we focus on a microscopic analysis of Gilbert damping due to Dyakonov-Perel and Elliot-Yafet mechanisms. We apply the theory to a model of a two-dimensional Néel anti-ferromagnet with a honeycomb magnetic lattice.

Two-dimensional magnets typically exhibit either easy-plane or easy-axis anisotropy, and play crucial roles in stabilizing magnetism at finite temperatures[17,18]. Easy-axis anisotropy selects a specific direction for magnetization, thereby defining an axis for the magnetic order. In contrast, easy-plane anisotropy does not select a particular in-plane direction for the Néel vector, allowing it to freely rotate within the plane. This situation is analogous to the XY model, where the system's continuous symmetry leads to the suppression of out-of-plane fluctuations rather than fixing the magnetization in a specific in-plane direction[19,20]. Without this anisotropy, the magnonic fluctuations in a two-dimensional crystal can grow uncontrollably large to destroy any long-range magnetic order, according to the Mermin-Wagner theorem[21].

Recent density-functional-theory calculations for single-layer transition metal trichalgenides[22], predict the existence of a large number of metallic anti-ferromagnets with honeycomb lattice and different types of magnetic order as shown in Fig. 1. Many of these crystals may have the Néel magnetic order as shown in Fig. 1a and are metallic: $FeSiSe_3$, $FeSiTe_3$, $VGeTe_3$, $MnGeS_3$, $FeGeSe_3$, $FeGeTe_3$, $NiGeSe_3$, $MnSnS_3$, $MnSnS_3$, $MnSnSe_3$, $FeSnSe_3$, $NiSnS_3$. Apart from that it has been predicted that anti-ferromagnetism can be induced in graphene by bringing it in proximity to $MnPSe_3$[23] or by bringing it in double proximity between a layer of $Cr_2Ge_2Te_6$ and WS_2[24].

Partly inspired by these predictions and recent technological advances in producing single-layer anti-ferromagnet crystals, we propose an effective model to study spin relaxation in 2D honeycomb anti-ferromagnet with Néel magnetic order. The same system was studied by us in Ref. 25, where we found that spin-orbit coupling introduces a weak anisotropy in spin-orbit torque and electric conductivity. Strong spin-orbit coupling was shown to lead to a giant anisotropy of Gilbert damping.

Our analysis below is built upon the results of Ref. 25, and we investigate and identify three separate regimes of spin-orbit strength. Each regime is characterized by qualitatively different dependence of Gilbert damping on spin-orbit interaction and conduction electron transport time. The regime of weak spin-orbit interaction is dominated by exchange field relaxation of electron spin, and the regime of moderate spin-orbit strength is dominated by Elliot-Yafet spin relaxation. These two regimes are characterized also by a universal factor of 2 anisotropy of Gilbert damping. The regime of strong spin-orbit strength, which leads to substantial splitting of electron Fermi surfaces, is characterized by Dyakonov-Perel relaxation of the in-plane spin component and Elliot-Yafet re-

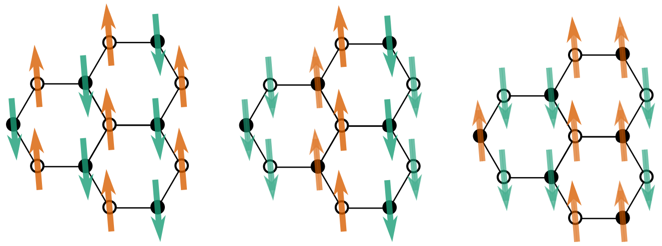


FIG. 1. Three anti-ferromagnetic phases commonly found among van-der-Waals magnets. Left-to-right: Néel, zig-zag, and stripy.

laxation of the perpendicular-to-the-plane Gilbert damping which leads to a giant damping anisotropy. Isotropic Gilbert damping is restored only for finite magnon wave vectors such that the magnon wavelength is smaller than the spin-orbit length.

Gilbert damping in a metallic anti-ferromagnet can be qualitatively understood in terms of the Fermi surface breathing[26]. A change in the magnetization direction gives rise to a change in the Fermi surface to which the conduction electrons have to adjust. This electronic reconfiguration is achieved through the scattering of electrons off impurities, during which angular momentum is transferred to the lattice. Gilbert damping, then, should be proportional to both (i) the ratio of the spin life-time and momentum life-time of conduction electrons, and (ii) the electric conductivity. Keeping in mind that the conductivity itself is proportional to momentum life-time, one may conclude that the Gilbert damping is linearly proportional to the spin life-time of conduction electrons. At the same time, the spin life-time of localized spins is *inversely* proportional to the spin life-time of conduction electrons. A similar relation between the spin life-times of conduction and localized electrons also holds for relaxation mechanisms that involve electron-magnon scattering[27].

Our approach formally decomposes the magnetic system into a classical sub-system of localized magnetic moments and a quasi-classical subsystem of conduction electrons. A local magnetic exchange couples these subsystems. Localized magnetic moments in transition-metal chalcogenides and halides form a hexagonal lattice. Here we focus on the Néel type anti-ferromagnet that is illustrated in Fig. 1a. In this case, one can define two sub-lattices A and B that host local magnetic moments $\mathbf{S}^{\mathrm{A}}$ and $\mathbf{S}^{\mathrm{B}}$, respectively. For the discussion of Gilbert damping, we ignore the weak dependence of both fields on atomic positions and assume that the modulus $S = |\mathbf{S}^{\mathrm{A(B)}}|$ is time-independent.

Under these assumptions, the magnetization dynamics of localized moments may be described in terms of two fields

$$\mathbf{m} = \frac{1}{2S}\big(\mathbf{S}^{\mathrm{A}} + \mathbf{S}^{\mathrm{B}}\big), \quad \mathbf{n} = \frac{1}{2S}\big(\mathbf{S}^{\mathrm{A}} - \mathbf{S}^{\mathrm{B}}\big), \tag{1}$$

which are referred to as the magnetization and staggered magnetization (or Néel vector), respectively. Within the mean-field approach, the vector fields yield the equations of motion

$$\dot{\mathbf{n}} = -J\,\mathbf{n}\times\mathbf{m} + \mathbf{n}\times\delta\mathbf{s}^{+} + \mathbf{m}\times\delta\mathbf{s}^{-}, \tag{2a}$$

$$\dot{\mathbf{m}} = \mathbf{m}\times\delta\mathbf{s}^{+} + \mathbf{n}\times\delta\mathbf{s}^{-}, \tag{2b}$$

where dot stands for the time derivative, while $\delta\mathbf{s}^{+}$ and $\delta\mathbf{s}^{-}$ stand for the mean staggered and non-staggered non-equilibrium fields that are proportional to the variation of the corresponding spin-densities of conduction electrons caused by the time dynamics of $\mathbf{n}$ and $\mathbf{m}$ fields. The energy J is proportional to the anti-ferromagnet exchange energy for localized momenta.

In Eqs. (2) we have omitted terms that are proportional to easy axis anisotropy for the sake of compactness. These terms are, however, important and will be introduced later in the text.

In the framework of Eqs. (2) the Gilbert damping can be computed as the linear response of the electron spin-density variation to a time change in both the magnetization and the Néel vector (see e. g. Refs.[25,28,29]).

In this definition, Gilbert damping describes the relaxation of localized spins by transferring both total and staggered angular momenta to the lattice by means of conduction electron scattering off impurities. Such a transfer is facilitated by spin-orbit interaction.

The structure of the full Gilbert damping tensor can be rather complicated as discussed in Ref. 25. However, by taking into account easy axis or easy plane anisotropy we may reduce the complexity of relevant spin configurations to parameterize

$$\delta\mathbf{s}^{+} = \alpha_m^{\parallel}\dot{\mathbf{m}}_{\parallel} + \alpha_m^{\perp}\dot{\mathbf{m}}_{\perp} + \alpha_m\mathbf{n}_{\parallel}\times(\mathbf{n}_{\parallel}\times\dot{\mathbf{m}}_{\parallel}), \tag{3a}$$

$$\delta\mathbf{s}^{-} = \alpha_n^{\parallel}\dot{\mathbf{n}}_{\parallel} + \alpha_n^{\perp}\dot{\mathbf{n}}_{\perp} + \alpha_n\mathbf{n}_{\parallel}\times(\mathbf{n}_{\parallel}\times\dot{\mathbf{n}}_{\parallel}), \tag{3b}$$

where the superscripts $\parallel$ and $\perp$ refer to the in-plane and perpendicular-to-the-plane projections of the corresponding vectors, respectively. The six coefficients $\alpha_m^{\parallel}$, $\alpha_m^{\perp}$, α_m, $\alpha_n^{\parallel}$, $\alpha_n^{\perp}$, and α_n parameterize the Gilbert damping.

Inserting Eqs. (3) into the equations of motion of Eqs. (2) produces familiar Gilbert damping terms. The damping proportional to time-derivatives of the Néel vector $\mathbf{n}$ is in general many orders of magnitude smaller than that proportional to the time-derivatives of the magnetization vector $\mathbf{m}$[25,30]. Due to the same reason, the higher harmonics term $\alpha_m\mathbf{n}_{\parallel}\times(\mathbf{n}_{\parallel}\times\partial_t\mathbf{m}_{\parallel})$ can often be neglected.

Thus, in the discussion below we may focus mostly on the coefficients $\alpha_m^{\parallel}$ and $\alpha_m^{\perp}$ that play the most important role in the magnetization dynamics of our system. The terms proportional to the time-derivative of $\mathbf{n}$ correspond to the transfer of angular momentum between the sub-lattices and are usually less relevant. We refer to the results of Ref. 25 when discussing these terms.

All Gilbert damping coefficients are intimately related to the electron spin relaxation time. The latter is relatively well understood in non-magnetic semiconductors with spin-orbital coupling. When a conducting electron moves in a steep potential it feels an effective magnetic field caused by relativistic effects. Thus, in a disordered system, the electron spin is subject to a random magnetic field each time it scatters off an impurity. At the same time, an electron also experiences precession around an effective spin-orbit field when it moves in between the collisions. Changes in spin direction *between* collisions are referred to as Dyakonov-Perel relaxation[31,32], while changes in spin-direction *during* collisions are referred to as Elliot-Yafet relaxation[33,34].

The spin-orbit field in semiconductors induces a characteristic frequency of spin precession $\Omega_{\rm s}$, while scalar disorder leads to a finite transport time τ of the conducting electrons. One may, then, distinguish two limits: (i) $\Omega_{\rm s}\tau \ll 1$ in which case the electron does not have sufficient time to change its direction between consecutive scattering events (Elliot-Yafet relaxation), and (ii) $\Omega_{\rm s}\tau \gg 1$ in which case the electron spin has multiple precession cycles in between the collisions (Dyakonov-Perel relaxation).

The corresponding processes define the so-called spin relaxation time, τ_s. In a 2D system the spin life-time $\tau_s^{\parallel}$, for the in-plane spin components, appears to be double the size of the life-time of the spin component that is perpendicular to the plane, $\tau_s^{\perp}$[32]. This geometric effect has largely been overlooked. For non-magnetic 2D semiconductor one can estimate[35,36]

$$\frac{1}{\tau_s^{\parallel}} \sim \begin{cases} \Omega_s^2\tau, & \Omega_{\rm s}\tau \ll 1 \\ 1/\tau, & \Omega_{\rm s}\tau \gg 1 \end{cases}, \qquad \tau_s^{\parallel} = 2\tau_s^{\perp}. \tag{4}$$

A pedagogical derivation and discussion of Eq. 4 can be found in Refs. 35 and 36. Because electrons are confined in two dimensions the random spin-orbit field is always directed in-plane, which leads to a decrease in the in-plane spin-relaxation rate by a factor of two compared to the out-of-plane spin-relaxation rate as demonstrated first in Ref. 32 (see Refs. 36–40 as well). The reason is that the perpendicular-to-the-plane component of spin is influenced by two components of the randomly changing magnetic field, i. e. x and y, whereas the parallel-to-the-plane spin components are only influenced by a single component of the fluctuating fields, i. e. the x spin projection is influenced only by the y component of the field and vice-versa. The argument has been further generalized in Ref. 25 to the case of strongly separated spin-orbit split Fermi surfaces. In this limit, the perpendicular-to-the-plane spin-flip processes on scalar disorder potential become fully suppressed. As a result, the perpendicular-to-the-plane spin component becomes nearly conserved, which results in a giant anisotropy of Gilbert damping in this regime.

In magnetic systems that are, at the same time, conducting there appears to be at least one additional energy scale, $\Delta_{\rm sd}$, that characterizes exchange coupling of conduction electron spin to the average magnetic moment of localized electrons. (In the case of s-d model description it is the magnetic exchange between the spin of conduction s electron and the localized magnetic moment of d or f electron on an atom.) This additional energy scale complicates the simple picture of Eq. (4) especially in the case of an anti-ferromagnet. The electron spin precession is now defined not only by spin-orbit field but also by $\Delta_{\rm sd}$. As the result the conditions $\Omega_{\rm s}\tau \ll 1$ and $\Delta_{\rm sd}\tau \gg 1$ may easily coexist. This dissolves the distinction between Elliot-Yafet and Dyakonov-Perel mechanisms of spin relaxation. One may, therefore, say that both Elliot-Yafet and Dyakonov-Perel mechanisms may act simultaneously in a typical 2D metallic magnet with spin-orbit coupling. The Gilbert damping computed from the microscopic model that we formulate below will always contain both contributions to spin-relaxation.

II. MICROSCOPIC MODEL AND RESULTS

The microscopic model that we employ to calculate Gilbert damping is the so-called s–d model that couples localized magnetic momenta $\mathbf{S}^{\rm A}$ and $\mathbf{S}^{\rm B}$ and conducting electron spins via the local magnetic exchange $\Delta_{\rm sd}$. Our effective low-energy Hamiltonian for conduction electrons reads

$$H = v_f\, \mathbf{p}\cdot\boldsymbol{\Sigma} + \frac{\lambda}{2}\left[\boldsymbol{\sigma}\times\boldsymbol{\Sigma}\right]_z - \Delta_{\rm sd}\,\mathbf{n}\cdot\boldsymbol{\sigma}\,\Sigma_z\Lambda_z + V(\mathbf{r}), \tag{5}$$

where the vectors $\boldsymbol{\Sigma}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\Lambda}$ denote the vectors of Pauli matrices acting on sub-lattice, spin and valley space, respectively. We also introduce the Fermi velocity v_f, Rashba-type spin-orbit interaction λ, and a random impurity potential $V(\mathbf{r})$.

The Hamiltonian of Eq. (5) can be viewed as the graphene electronic model where conduction electrons have 2D Rashba spin-orbit coupling and are also coupled to anti-ferromagnetically ordered classical spins on the honeycomb lattice.

The coefficients $\alpha_m^{\parallel}$ and $\alpha_m^{\perp}$ are obtained using linear response theory for the response of spin-density $\delta\mathbf{s}^{+}$ to the time-derivative of magnetization vector $\partial_t\mathbf{m}$. Impurity potential $V(\mathbf{r})$ is important for describing momentum relaxation to the lattice. This is related to the angular momentum relaxation due to spin-orbit coupling. The effect of random impurity potential is treated perturbatively in the (diffusive) ladder approximation that involves a summation over diffusion ladder diagrams. The details of the microscopic calculation can be found in the Appendices.

Before presenting the disorder-averaged quantities $\alpha_m^{\parallel,\perp}$, it is instructive to consider first the contribution to Gilbert damping originating from a small number of electron-impurity collisions. This clarifies how the number of impurity scattering effects will affect the final result.

Let us annotate the Gilbert damping coefficients with an additional superscript (l) that denotes the number of scattering events that are taken into account. This

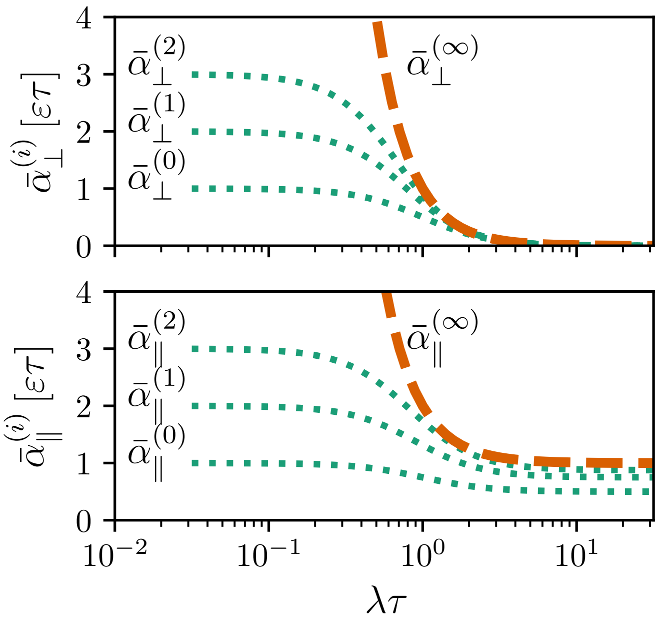


FIG. 2. Gilbert damping in the limit $\Delta_{\rm sd}=0$. Dotted (green) lines correspond to the results of the numerical evaluation of $\bar{\alpha}_{m,\perp,\parallel}^{(l)}$ for $l=0,1,2$ as a function of the parameter $\lambda\tau$. The dashed (orange) line corresponds to the diffusive (fully vertex corrected) results for $\bar{\alpha}_m^{\perp,\parallel}$.

means, in the diagrammatic language, that the corresponding quantity is obtained by summing up the ladder diagrams with $\leq l$ disorder lines. Each disorder line corresponds to a quasi-classical scattering event from a single impurity. The corresponding Gilbert damping coefficient is, therefore, obtained in the approximation where conduction electrons have scattered at most l number of times before releasing their non-equilibrium magnetic moment into a lattice.

To make final expressions compact we define the dimensionless Gilbert damping coefficients $\bar{\alpha}_m^{\parallel,\perp}$ by extracting the scaling factor

$$\alpha_m^{\parallel,\perp}=\frac{\mathcal{A}\Delta_{\rm sd}^2}{\pi\hbar^2 v_f^2 S}\,\bar{\alpha}_m^{\parallel,\perp}, \tag{6}$$

where $\mathcal{A}$ is the area of the unit cell, v_f is the Fermi velocity of the conducting electrons and $\hbar=h/2\pi$ is the Planck's constant. We also express the momentum scattering time τ in inverse energy units, $\tau\to\hbar\tau$.

Let us start by computing the coefficients $\bar{\alpha}_m^{\parallel,\perp(l)}$ in the formal limit $\Delta_{\rm sd}\to 0$. We can start with the "bare bubble" contribution which describes spin relaxation without a single scattering event. The corresponding results read

$$\bar{\alpha}_{m,\perp}^{(0)}=\varepsilon\tau\frac{1-\lambda^2/4\varepsilon^2}{1+\lambda^2\tau^2}, \tag{7a}$$

$$\bar{\alpha}_{m,\parallel}^{(0)}=\varepsilon\tau\left(\frac{1+\lambda^2\tau^2/2}{1+\lambda^2\tau^2}-\frac{\lambda^2}{8\varepsilon^2}\right), \tag{7b}$$

where ε denotes the Fermi energy which we consider positive (electron-doped system).

In all realistic cases, we have to consider $\lambda/\varepsilon\ll 1$, while the parameter $\lambda\tau$ may in principle be arbitrary. For $\lambda\tau\ll 1$ the disorder-induced broadening of the electron Fermi surfaces exceeds the spin-orbit induced splitting. In this case one basically finds no anisotropy of "bare" damping: $\bar{\alpha}_{m,\perp}^{(0)}=\bar{\alpha}_{m,\parallel}^{(0)}$. In the opposite limit of substantial spin-orbit splitting one gets an ultimately anisotropic damping $\bar{\alpha}_{m,\perp}^{(0)}\ll\bar{\alpha}_{m,\parallel}^{(0)}$. This asymptotic behavior can be summarized as

$$\bar{\alpha}_{m,\perp}^{(0)}=\varepsilon\tau\begin{cases}1 & \lambda\tau\ll 1,\\ (\lambda\tau)^{-2} & \lambda\tau\gg 1,\end{cases} \tag{8a}$$

$$\bar{\alpha}_{m,\parallel}^{(0)}=\varepsilon\tau\begin{cases}1 & \lambda\tau\ll 1,\\ \frac{1}{2}\left(1+(\lambda\tau)^{-2}\right) & \lambda\tau\gg 1,\end{cases} \tag{8b}$$

where we have used that $\varepsilon\gg\lambda$.

The results of Eq. (8) modify by electron diffusion. By taking into account up to l scattering events we obtain

$$\bar{\alpha}_{m,\perp}^{(l)}=\varepsilon\tau\begin{cases}l+\mathcal{O}(\lambda^2\tau^2) & \lambda\tau\ll 1,\\ (1+\delta_{l0})/(\lambda\tau)^2 & \lambda\tau\gg 1,\end{cases} \tag{9a}$$

$$\bar{\alpha}_{m,\parallel}^{(l)}=\varepsilon\tau\begin{cases}l+\mathcal{O}(\lambda^2\tau^2) & \lambda\tau\ll 1,\\ 1-(1/2)^{l+1}+\mathcal{O}((\lambda\tau)^{-2}) & \lambda\tau\gg 1,\end{cases} \tag{9b}$$

where we have used $\varepsilon\gg\lambda$ again.

From Eqs. (9) we see that the Gilbert damping for $\lambda\tau\ll 1$ gets an additional contribution of $\varepsilon\tau$ from each scattering event as illustrated numerically in Fig. 2. This leads to a formal divergence of Gilbert damping in the limit $\lambda\tau\ll 1$. While, at first glance, the divergence looks like a strong sensitivity of damping to impurity scattering, in reality, it simply reflects a diverging spin life-time. Once a non-equilibrium magnetization $\mathbf{m}$ is created it becomes almost impossible to relax it to the lattice in the limit of weak spin-orbit coupling. The formal divergence of $\alpha_m^{\perp}=\alpha_m^{\parallel}$ simply reflects the conservation law for electron spin polarization in the absence of spin-orbit coupling such that the corresponding spin life-time becomes arbitrarily large as compared to the momentum scattering time τ.

By taking the limit $l\to\infty$ (i. e. by summing up the entire diffusion ladder) we obtain compact expressions

$$\bar{\alpha}_m^{\perp}\equiv\bar{\alpha}_{m,\perp}^{(\infty)}=\varepsilon\tau\,\frac{1}{2\lambda^2\tau^2}, \tag{10a}$$

$$\bar{\alpha}_m^{\parallel}\equiv\bar{\alpha}_{m,\parallel}^{(\infty)}=\varepsilon\tau\,\frac{1+\lambda^2\tau^2}{\lambda^2\tau^2}, \tag{10b}$$

which assume $\bar{\alpha}_m^{\perp}\ll\bar{\alpha}_m^{\parallel}$ for $\lambda\tau\gg 1$ and $\bar{\alpha}_m^{\perp}=\bar{\alpha}_m^{\parallel}/2$ for $\lambda\tau\ll 1$. The factor of 2 difference that we observe when $\lambda\tau\ll 1$, corresponds to a difference in the electron spin life-times $\tau_s^{\perp}=\tau_s^{\parallel}/2$ that was discussed in the introduction[32].

Strong spin-orbit coupling causes a strong out-of-plane anisotropy of damping, $\bar{\alpha}_m^{\perp}\ll\bar{\alpha}_m^{\parallel}$ which corresponds to a suppression of the perpendicular-to-the-plane damping component. As a result, the spin-orbit interaction makes it much easier to relax the magnitude of the m_z component of magnetization than that of in-plane components.

Let us now turn to the dependence of $\bar{\alpha}_m$ coefficients on $\Delta_{\rm sd}$ that is illustrated numerically in Fig. 3. We consider first the case of absent spin-orbit coupling $\lambda = 0$. In this case, the combination of spin-rotational and sub-lattice symmetry (the equivalence of A and B sub-lattice) must make Gilbert damping isotropic (see e. g.[25,41]). The direct calculation for $\lambda = 0$ does, indeed, give rise to the isotropic result $\bar{\alpha}_m^{\perp} = \bar{\alpha}_m^{\parallel} = \varepsilon\tau(\varepsilon^2 + \Delta_{\rm sd}^2)/2\Delta_{\rm sd}^2$, which is, however, in contradiction to the limit $\lambda \to 0$ in Eq. (10).

At first glance, this contradiction suggests the existence of a certain energy scale for λ over which the anisotropy emerges. The numerical analysis illustrated in Fig. 4 reveals that this scale does not depend on the values of $1/\tau$, $\Delta_{\rm sd}$, or ε. Instead, it is defined solely by numerical precision. In other words, an isotropic Gilbert damping is obtained only when the spin-orbit strength λ is set below the numerical precision in our model. We should, therefore, conclude that the transition from isotropic to anisotropic (factor of 2) damping occurs exactly at $\lambda = 0$. Interestingly, the factor of 2 anisotropy is absent in Eqs. (8) and emerges only in the diffusive limit.

We will see below that this paradox can only be resolved by analyzing the Gilbert damping beyond the infinite wave-length limit.

One can see from Fig. 3 that the main effect of finite $\Delta_{\rm sd}$ is the regularization of the Gilbert damping divergency $(\lambda\tau)^{-2}$ in the limit $\lambda\tau \ll 1$. Indeed, the limit of weak spin-orbit coupling is non-perturbative for $\Delta_{\rm sd}/\varepsilon \ll \lambda\tau \ll 1$, while, in the opposite limit, $\lambda\tau \ll \Delta_{\rm sd}/\varepsilon \ll 1$, the results of Eqs. (10) are no longer valid. Assuming $\Delta_{\rm sd}/\varepsilon \ll 1$ we obtain the asymptotic expressions for the results presented in Fig. 3 as

$$\bar{\alpha}_m^{\perp} = \frac{1}{2}\varepsilon\tau \begin{cases} \frac{2}{3}\frac{\varepsilon^2+\Delta_{\rm sd}^2}{\Delta_{\rm sd}^2} & \lambda\tau \ll \Delta_{\rm sd}/\varepsilon, \\ \frac{1}{\lambda^2\tau^2} & \lambda\tau \gg \Delta_{\rm sd}/\varepsilon, \end{cases} \tag{11a}$$

$$\bar{\alpha}_m^{\parallel} = \varepsilon\tau \begin{cases} \frac{2}{3}\frac{\varepsilon^2+\Delta_{\rm sd}^2}{\Delta_{\rm sd}^2} & \lambda\tau \ll \Delta_{\rm sd}/\varepsilon, \\ 1 + \frac{1}{\lambda^2\tau^2} & \lambda\tau \gg \Delta_{\rm sd}/\varepsilon, \end{cases} \tag{11b}$$

which suggest that $\bar{\alpha}_m^{\perp}/\bar{\alpha}_m^{\parallel} = 2$ for $\lambda\tau \ll 1$. In the opposite limit, $\lambda\tau \gg 1$, the anisotropy of Gilbert damping grows as $\bar{\alpha}_m^{\parallel}/\bar{\alpha}_m^{\perp} = 2\lambda^2\tau^2$.

The results of Eqs. (11) can also be discussed in terms of the electron spin life-time, $\tau_s^{\perp(\parallel)} = \bar{\alpha}_m^{\perp(\parallel)}/\varepsilon$. For the inverse in-plane spin life-time we find

$$\frac{1}{\tau_s^{\parallel}} = \begin{cases} 3\Delta_{\rm sd}^2/2\varepsilon^2\tau & \lambda\tau \ll \Delta_{\rm sd}/\varepsilon, \\ \lambda^2\tau & \Delta_{\rm sd}/\varepsilon \ll \lambda\tau \ll 1, \\ 1/\tau & 1 \ll \lambda\tau, \end{cases} \tag{12}$$

that, for $\Delta_{\rm sd} = 0$, is equivalent to the known result of Eq. (4). Indeed, for $\Delta_{\rm sd} = 0$, the magnetic exchange plays no role and one observes the cross-over from Elliot-Yafet ($\lambda\tau \ll 1$) to Dyakonov-Perel ($\lambda\tau \gg 1$) spin relaxation.

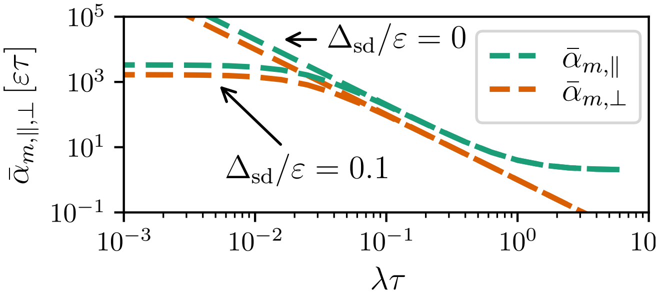


FIG. 3. Numerical results for the Gilbert damping components in the diffusive limit (vertex corrected)as the function of the spin-orbit coupling strength λ. The results correspond to $\varepsilon\tau = 50$ and $\Delta_{\rm sd}\tau = 0.1$ and agree with the asymptotic expressions of Eq. (11). Three different regimes can be distinguished for $\bar{\alpha}_m^{\parallel}$: i) spin-orbit independent damping $\bar{\alpha}_m^{\parallel} \propto \varepsilon^3\tau/\Delta_{\rm sd}^2$ for the exchange dominated regime, $\lambda\tau \ll \Delta_{\rm sd}/\varepsilon$, ii) the damping $\bar{\alpha}_m^{\parallel} \propto \varepsilon/\lambda^2\tau$ for Elliot-Yafet relaxation regime, $\Delta_{\rm sd}/\varepsilon \ll \lambda\tau \ll 1$, and iii) the damping $\bar{\alpha}_m^{\parallel} \propto \varepsilon\tau$ for the Dyakonov-Perel relaxation regime, $\lambda\tau \gg 1$. The latter regime is manifestly absent for $\bar{\alpha}_m^{\perp}$ in accordance with Eqs. (12,13).

This cross-over is, however, absent in the relaxation of the perpendicular spin component

$$\frac{1}{\tau_s^{\perp}} = 2 \begin{cases} 3\Delta_{\rm sd}^2/2\varepsilon^2\tau & \lambda\tau \ll \Delta_{\rm sd}/\varepsilon, \\ \lambda^2\tau & \Delta_{\rm sd}/\varepsilon \ll \lambda\tau, \end{cases} \tag{13}$$

where Elliot-Yafet-like relaxation extends to the regime $\lambda\tau \gg 1$.

As mentioned above, the factor of two anisotropy in spin-relaxation of $2D$ systems, $\tau_s^{\parallel} = 2\tau_s^{\perp}$, is known in the literature[32] (see Refs.[36–38] as well). Unlimited growth of spin life-time anisotropy, $\tau_s^{\parallel}/\tau_s^{\perp} = 2\lambda^2\tau^2$, in the regime $\lambda\tau \ll 1$ has been described first in Ref. 25. It can be qualitatively explained by a strong suppression of spin-flip processes for z spin component due to spin-orbit induced splitting of Fermi surfaces. The mechanism is effective only for scalar (non-magnetic) disorder. Even though such a mechanism is general for any magnetic or non-magnetic 2D material with Rashba-type spin-orbit coupling, the effect of the spin life-time anisotropy on Gilbert damping is much more relevant for anti-ferromagnets. Indeed, in an anti-ferromagnetic system the modulus of $\mathbf{m}$ is, by no means, conserved, hence the variations of perpendicular and parallel components of the magnetization vector are no longer related.

In the regime, $\lambda\tau \ll \Delta_{\rm sd}/\varepsilon$ the spin life-time is defined by exchange interaction and the distinction between Dyakonov-Perel and Elliot-Yafet mechanisms of spin relaxation is no longer relevant. In this regime, the spin-relaxation time is by a factor $(\varepsilon/\Delta_{\rm sd})^2$ larger than the momentum relaxation time.

Let us now return to the problem of emergency of the

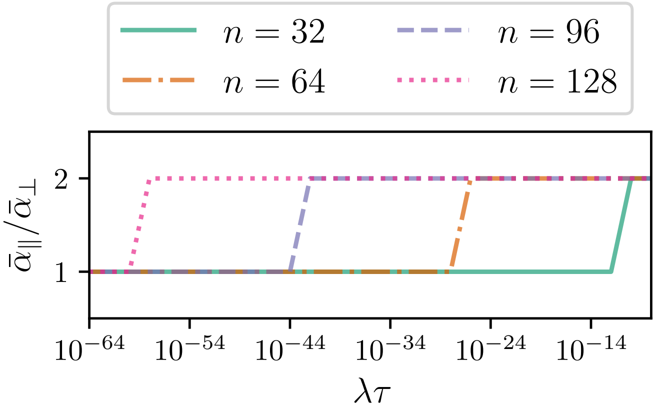


FIG. 4. Numerical evaluation of Gilbert damping anisotropy in the limit $\lambda \to 0$. Isotropic damping tensor is restored only if $\lambda = 0$ with ultimate numerical precision. The factor of 2 anisotropy emerges at any finite λ, no matter how small it is, and only depends on the numerical precision n, i.e. the number of digits contained in each variable during computation. The crossover from isotropic to anisotropic damping can be understood only by considering finite, though vanishingly small, magnon q vectors.

factor of 2 anisotropy of Gilbert damping at $\lambda = 0$. We have seen above (see Fig. 4) that, surprisingly, there exists no energy scale for the anisotropy to emerge. The transition from the isotropic limit ($\lambda = 0$) to a finite anisotropy appeared to take place exactly at $\lambda = 0$. We can, however, generalize the concept of Gilbert damping by considering the spin density response function at a finite wave vector $\mathbf{q}$.

To generalize the Gilbert damping, we are seeking a response of spin density at a point $\mathbf{r}$, $\delta\mathbf{s}_+(\mathbf{r})$ to a time derivative of magnetization vectors $\dot{\mathbf{m}}_\parallel$ and $\dot{\mathbf{m}}_\perp$ at the point $\mathbf{r}'$. The Fourier transform with respect to $\mathbf{r} - \mathbf{r}'$ gives the Gilbert damping for a magnon with the wave-vector $\mathbf{q}$.

The generalization to a finite $\mathbf{q}$-vector shows that the limits $\lambda \to 0$ and $q \to 0$ cannot be interchanged. When the limit $\lambda \to 0$ is taken before the limit $q \to 0$ one finds an isotropic Gilbert damping, while for the opposite order of limits, it becomes a factor of 2 anisotropic. In a realistic situation, the value of q is limited from below by an inverse size of a typical magnetic domain $1/L_{\mathrm{m}}$, while the spin-orbit coupling is effective on the length scale $L_\lambda = 2\pi\hbar v_f/\lambda$. In this picture, the isotropic Gilbert damping is characteristic for the case of sufficiently small domain size $L_{\mathrm{m}} \ll L_\lambda$, while the anisotropic Gilbert damping corresponds to the case $L_\lambda \ll L_m$.

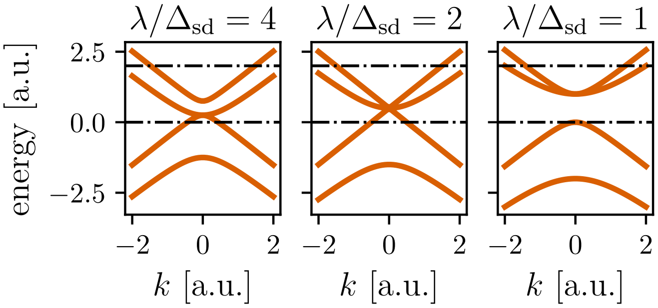


FIG. 5. Band-structure for the effective model of Eq. (5) in a vicinity of $\mathbf{K}$ valley assuming $n_z = 1$. Electron bands touch for $\lambda = 2\Delta_{\mathrm{sd}}$. The regime $\lambda \le 2\Delta_{\mathrm{sd}}$ corresponds to spin-orbit band inversion. The band structure in the valley $\mathbf{K}'$ is inverted. Our microscopic analysis is performed in the electron-doped regime for the Fermi energy above the gap as illustrated by the top dashed line. The bottom dashed line denotes zero energy (half-filling).

In the limit $q\ell \ll 1$, where $\ell = v_f\tau$ is the electron mean free path, we can summarize our results as

$$\bar{\alpha}^{\perp}_{m} = \varepsilon\tau \begin{cases} \frac{\varepsilon^2+\Delta^2_{\mathrm{sd}}}{2\Delta^2_{\mathrm{sd}}} & \lambda\tau \ll q\ell \ll \Delta_{\mathrm{sd}}/\varepsilon, \\ \frac{1}{3}\frac{\varepsilon^2+\Delta^2_{\mathrm{sd}}}{\Delta^2_{\mathrm{sd}}} & q\ell \ll \lambda\tau \ll \Delta_{\mathrm{sd}}/\varepsilon, \\ \frac{1}{2\lambda^2\tau^2} & \lambda\tau \gg \Delta_{\mathrm{sd}}/\varepsilon, \end{cases}, \tag{14a}$$

$$\bar{\alpha}^{\parallel}_{m} = \varepsilon\tau \begin{cases} \frac{\varepsilon^2+\Delta^2_{\mathrm{sd}}}{2\Delta^2_{\mathrm{sd}}} & \lambda\tau \ll q\ell \ll \Delta_{\mathrm{sd}}/\varepsilon, \\ \frac{2}{3}\frac{\varepsilon^2+\Delta^2_{\mathrm{sd}}}{\Delta^2_{\mathrm{sd}}} & q\ell \ll \lambda\tau \ll \Delta_{\mathrm{sd}}/\varepsilon, \\ 1+\frac{1}{\lambda^2\tau^2} & \lambda\tau \gg \Delta_{\mathrm{sd}}/\varepsilon, \end{cases} \tag{14b}$$

which represent a simple generalization of Eqs. (11).

The results of Eqs. (14) correspond to a simple behavior of Gilbert damping anisotropy,

$$\bar{\alpha}^{\parallel}_{m}/\bar{\alpha}^{\perp}_{m} = \begin{cases} 1 & \lambda\tau \ll q\ell, \\ 2\left(1+\lambda^2\tau^2\right) & q\ell \ll \lambda\tau, \end{cases} \tag{15}$$

where we still assume $q\ell \ll 1$.

III. ANTI-FERROMAGNETIC RESONANCE

The broadening of the anti-ferromagnet resonance peak is one obvious quantity that is sensitive to Gilbert damping. The broadening is however not solely defined by a particular Gilbert damping component but depends also on both magnetic anisotropy and anti-ferromagnetic exchange.

To be more consistent we can use the model of Eq. (5) to analyze the contribution of conduction electrons to an easy axis anisotropy. The latter is obtained by expanding the free energy for electrons in the value of n_z, which has a form $E = -Kn_z^2/2$. With the conditions $\varepsilon/\lambda \gg 1$ and $\varepsilon/\Delta_{\mathrm{sd}} \gg 1$ we obtain the anisotropy constant as

$$K = \frac{\mathcal{A}}{2\pi\hbar^2 v^2} \begin{cases} \Delta^2_{\mathrm{sd}}\lambda & 2\Delta_{\mathrm{sd}}/\lambda \le 1, \\ \Delta_{\mathrm{sd}}\lambda^2/2 & 2\Delta_{\mathrm{sd}}/\lambda \ge 1, \end{cases} \tag{16}$$

where $\mathcal{A}$ is the area of the unit cell. Here we assume both λ and $\Delta_{\rm sd}$ positive, therefore, the model naturally gives rise to an easy axis anisotropy with $K > 0$. In real materials, there exist other sources of easy axis or easy plane anisotropy. In-plane magneto-crystalline anisotropy also plays an important role. For example, Néel-type anti-ferromagnets with easy-axis anisotropy are $\mathrm{FePS_3}$, $\mathrm{FePSe_3}$ or $\mathrm{MnPS_3}$, whereas those with easy plane and in-plane magneto-crystalline anisotropy are $\mathrm{NiPS_3}$ and $\mathrm{MnPSe_3}$. Many of those materials are, however, Mott insulators. Our qualitative theory may still apply to materials like $\mathrm{MnPS_3}$ monolayers at strong electron doping.

The transition from $2\Delta_{\rm sd}/\lambda \geq 1$ to $2\Delta_{\rm sd}/\lambda \leq 1$ in Eq. (16) corresponds to the touching of two bands in the model of Eq. (5) as illustrated in Fig. 5.

Anti-ferromagnetic magnon frequency and life-time in the limit $q \to 0$ are readily obtained by linearizing the equations of motion

$$\dot{\mathbf{n}} = -J\,\mathbf{n}\times\mathbf{m} + K\,\mathbf{m}\times\mathbf{n}_\perp + \mathbf{n}\times(\hat{\alpha}_m\dot{\mathbf{m}})\,, \tag{17a}$$

$$\dot{\mathbf{m}} = K\,\mathbf{n}\times\mathbf{n}_\perp + \mathbf{n}\times(\hat{\alpha}_n\dot{\mathbf{n}})\,, \tag{17b}$$

where we took into account easy axis anisotropy K and disregarded irrelevant terms $\mathbf{m}\times(\hat{\alpha}_n\dot{\mathbf{n}})$ and $\mathbf{m}\times(\hat{\alpha}_m\dot{\mathbf{m}})$. We have also defined Gilbert damping tensors such as $\hat{\alpha}_m\dot{\mathbf{m}} = \alpha_m^{\parallel}\dot{\mathbf{m}}_{\parallel} + \alpha_m^{\perp}\dot{\mathbf{m}}_{\perp}$, $\hat{\alpha}_n\dot{\mathbf{n}} = \alpha_n^{\parallel}\dot{\mathbf{n}}_{\parallel} + \alpha_n^{\perp}\dot{\mathbf{n}}_{\perp}$.

In the case of easy axis anisotropy we can use the linearized modes $\mathbf{n} = \hat{\mathbf{z}} + \delta\mathbf{n}_{\parallel}\,e^{i\omega t}$, $\mathbf{m} = \delta\mathbf{m}_{\parallel}\,e^{i\omega t}$, hence we get the energy of $q = 0$ magnon as

$$\omega = \omega_0 - i\Gamma/2, \tag{18}$$

$$\omega_0 = \sqrt{JK}, \qquad \Gamma = J\alpha_n^{\parallel} + K\alpha_m^{\parallel} \tag{19}$$

where we took into account that $K \ll J$. The expression for ω_0 is well known due to Kittel and Keffer[42,43].

Using Ref. 25 we find out that $\alpha_n^{\parallel} \simeq \alpha_m^{\perp}(\lambda/\varepsilon)^2$ and $\alpha_n^{\perp} \simeq \alpha_m^{\parallel}(\lambda/\varepsilon)^2$, hence

$$\Gamma \simeq \alpha_m^{\parallel}\left(K + \frac{J/2}{\varepsilon^2/\lambda^2 + \varepsilon^2\tau^2}\right), \tag{20}$$

where we have simply used Eqs. (10). Thus, one may often ignore the contribution $J\alpha_n^{\parallel}$ as compared to $K\alpha_m^{\parallel}$ despite the fact that $K \ll J$.

In the context of anti-ferromagnets, spin-pumping terms are usually associated with the coefficients $\alpha_n^{\parallel}$ in Eq. (3b) that are not in the focus of the present study. Those coefficients have been analyzed for example in Ref. 25. In this manuscript we simply use the known results for α_n in Eqs. (17-19), where we illustrate the effect of both spin-pumping coefficient α_n and the direct Gilbert damping α_m on the magnon life time. One can see from Eqs. (19,20) that the spin-pumping contributions do also contribute, though indirectly, to the magnon decay. The spin pumping contributions become more important in magnetic materials with small magnetic anisotropy. The processes characterized by the coefficients α_n may also be interpreted in terms of angular momentum transfer from one AFM sub-lattice to another. In that respect, the spin pumping is specific to AFM, and is qualitatively different from the direct Gilbert damping processes (α_m) that describe the direct momentum relaxation to the lattice.

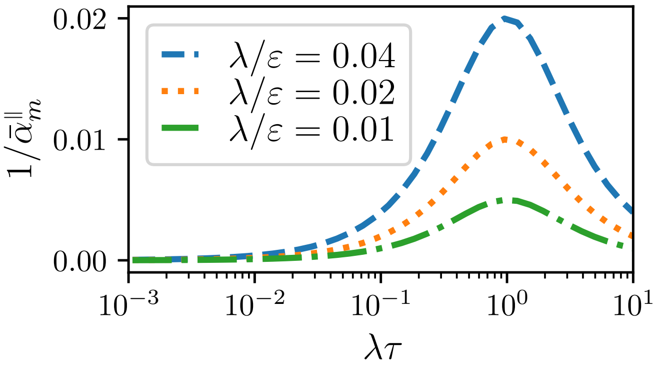


FIG. 6. Numerical evaluation of the inverse Gilbert damping $1/\bar{\alpha}_m^{\parallel}$ as a function of the momentum relaxation time τ. The inverse damping is peaked at $\tau \propto 1/\lambda$ which also corresponds to the maximum of the anti-ferromagnetic resonance quality factor in accordance with Eq. (21).

As illustrated in Fig. 6 the quality factor of the anti-ferromagnetic resonance (for a metallic anti-ferromagnet with easy-axis anisotropy) is given by

$$Q = \frac{\omega_0}{\Gamma} \simeq \frac{1}{\alpha_m^{\parallel}}\sqrt{\frac{J}{K}}. \tag{21}$$

Interestingly, the quality factor defined by Eq. (21) is maximized for $\lambda\tau \simeq 1$, i.e. for the electron spin-orbit length being of the order of the scattering mean free path.

The quantities $1/\sqrt{K}$ and $1/\bar{\alpha}_m^{\parallel}$ are illustrated in Fig. 6 from the numerical analysis. As one would expect, the quality factor vanishes in both limits $\lambda \to 0$ and $\lambda \to \infty$. The former limit corresponds to an overdamped regime hence no resonance can be observed. The latter limit corresponds to a constant $\alpha_m^{\parallel}$, but the resonance width Γ grows faster with λ than ω_0 does, hence the vanishing quality factor.

It is straightforward to check that the results of Eqs. (20,21) remain consistent when considering systems with either easy-plane or in-plane magneto-crystalline anisotropy. Thus, the coefficient $\alpha_m^{\perp}$ normally does not enter the magnon damping, unless the system is brought into a vicinity of spin-flop transition by a strong external field.

IV. CONCLUSION

In conclusion, we have analyzed the Gilbert damping tensor in a model of a two-dimensional anti-ferromagnet on a honeycomb lattice. We consider the damping mechanism that is dominated by a finite electron spin life-time due to a combination of spin-orbit coupling and impurity scattering of conduction electrons. In the case of a 2D electron system with Rashba spin-orbit coupling λ, the Gilbert damping tensor is characterized by two components $\alpha_m^{\parallel}$ and $\alpha_m^{\perp}$. We show that the anisotropy of Gilbert damping depends crucially on the parameter $\lambda\tau$, where τ is the transport scattering time for conduction electrons. For $\lambda\tau \ll 1$ the anisotropy is set by a geometric factor of 2, $\alpha_m^{\parallel} = 2\alpha_m^{\perp}$, while it becomes infinitely large in the opposite limit, $\alpha_m^{\parallel} = (\lambda\tau)^2\alpha_m^{\perp}$ for $\lambda\tau \gg 1$. Gilbert damping becomes isotropic exactly for $\lambda = 0$, or, strictly speaking, for the case $\lambda \ll \hbar v_f q$, where q is the magnon wave vector.

This factor of 2 is essentially universal, and is a geometric effect: the z-component relaxation results from fluctuations in two in-plane spin components, whereas in-plane relaxation stems from fluctuations of the z-component alone. This reflects the subtleties of our microscopic model, where the mechanism for damping is activated by the decay of conduction electron momenta, linked to spin-relaxation through spin-orbit interactions.

We find that Gilbert damping is insensitive to magnetic order for $\lambda \gg \Delta_{\rm sd}/\varepsilon\tau$, where $\Delta_{\rm sd}$ is an effective exchange coupling between spins of conduction and localized electrons. In this case, the electron spin relaxation can be either dominated by scattering (Dyakonov-Perel relaxation) or by spin-orbit precession (Elliot-Yafet relaxation). We find that the Gilbert damping component $\alpha_m^{\perp} \simeq \varepsilon/\lambda^2\tau$ is dominated by Elliot-Yafet relaxation irrespective of the value of the parameter $\lambda\tau$, while the other component crosses over from $\alpha_m^{\parallel} \simeq \varepsilon/\lambda^2\tau$ (Elliot-Yafet relaxation) for $\lambda\tau \ll 1$, to $\alpha_m^{\parallel} \simeq \varepsilon\tau$ (Dyakonov-Perel relaxation) for $\lambda\tau \gg 1$. For the case $\lambda \ll \Delta_{\rm sd}/\varepsilon\tau$ the spin relaxation is dominated by interaction with the exchange field.

Crucially, our results are not confined solely to the Néel order on the honeycomb lattice: we anticipate a broader applicability across various magnetic orders, including the zigzag order. This universality stems from our focus on the large magnon wavelength limit. The choice of the honeycomb lattice arises from its unique ability to maintain isotropic electronic spectra within the plane, coupled with the ability to suppress anisotropy concerning in-plane spin rotations. Strong anisotropic electronic spectra would naturally induce strong anisotropic in-plane Gilbert damping, which are absent in our results.

Finally, we show that the anti-ferromagnetic resonance width is mostly defined by $\alpha_m^{\parallel}$ and demonstrate that the resonance quality factor is maximized for $\lambda\tau \approx 1$. Our microscopic theory predictions may be tested for systems such as MnPS$_3$ monolayer on Pt and similar heterostructures.

ACKNOWLEDGMENTS

We are grateful to O. Gomonay, R. Duine, J. Sinova, and A. Mauri for helpful discussions. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 873028.

Appendix A: Microscopic framework

The microscopic model that we employ to calculate Gilbert damping belongs to a class of so-called s–d models that describe the physical system in the form of a Heisenberg model for localized spins and a tight-binding model for conduction electrons that are weakly coupled by a local magnetic exchange interaction of the strength $\Delta_{\rm sd}$.

Our effective electron Hamiltonian for a metallic hexagonal anti-ferromagnet is given by[25]

$$H_0 = v_f \mathbf{p}\cdot\mathbf{\Sigma} + \frac{\lambda}{2}\left[\boldsymbol{\sigma}\times\mathbf{\Sigma}\right]_z - \Delta_{\rm sd}\boldsymbol{n}\cdot\boldsymbol{\sigma}\,\Sigma_z\Lambda_z, \qquad \text{(A1)}$$

where the vectors $\mathbf{\Sigma}$, $\boldsymbol{\sigma}$ and $\mathbf{\Lambda}$ denote the vectors of Pauli-matrices acting on sub-lattice, spin and valley space respectively. We also introduce the Fermi velocity v_f, Rashba-type spin-orbit interaction λ.

To describe Gilbert damping of the localized field $\boldsymbol{n}$ we have to add the relaxation mechanism. This is provided in our model by adding a weak impurity potential $H = H_0 + V(\mathbf{r})$. The momentum relaxation due to scattering on impurities leads indirectly to the relaxation of Heisenberg spins due to the presence of spin-orbit coupling and exchange couplings.

For modeling the impurity potential, we adopt a delta-correlated random potential that corresponds to the point scatter approximation, where the range of the impurity potential is much shorter than that of the mean free path (see e.g. section 3.8 of Ref. 44), i.e.

$$\langle V(\mathbf{r})V(\mathbf{r}')\rangle = 2\pi\alpha(\hbar v_f)^2\delta(\mathbf{r}-\mathbf{r}'), \qquad \text{(A2)}$$

where the dimensionless coefficient $\alpha \ll 1$ characterizes the disorder strength. The corresponding scattering time for electrons is obtained as $\tau = \hbar/\pi\alpha\epsilon$, which is again similar to the case of graphene.

The response of symmetric spin-polarization $\delta\boldsymbol{s}^+$ to the time-derivative of non-staggered magnetization, $\partial_t\boldsymbol{m}$, is defined by the linear relation

$$\delta s_\alpha^+ = \sum_\beta \mathcal{R}_{\alpha\beta}\big|_{\omega=0}\,\dot{m}_\beta, \qquad \text{(A3)}$$

where the response tensor is taken at zero frequency[25,45]. The linear response is defined generally by the tensor

$$\mathcal{R}_{\alpha\beta} = \frac{\mathcal{A}\Delta_{\rm sd}^2}{2\pi S}\int\frac{\mathrm{d}\boldsymbol{p}}{(2\pi\hbar)^2}\left\langle \mathrm{Tr}\left[G^{\rm R}_{\varepsilon,\mathbf{p}}\sigma_\alpha G^{\rm A}_{\varepsilon+\hbar\omega,\mathbf{p}}\sigma_\beta\right]\right\rangle, \qquad \text{(A4)}$$

where $G^{\mathrm{R(A)}}_{\varepsilon,\boldsymbol{p}}$ are standing for retarded(advanced) Green functions and the angular brackets denote averaging over disorder fluctuations.

The standard recipe for disorder averaging is the diffusive approximation[46,47] that is realized by replacing the bare Green functions in Eq. (A4) with disorder-averaged Green functions and by replacing one of the vertex operators σ_x or σ_y with the corresponding *vertex-corrected* operator that is formally obtained by summing up ladder impurity diagrams (diffusons).

In models with spin-orbit coupling, the controllable diffusive approximation for non-dissipative quantities may become, however, more involved as was noted first in Ref. 48. For Gilbert damping it is, however, sufficient to consider the ladder diagram contributions only.

The disorder-averaged Green function is obtained by including an imaginary part of the self-energy Σ^{R} (not to be confused here with the Pauli matrix $\Sigma_{0,x,y,z}$) that is evaluated in the first Born approximation

$$\operatorname{Im}\Sigma^{\mathrm{R}} = 2\pi\alpha v_f^2 \int \frac{\mathrm{d}\boldsymbol{p}}{(2\pi)^2}\,\operatorname{Im}\frac{1}{\varepsilon - H_0 + i0}. \qquad \text{(A5)}$$

The real part of the self-energy leads to the renormalization of the energy scales ε, λ and Δ_{sd}.

In the first Born approximation, the disorder-averaged Green function is given by

$$G^{\mathrm{R}}_{\varepsilon,\boldsymbol{p}} = \frac{1}{\varepsilon - H_0 - i\operatorname{Im}\Sigma^R}. \qquad \text{(A6)}$$

The vertex corrections are computed in the diffusive approximation. The latter involves replacing the vertex σ_α with the *vertex-corrected* operator,

$$\sigma^{\mathrm{vc}}_\alpha = \sum_{l=0}^{\infty}\sigma^{(l)}_\alpha, \qquad \text{(A7)}$$

where the index l corresponds to the number of disorder lines in the ladder.

The operators $\sigma^{(l)}_\alpha$ can be defined recursively as

$$\sigma^{(l)}_\alpha = \frac{2\hbar v_f^2}{\varepsilon\tau}\int\frac{\mathrm{d}\boldsymbol{p}}{(2\pi)^2} G^{\mathrm{R}}_{\varepsilon,\boldsymbol{p}}\sigma^{(l-1)}_\alpha G^{\mathrm{A}}_{\varepsilon+\hbar\omega,\boldsymbol{p}}, \qquad \text{(A8)}$$

where $\sigma^{(0)}_\alpha = \sigma_\alpha$.

The summation in Eq. (A7) can be computed in the full operator basis, $B_{i=\{\alpha,\beta,\gamma\}} = \sigma_\alpha\Sigma_\beta\Lambda_\gamma$, where each index α, β and γ takes on 4 possible values (with zero standing for the unity matrix). We may always normalize $\operatorname{Tr} B_i B_j = 2\delta_{ij}$ in an analogy to the Pauli matrices. The operators B_i are, then, forming a finite-dimensional space for the recursion of Eq. (A8).

The vertex-corrected operators B^{vc}_i are obtained by summing up the matrix geometric series

$$B^{\mathrm{vc}}_i = \sum_j \left(\frac{1}{1-\mathcal{F}}\right)_{ij} B_j, \qquad \text{(A9)}$$

where the entities of the matrix $\mathcal{F}$ are given by

$$\mathcal{F}_{ij} = \frac{\hbar v_f^2}{\varepsilon\tau}\int\frac{\mathrm{d}\boldsymbol{p}}{(2\pi)^2}\operatorname{Tr}\left[G^{\mathrm{R}}_{\varepsilon,\boldsymbol{p}} B_i G^{\mathrm{A}}_{\varepsilon+\hbar\omega,\boldsymbol{p}} B_j\right]. \qquad \text{(A10)}$$

Our operators of interest σ_x and σ_y can always be decomposed in the operator basis as

$$\sigma_\alpha = \frac{1}{2}\sum_i B_i\,\operatorname{Tr}(\sigma_\alpha B_i), \qquad \text{(A11)}$$

hence the vertex-corrected spin operator is given by

$$\sigma^{\mathrm{vc}}_\alpha = \frac{1}{2}\sum_{ij} B^{\mathrm{vc}}_i \operatorname{Tr}(\sigma_\alpha B_i). \qquad \text{(A12)}$$

Moreover, the computation of the entire response tensor of Eq. (A4) in the diffusive approximation can also be expressed via the matrix $\mathcal{F}$ as

$$\mathcal{R}_{\alpha\beta} = \frac{\alpha_0\varepsilon\tau}{8\hbar}\sum_{ij}\left[\operatorname{Tr}\sigma_\alpha B_i\right]\left[\frac{\mathcal{F}}{1-\mathcal{F}}\right]_{ij}\left[\operatorname{Tr}\sigma_\beta B_j\right], \qquad \text{(A13)}$$

where $\alpha_0 = \mathcal{A}\Delta^2_{\mathrm{sd}}/\pi\hbar^2 v_f^2 S$ is the coefficient used in Eq. (6) to define the unit of the Gilbert damping.

It appears that one can always choose the basis of B_i operators such that the computation of Eq. (A13) is closed in a subspace of just three B_i operators with $i = 1, 2, 3$. This enables us to make analytical computations of Eq. (A13).

Appendix B: Magnetization dynamics

The representation of the results can be made somewhat simpler by choosing x axis in the direction of the in-plane projection $\mathbf{n}_\parallel$ of the Néel vector, hence $n_y = 0$. In this case, one can represent the result as

$$\delta\mathbf{s}^+ = c_1\mathbf{n}_\parallel \times (\mathbf{n}_\parallel \times \partial_t\mathbf{m}_\parallel) + c_2\partial_t\mathbf{m}_\parallel + c_3\partial_t\mathbf{m}_\perp + c_4\mathbf{n},$$

where $\mathbf{n}$ dependence of the coefficients c_i may be parameterized as

$$c_1 = \frac{r_{11} - r_{22} - r_{31}(1-n_z^2)/(n_x n_z)}{1-n_z^2}, \qquad \text{(B1a)}$$

$$c_2 = r_{11} - r_{31}(1-n_z^2)/(n_x n_z), \qquad \text{(B1b)}$$

$$c_3 = r_{33}, \qquad \text{(B1c)}$$

$$c_4 = (r_{31}/n_z)\,\partial_t m_z + \zeta(\partial_t\mathbf{m})\cdot\mathbf{n}. \qquad \text{(B1d)}$$

The analytical results in the paper correspond to the evaluation of $\delta s^\pm$ up to the second order in Δ_{sd} using perturbative analysis. Thus, zero approximation corresponds to setting $\Delta_{\mathrm{sd}} = 0$ in Eqs. (A1,A5).

The equations of motion on $\mathbf{n}$ and $\mathbf{m}$ are given by Eqs. (2),

$$\partial_t\mathbf{n} = -J\,\mathbf{n}\times\mathbf{m} + \mathbf{n}\times\delta\mathbf{s}^+ + \mathbf{m}\times\delta\mathbf{s}^-, \qquad \text{(B2a)}$$

$$\partial_t\mathbf{m} = \mathbf{m}\times\delta\mathbf{s}^+ + \mathbf{n}\times\delta\mathbf{s}^-, \qquad \text{(B2b)}$$

It is easy to see that the following transformation leaves the above equations invariant,

$$\delta\mathbf{s}^+ \to \delta\mathbf{s}^+ - \xi\,\mathbf{n}, \qquad \delta\mathbf{s}^- \to \delta\mathbf{s}^- - \xi\,\mathbf{m}, \tag{B3}$$

for an arbitrary value of ξ.

Such a gauge transformation can be used to prove that the coefficient c_4 is irrelevant in Eqs. (B2).

In this paper, we compute $\delta\mathbf{s}^\pm$ to the zeroth order in $|\mathbf{m}|$ – the approximation which is justified by the sub-lattice symmetry in the anti-ferromagnet. A somewhat more general model has been analyzed also in Ref. 25 to which we refer the interested reader for more technical details.

Appendix C: Anisotropy constant

The anisotropy constant is obtained from the grand potential energy Ω for conducting electrons. For the model of Eq. (A1) the latter can be expressed as

$$\Omega = -\sum_{\varsigma=\pm} \frac{1}{\beta} \int \mathrm{d}\varepsilon\, g(\varepsilon)\nu_\varsigma(\varepsilon), \tag{C1}$$

where $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature, $\varsigma = \pm$ is the valley index (for the valleys $\boldsymbol{K}$ and $\boldsymbol{K}'$), $G^{\mathrm{R}}_{\varsigma,\boldsymbol{p}}$ is the bare retarded Green function with momentum $\boldsymbol{p}$ and in the valley ς. We have also defined the function

$$g(\varepsilon) = \ln\left(1 + \exp[\beta(\mu - \varepsilon)]\right), \tag{C2}$$

where μ is the electron potential, and the electron density of states in each of the valleys is given by,

$$\nu_\varsigma(\varepsilon) = \frac{1}{\pi} \int \frac{\mathrm{d}\mathbf{p}}{(2\pi\hbar)^2}\, \mathrm{Im}\,\mathrm{Tr}\, G^{\mathrm{R}}_{\varsigma,\mathbf{p}}, \tag{C3}$$

where the trace is taken only over spin and sub-lattice space,

In the metal regime considered, the chemical potential is assumed to be placed in the upper electronic band. In this case, the energy integration can be taken only for positive energies. The two valence bands are always filled and can only add a constant shift to the grand potential Ω that we disregard.

The evaluation of Eq. (C1) yields the following density of states

$$\nu_\tau(\varepsilon) = \frac{1}{2\pi\hbar^2 v_f^2} \begin{cases} 0 & 0 < \varepsilon < \varepsilon_2 \\ \varepsilon/2 + \lambda/4 & \varepsilon_2 < \varepsilon < \varepsilon_1, \\ \varepsilon & \varepsilon > \varepsilon_1, \end{cases} \tag{C4}$$

where the energies $\varepsilon_{1,2}$ correspond to the extremum points (zero velocity) for the electronic bands. These energies, for each of the valleys, are given by

$$\varepsilon_{1,\varsigma} = \frac{1}{2}\big(+ \lambda + \sqrt{4\Delta^2 + \lambda^2 - 4\varsigma\Delta\lambda n_z}\big), \tag{C5a}$$

$$\varepsilon_{2,\varsigma} = \frac{1}{2}\big(- \lambda + \sqrt{4\Delta^2 + \lambda^2 + 4\varsigma\Delta\lambda n_z}\big) \tag{C5b}$$

where $\varsigma = \pm$ is the valley index.

In the limit of zero temperature we can approximate Eq. (C1) as

$$\Omega = -\sum_{\varsigma=\pm} \frac{1}{\beta} \int_0^\infty \mathrm{d}\varepsilon\,(\mu - \varepsilon)\nu_\varsigma(\varepsilon). \tag{C6}$$

Then, with the help of Eq. (C1) we find,

$$\begin{aligned}\Omega = -\frac{1}{24\pi\hbar^2 v_f^2} \sum_{\varsigma=\pm} \big[&(\varepsilon_{1,\varsigma} - \mu)^2(4\varepsilon_{1,\varsigma} - 3\lambda + 2\mu) \\ &+(\varepsilon_{2,\varsigma} - \mu)^2(4\varepsilon_{2,\varsigma} + 3\lambda + 2\mu)\big]\,.\end{aligned} \tag{C7}$$

By substituting the results of Eqs. (C5) into the above equation we obtain

$$\begin{aligned}\Omega = -\frac{1}{24\pi\hbar^2 v_f^2} \Big[&(4\Delta^2 - 4n_z\Delta\lambda + \lambda^2)^{2/3} \\ &+(4\Delta^2 + 4n_z\Delta\lambda + \lambda^2)^{2/3} - 24\Delta\mu + 8\mu^3\Big]\,.\end{aligned} \tag{C8}$$

A careful analysis shows that the minimal energy corresponds to $n_z = \pm 1$ so that the conducting electrons prefer an easy-axis magnetic anisotropy. By expanding in powers of n_z^2 around $n_z = \pm 1$ we obtain $\Omega = -Kn_z^2/2$, where

$$K = \frac{1}{2\pi\hbar^2 v^2} \begin{cases} |\Delta^2\lambda| & |\lambda/2\Delta| \geq 1, \\ |\Delta\lambda^2|/2 & |\lambda/2\Delta| \leq 1. \end{cases} \tag{C9}$$

This provides us with the easy axis anisotropy of Eq. (16).

---

[1] S. A. Siddiqui, J. Sklenar, K. Kang, M. J. Gilbert, A. Schleife, N. Mason, and A. Hoffmann, Journal of Applied Physics **128**, 040904 (2020), https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/5.0009445/15249168/040904_1_online.pdf.

[2] V. V. Mazurenko, Y. O. Kvashnin, A. I. Lichtenstein, and M. I. Katsnelson, Journal of Experimental and Theoretical Physics **132**, 506 (2021).

[3] L. Šmejkal, J. Sinova, and T. Jungwirth, Phys. Rev. X **12**, 040501 (2022).

[4] B. A. Bernevig, C. Felser, and H. Beidenkopf, Nature **603**, 41 (2022).

[5] T. G. H. Blank, K. A. Grishunin, B. A. Ivanov, E. A. Mashkovich, D. Afanasiev, and A. V. Kimel, Phys. Rev. Lett. **131**, 096701 (2023).

[6] W. Wu, C. Yaw Ameyaw, M. F. Doty, and M. B. Jungfleisch, Journal of Applied Physics **130**, 091101 (2021), https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/5.0057536/13478272/091101_1_online.pdf.

[7] M. Gibertini, M. Koperski, A. F. Morpurgo, and K. S. Novoselov, Nature Nanotechnology **14**, 408 (2019).

[8] K. M. D. Hals, Y. Tserkovnyak, and A. Brataas, Physical Review Letters **106**, 107206 (2011), publisher: American Physical Society.

[9] R. Cheng, D. Xiao, and A. Brataas, Physical Review Letters **116**, 207603 (2016), publisher: American Physical Society.

[10] S. Urazhdin and N. Anthony, Physical Review Letters **99**, 046602 (2007), publisher: American Physical Society.

[11] R. Cheng, M. W. Daniels, J.-G. Zhu, and D. Xiao, Physical Review B **91**, 064423 (2015), publisher: American Physical Society.

[12] R. Khymyn, I. Lisenkov, V. Tiberkevich, B. A. Ivanov, and A. Slavin, Scientific Reports **7**, 43705 (2017), number: 1 Publisher: Nature Publishing Group.

[13] R. Cheng, J. Xiao, Q. Niu, and A. Brataas, Physical Review Letters **113**, 057601 (2014), publisher: American Physical Society.

[14] A. Mougin, M. Cormier, J. P. Adam, P. J. Metaxas, and J. Ferré, Europhysics Letters (EPL) **78**, 57007 (2007), publisher: IOP Publishing.

[15] A. A. Thiele, Physical Review Letters **30**, 230 (1973), publisher: American Physical Society.

[16] R. Weber, D.-S. Han, I. Boventer, S. Jaiswal, R. Lebrun, G. Jakob, and M. Kläui, Journal of Physics D: Applied Physics **52**, 325001 (2019), publisher: IOP Publishing.

[17] V. Y. Irkhin, A. A. Katanin, and M. I. Katsnelson, Phys. Rev. B **60**, 1082 (1999).

[18] D. V. Spirin, Journal of Magnetism and Magnetic Materials **264**, 121 (2003).

[19] J. Kosterlitz, Journal of Physics C: Solid State Physics **7**, 1046 (1974).

[20] J. M. Kosterlitz and D. J. Thouless, in *Basic Notions Of Condensed Matter Physics* (CRC Press, 2018) pp. 493–515.

[21] N. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).

[22] B. L. Chittari, D. Lee, N. Banerjee, A. H. MacDonald, E. Hwang, and J. Jung, Phys. Rev. B **101**, 085415 (2020).

[23] P. Högl, T. Frank, K. Zollner, D. Kochan, M. Gmitra, and J. Fabian, Phys. Rev. Lett. **124**, 136403 (2020).

[24] K. Dolui, M. D. Petrović, K. Zollner, P. Plecháč, J. Fabian, and B. K. Nikolić, Nano Lett. **20**, 2288 (2020).

[25] M. Baglai, R. J. Sokolewicz, A. Pervishko, M. I. Katsnelson, O. Eriksson, D. Yudin, and M. Titov, Physical Review B **101**, 104403 (2020), publisher: American Physical Society.

[26] M. Fähnle and D. Steiauf, Physical Review B **73**, 184427 (2006).

[27] H. T. Simensen, A. Kamra, R. E. Troncoso, and A. Brataas, Physical Review B **101**, 020403 (2020), publisher: American Physical Society.

[28] A. Brataas, Y. Tserkovnyak, and G. E. W. Bauer, Physical Review Letters **101**, 037207 (2008), publisher: American Physical Society.

[29] H. Ebert, S. Mankovsky, D. Ködderitzsch, and P. J. Kelly, Physical Review Letters **107**, 066603 (2011), publisher: American Physical Society.

[30] Q. Liu, H. Y. Yuan, K. Xia, and Z. Yuan, Physical Review Materials **1**, 061401 (2017), publisher: American Physical Society.

[31] M. Dyakonov and V. Perel, Sov. Phys. Solid State, Ussr **13**, 3023 (1972).

[32] M. Dyakonov and V. Kachorovskij, Fizika i tehnika poluprovodnikov **20**, 178 (1986).

[33] R. Elliott, Phys. Rev. **96**, 266 (1954).

[34] Y. Yafet, in *Solid State Physics*, Vol. 14, edited by F. Seitz and D. Turnbull (Elsevier, 1963) pp. 1–98.

[35] M. Dyakonov, arXiv:cond-mat/0401369 (2004).

[36] M. I. Dyakonov, ed., *Spin Physics in Semiconductors*, 2nd ed., Springer Series in Solid-State Sciences (Springer International Publishing, 2017).

[37] N. Averkiev, L. Golub, and M. Willander, Semiconductors **36**, 91 (2002).

[38] A. Burkov, A. S. Núñez, and A. MacDonald, Phys. Rev. B **70**, 155308 (2004).

[39] A. A. Burkov and L. Balents, Physical Review B **69**, 245312 (2004), publisher: American Physical Society.

[40] N. A. Sinitsyn and Y. V. Pershin, Reports on Progress in Physics **79**, 106501 (2016), publisher: IOP Publishing.

[41] A. Kamra, R. E. Troncoso, W. Belzig, and A. Brataas, Phys. Rev. B **98**, 184402 (2018).

[42] C. Kittel, Phys. Rev. **82**, 565 (1951).

[43] F. Keffer and C. Kittel, Phys. Rev. **85**, 329 (1952).

[44] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).

[45] I. Ado, O. A. Tretiakov, and M. Titov, Phys. Rev. B **95**, 094401 (2017).

[46] J. Rammer, *Quantum Transport Theory* (CRC Press, New York, 2018).

[47] G. D. Mahan, *Many-particle physics* (Springer Science & Business Media, 2013).

[48] I. Ado, I. Dmitriev, P. Ostrovsky, and M. Titov, EPL **111**, 37004 (2015).

First-order effect of electron-electron interactions on the anomalous Hall conductivity of massive Dirac fermions

A. Daria Dumitriu-I.,[*] Darius-A. Deaconu, Alexander E. Kazantsev, and Alessandro Principi[†]
Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom

(Dated: April 22, 2024)

We investigate the first-order correction to the anomalous Hall conductivity of 2D massive Dirac fermions arising from electron-electron interactions. In a fully gapped system in the limit of zero temperature, we find that this correction vanishes, confirming the absence of perturbative corrections to the topological Hall conductivity. At finite temperature or chemical potential, we find that the total Hall response decays faster than in the noninteracting case, depending on the strength of electron-electron interactions. These features, which could potentially be observed experimentally, show the importance of two-body interactions for anomalous Hall transport.

I. INTRODUCTION

After the discovery of the quantum Hall effect, theoretical effort was directed towards understanding the robustness of the quantization of the Hall conductivity σ_H, which in the absence of interactions is related to a topological invariant [1, 2]. In a seminal paper [3], Coleman and Hill showed that two-particle interactions do not modify the value of σ_H at zero temperature, if the Fermi energy lies in the bulk band gap, which was later generalized to nonrelativistic interactions in Ref. [4]. This raises the question whether the Hall conductivity is robust to the effects of interactions in systems at finite temperature or chemical potential. To answer this question, we investigate the impact of electron-electron (e-e) interactions in the archetypal model of the anomalous Hall effect [5], a two-dimensional system of massive Dirac fermions.

Aside from its foundational significance, this issue also has practical implications since real-world materials exist at nonzero temperatures and are rarely completely free of doping. It has been demonstrated that certain many-body interactions, for example, that between electrons [6–8], or that between electrons and phonons [9], can substantially influence Hall responses. For instance, in the anomalous Hall and spin Hall effects, the presence of quenched disorder can yield outcomes remarkably at odds with noninteracting results [10, 11]. These include a faster decay of Hall responses with growing chemical potential, changes in sign, and, in some cases, even the complete elimination of these effects. However, while it may be possible to mitigate disorder in a material by refining the growth process, the omnipresent e-e interactions cannot be easily eliminated. Thus, it is fundamental to understand how they affect the Hall response to be able to predict and explain experimental results.

In this paper, we study the correction to the anomalous Hall effect of massive Dirac fermions to first order in the strength of e-e interactions. We start by introducing the model Hamiltonian and calculate the zeroth- and first-order response functions. We find that the first-order response is divergent. Divergences can be removed by renormalizing the bare parameters, which enter the model. In this way, we obtain a finite expression for the first-order correction to the Hall conductivity.

The first-order correction naturally depends on the form of the e-e interaction. Here, we initially address the case of a contact potential, whose relative simplicity allows us to do significant analytical progress. We relate such a potential to the Coulomb interaction between electrons in an overscreened regime. We show that interaction corrections can be large enough to counterbalance the noninteracting contribution to the Hall conductivity. In particular, we show that the anomalous Hall conductivity can decay faster with chemical potential than what is predicted in the noninteracting case. We then compute the same correction for an unscreened Coulomb potential, and show that it exhibits similar features. Thus, we confirm that they are robust irrespective of the precise form of the interaction.

II. DESCRIPTION OF THE MODEL

We consider massive Dirac fermions in two dimensions, whose dynamics is described by the many-body Hamiltonian (hereafter, we set $\hbar = 1$)

$$\hat{\mathcal{H}} = \sum_{\alpha\beta,\boldsymbol{k}} \hat{\psi}^{\dagger}_{\alpha,\boldsymbol{k}} (\boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma}_{\alpha\beta} - \mu\,\delta_{\alpha\beta})\hat{\psi}_{\beta,\boldsymbol{k}} + \frac{1}{2}\sum_{\boldsymbol{q}\neq 0}\sum_{\alpha\boldsymbol{k},\beta\boldsymbol{k}'} \hat{\psi}^{\dagger}_{\alpha,\boldsymbol{k}-\frac{\boldsymbol{q}}{2}} \hat{\psi}^{\dagger}_{\beta,\boldsymbol{k}'+\frac{\boldsymbol{q}}{2}} V_{\boldsymbol{q}} \hat{\psi}_{\beta,\boldsymbol{k}'-\frac{\boldsymbol{q}}{2}} \hat{\psi}_{\alpha,\boldsymbol{k}+\frac{\boldsymbol{q}}{2}}\,, \tag{1}$$

where $\psi_{\lambda,\boldsymbol{k}}$ ($\psi^{\dagger}_{\lambda,\boldsymbol{k}}$) destroys (creates) a particle with momentum $\boldsymbol{k}$ and pseudospin λ, μ is the chemical potential, $H(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma}$ is the single-particle Hamiltonian, $\boldsymbol{d}(\boldsymbol{k}) = (v_{\rm F}k_x, v_{\rm F}k_y, \Delta)$ and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is a vector of Pauli matrices. Here, Δ is half the band gap, while $v_{\rm F}$ is the Fermi velocity. The energy of these particles is $\pm\varepsilon_k$, where $\varepsilon_k = \sqrt{(v_{\rm F}k)^2 + \Delta^2}$ and the $\pm$ sign identifies the conduction and valence bands, respectively. Finally, $V_{\boldsymbol{q}}$ is the e-e interaction, which will be specified

---

[*] alexandra-daria.dumitriu-iovanescu@outlook.com
[†] alessandro.principi@manchester.ac.uk

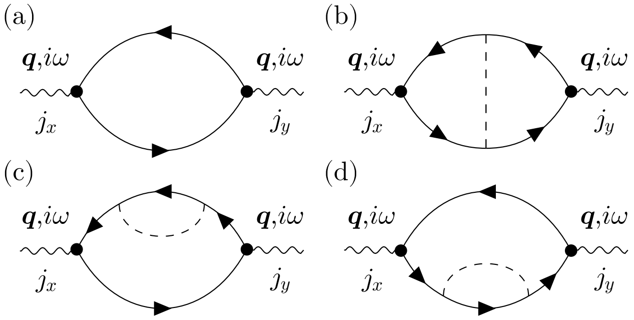


FIG. 1. (a) Noninteracting contribution to the current-current response function $\chi_{j_xj_y}(\boldsymbol{q},i\omega)$. [(b)–(d)] The first-order contributions to $\chi_{j_xj_y}(\boldsymbol{q},i\omega)$ due to vertex corrections (b), and self-energy insertions [(c)–(d)]. In these diagrams, solid lines are MGFs, dashed lines represent electron-electron interactions, while solid dots stand for current vertices.

later. The noninteracting Matsubara Green's function (MGF) corresponding to our Hamiltonian takes the form $G^{(0)}(\boldsymbol{k},i\omega_n) = \left[(i\omega_n+\mu)\sigma_0 - H(\boldsymbol{k})\right]^{-1}$, where $i\omega_n$ is a fermionic Matsubara frequency and σ_0 is the unit matrix.

Within the linear-response formalism [12], the Hall conductivity is defined in terms of the current-current response function $\chi_{j_xj_y}(\boldsymbol{q},\omega)$ as

$$\sigma_H = \lim_{\omega\to 0}\left[\frac{ie^2}{\omega}\chi_{j_xj_y}(\boldsymbol{q}=0,\omega)\right]. \tag{2}$$

Figure 1 summarizes the diagrammatic calculation of $\chi_{j_xj_y}(\boldsymbol{q},i\omega)$ to zeroth [panel (a)] and first order [panels (b)–(d)] in the e-e interaction. There, solid lines represent noninteracting MGFs, while dashed lines represent the e-e interaction. Finally, the solid dots are the current operators $j_{x/y} = v_{\rm F}\sigma_{x/y}$.

III. ZEROTH-ORDER CONTRIBUTION

At zeroth order in interaction we find the Hall conductivity [12](see Appendix A)

$$\sigma_H^{(0)}(\mu,T) = -\frac{e^2}{2h}\left(1-\mathcal{F}^+_{-2}(\bar\mu,\bar T)\right), \tag{3}$$

which recovers [13] the well-known zero-temperature result for the anomalous Hall conductivity at a finite doping (see, e.g., Ref. [14]), as well as the finite-temperature generalization of the Chern number, the so-called "Ulhmann number" [15]. In Eq. (3) we defined the integral

$$\mathcal{F}_n^{\pm} = \int_1^\infty dx\, x^n\left(\frac{1}{e^{(x-\bar\mu)/\bar T}+1} \pm \frac{1}{e^{(x+\bar\mu)/\bar T}+1}\right), \tag{4}$$

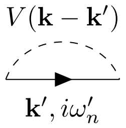
where $\bar\mu = \mu/\Delta$, $\bar T = k_{\rm B}T/\Delta$ are the chemical potential and the temperature, measured in units of (half) the gap.

IV. FIRST-ORDER CONTRIBUTION

Figures 1(b) and 1(c)–1(d) show the first-order diagrams, which yield the exchange and self-energy corrections to σ_H, respectively [16].

We start by evaluating the self-energy insertion, i.e.

$$-\Sigma(\boldsymbol{k}) =$$



$$= -\Sigma_0(\boldsymbol{k})\sigma_0 - \boldsymbol{\Sigma}(\boldsymbol{k})\cdot\boldsymbol{\sigma}, \tag{5}$$

where

$$\Sigma_0(\boldsymbol{k}) = -\frac{1}{2}\int\frac{d^2\boldsymbol{k}'}{(2\pi)^2}V(\boldsymbol{k}-\boldsymbol{k}')\Phi^+_{k'}, \tag{6}$$

$$\boldsymbol{\Sigma}(\boldsymbol{k}) = \frac{1}{2}\int\frac{d^2\boldsymbol{k}'}{(2\pi)^2}\frac{\boldsymbol{d}(\boldsymbol{k}')}{\varepsilon_{k'}}V(\boldsymbol{k}-\boldsymbol{k}')\Phi^-_{k'}, \tag{7}$$

are independent of the external frequency. We defined $\Phi_k^{\pm} = 1 - f_h(\varepsilon_k) \pm f_e(\varepsilon_k)$, where $f_e(\varepsilon) = [e^{(\varepsilon-\mu)/(k_{\rm B}T)}+1]^{-1}$ and $f_h(\varepsilon) = [e^{(\varepsilon+\mu)/(k_{\rm B}T)}+1]^{-1}$ are the Fermi–Dirac distribution functions for electrons and holes, respectively.

Adding the contributions from diagrams 1(b)– 1(d), and expanding to first order in frequency, we find the first-order correction to the current-current response function

$$\chi^{(1)}_{j_xj_y}(\boldsymbol{q}=0,i\omega) \to -\frac{\omega v_{\rm F}^2\Delta}{4}\int\frac{d^2\boldsymbol{k}}{(2\pi)^2}\Bigg\{-\left[\frac{2\Sigma_0(\boldsymbol{k})}{\varepsilon_k^3}\frac{\partial\Phi_k^+}{\partial\varepsilon_k}\right] + \left[\frac{2\Sigma_\perp(\boldsymbol{k})}{\varepsilon_k^2\Delta}\frac{\partial\Phi_k^-}{\partial\varepsilon_k}\right] + \int\frac{d^2\boldsymbol{k}'}{(2\pi)^2}V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_{\rm F}^2(\boldsymbol{k}\cdot\boldsymbol{k}'-k'^2)}{\varepsilon_k^3\varepsilon_{k'}^2}\frac{\partial\Phi^-_{k'}}{\partial\varepsilon_{k'}}\Phi^-_k\Bigg\}, \tag{8}$$

where we split $\boldsymbol{\Sigma}$ into $\boldsymbol{\Sigma}_\parallel = (\Sigma_x,\Sigma_y)$ and $\Sigma_\perp = \Sigma_z$. In Appendix B we show that the terms in the self-energy containing $\boldsymbol{\Sigma}_\parallel$ cancel similar terms from the exchange diagram. This is a consequence of the Ward identity [17].

Although the contribution from $\boldsymbol{\Sigma}_\parallel$ cancels between diagrams, Eq. (8) still depends on $\Sigma_0(\boldsymbol{k})$ and $\Sigma_\perp(\boldsymbol{k})$, both of which suffer from ultraviolet (UV) divergences and must be regularized by introducing a cutoff Λ (see below). For a contact potential $V(\boldsymbol{k}) = V_c$, where V_c is a constant, which does not depend on the momentum carried by the interaction, these divergences are quadratic $\propto \Lambda^2$ (for Σ_0) or linear $\propto \Lambda$ (for $\boldsymbol{\Sigma}_\parallel$ and $\Sigma_\perp$). For a Coulomb potential $V(\boldsymbol{k}) = 2\pi e^2/(\kappa k)$ for a medium with dielectric constant κ, the quadratic divergence becomes linear, while the linear ones become logarithmic.

A. Regularization of divergences

In order to deal with these divergences, we introduce the UV cutoff in momentum Λ by employing a type of Pauli–Villars [17][18] regularization scheme: The interaction is modified in such a way as to vanish fast enough at large momentum, so that the integrals become convergent. More explicitly, the interaction becomes $V(\boldsymbol{k}) \to V(\boldsymbol{k})\big[1 + (|\boldsymbol{k}|/\Lambda)^n\big]^{-1}$, where n is a natural number. The integrals in Eq. (8) then contain cutoff-dependent pieces signifying the presence of divergences, as well as finite parts. The latter become independent of the cutoff in the limit $\Lambda \to \infty$.

UV divergences are absorbed into the redefinition of the parameters, which enter into the Hamiltonian. Bare parameters (x^{bare}, where $x \equiv \mu, v_{\text{F}}, \Delta$) are defined in terms of physical ones (x^{ph}) as $x^{\text{bare}} = x^{\text{ph}} + \delta x^{\text{CT}}$. Here, δx^{CT} are the counterterms. From here on, we assume the Hamiltonian is defined in terms of renormalized parameters and hence omit the label “ph”.

Using the Dyson equation and the expression given in Eq. (5), we write the dressed MGF as a function of the bare parameters

$$\begin{aligned} G^{-1}(\boldsymbol{k}) = &\big(i\omega_n + \mu + \delta\mu^{\text{CT}} - \Sigma_0(\boldsymbol{k})\big)\,\sigma_0 \\ &- \big(\boldsymbol{d}(\boldsymbol{k}) + \delta\boldsymbol{d}^{\text{CT}}(\boldsymbol{k}) + \boldsymbol{\Sigma}(\boldsymbol{k})\big)\cdot\boldsymbol{\sigma}, \end{aligned} \tag{9}$$

where $\delta\boldsymbol{d}^{\text{CT}}(\boldsymbol{k}) = (\delta v_{\text{F}}^{\text{CT}} k_x, \delta v_{\text{F}}^{\text{CT}} k_y, \delta\Delta^{\text{CT}})$. Thus, the divergences introduced by $\Sigma_0(\boldsymbol{k})$, $\boldsymbol{\Sigma}_\parallel(\boldsymbol{k})$, and $\Sigma_\perp(\boldsymbol{k})$ are canceled by the renormalization of the chemical potential, Fermi velocity, and energy gap, respectively. The explicit expressions of the counterterms, which are consistent with previous results from Refs. [19, 20] are given below for the two separate cases of e-e interaction.

We remark on a notable feature of the Hall response. Because of the topological nature of the Hall response, i.e. the fact that $\chi^{(0)}_{j_x j_y}$ is independent of v_{F}, no additional first-order counterterm diagram is produced from $\chi^{(0)}_{j_x j_y}$ as a consequence of the renormalization $v_{\text{F}} \to v_{\text{F}} + \delta v_{\text{F}}^{\text{CT}}$. At the same time, there is no divergence to first order that would have to be canceled by such a counterterm diagram. This is precisely due to the cancellation of the $\boldsymbol{\Sigma}_\parallel$ terms between self-energy and vertex diagrams. This in turn implies that the renormalization of the chemical potential and band gap are sufficient to ensure the finiteness of first-order response.

For the case of a contact potential V_c, we find that the counterterms are [see Eqs. (C1)–(C3)]

$$\frac{\delta\mu^{\text{CT}}}{\mu} = -\frac{N_0 V_c}{4\,\text{sinc}(2\pi/n)}\frac{\Delta}{\mu}\left(\frac{v_{\text{F}}\Lambda}{\Delta}\right)^2, \tag{10}$$

$$\frac{\delta v_{\text{F}}^{\text{CT}}}{v_{\text{F}}} = -\frac{N_0 V_c}{4\,\text{sinc}(\pi/n)}\left(\frac{v_{\text{F}}\Lambda}{\Delta}\right), \tag{11}$$

$$\frac{\delta\Delta^{\text{CT}}}{\Delta} = -\frac{N_0 V_c}{2}\left(\frac{v_{\text{F}}\Lambda}{\Delta\,\text{sinc}(\pi/n)} - 1\right), \tag{12}$$

where $N_0 = \Delta/(2\pi v_{\text{F}}^2)$ is the density of states (at zero temperature) at the bottom of the conduction band.

Thus, after some lengthy algebra (see Appendix C), we find the contribution from the first-order diagrams

$$\begin{aligned} \sigma_H^{(1)} = &\frac{\bar{\sigma}_H}{2}\big[\mathcal{F}^+_{-2}(1 - \mathcal{F}^+_{-2} - 2\mathcal{F}^+_0) - 2\mathcal{F}^-_1\mathcal{F}^-_{-3} \\ &+ \mathcal{F}^+_0(f_h(\Delta) + f_e(\Delta) + 1) - \mathcal{F}^-_1(f_h(\Delta) - f_e(\Delta))\big]. \end{aligned} \tag{13}$$

This equation is one of the central results of our paper. We reintroduced the factors of $\hbar$ and defined $\bar{\sigma}_H = N_0 V_c\, e^2/2h$ as our effective coupling constant multiplied by one fourth of the conductance quantum (since we neglect spin and valley degeneracies).

We evaluate Eq. (13) numerically as a function of both chemical potential and temperature. We start by discussing the case when the factor $N_0 V_c$ is constant, i.e. it is independent of temperature and chemical potential. In this case, the behavior of the first-order correction to the Hall conductivity as a function of doping and temperature is shown in Fig. 2. The particle-hole symmetry is clearly reflected in the fact that the correction is an even function of the chemical potential. An important feature of the Hall response that can be inferred from Fig. 2(a) is the fact that the first-order correction vanishes at $\bar{T} = 0$ when the chemical potential is placed inside the gap, i.e. $|\bar{\mu}| \leq 1$. This is not by chance, but it is a consequence of the Coleman–Hill theorem [3], which states that there are no perturbative corrections to the topological part of the Hall conductivity. It is worth noting that the validity of this theorem for our model is not a trivial statement, since the original proof assumed Lorentz invariance while nonretarded e-e interactions break Lorentz invariance [21]. The theorem was extended to nonretarded interactions between electrons described by tight-binding models in Ref. [4].

Figure 2(b) shows the temperature dependence of the correction $\sigma_H^{(1)}$ for a set of values of the chemical potential. All curves have the same asymptotic behavior, which highlights the logarithmic divergence at high temperatures, i.e. $\sigma_H^{(1)} \to \ln 2(\ln\bar{T} + \mathcal{C})\bar{\sigma}_H$, where $\mathcal{C} \approx 0.98$.

Figure 2(c) shows $\sigma_H^{(1)}$ as a function of chemical potential for fixed values of temperature. The curve at $\bar{T} = 0$ clearly proves the validity of the Coleman–Hill theorem, i.e. the vanishing of the correction to the Hall conductivity within the band gap. In the Fermi liquid regime

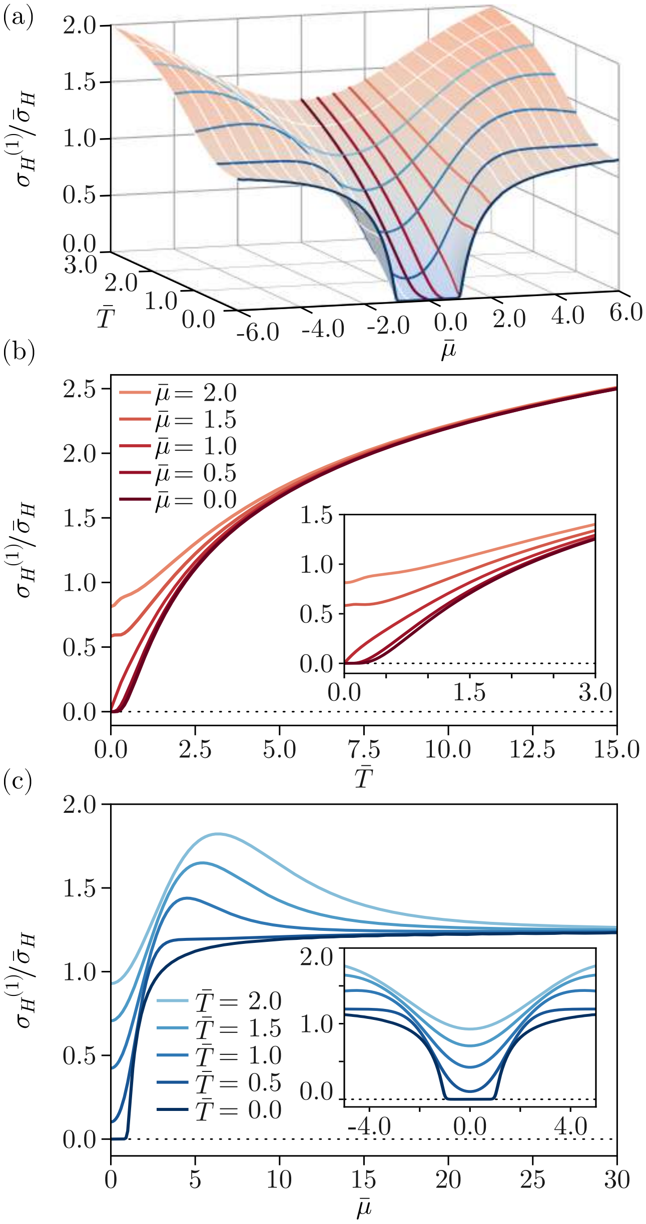


FIG. 2. (a) The first-order correction due to contact e-e interactions to the Hall conductivity $\sigma_H^{(1)}$ as a function of both chemical potential and temperature. (b) The correction $\sigma_H^{(1)}$ as a function of temperature, for various chemical potentials, along the cuts shown in the same colours in (a). A logarithmic divergence at high temperatures can be observed. (c) The correction $\sigma_H^{(1)}$ as a function of the chemical potential, along the cuts shown in the same colours in (a), for fixed values of the temperature. At large doping, curves converge to the value of 5/4. (Insets) Magnifications of the main plots in a similar range of values as the 3D plot.

(i.e. for small temperatures and $|\bar{\mu}| > 1$) the correction approaches the limit $\sigma_H^{(1)} \to (5\bar{\mu}^2 - 2|\bar{\mu}| - 3)/(4\bar{\mu}^2)\bar{\sigma}_H$.

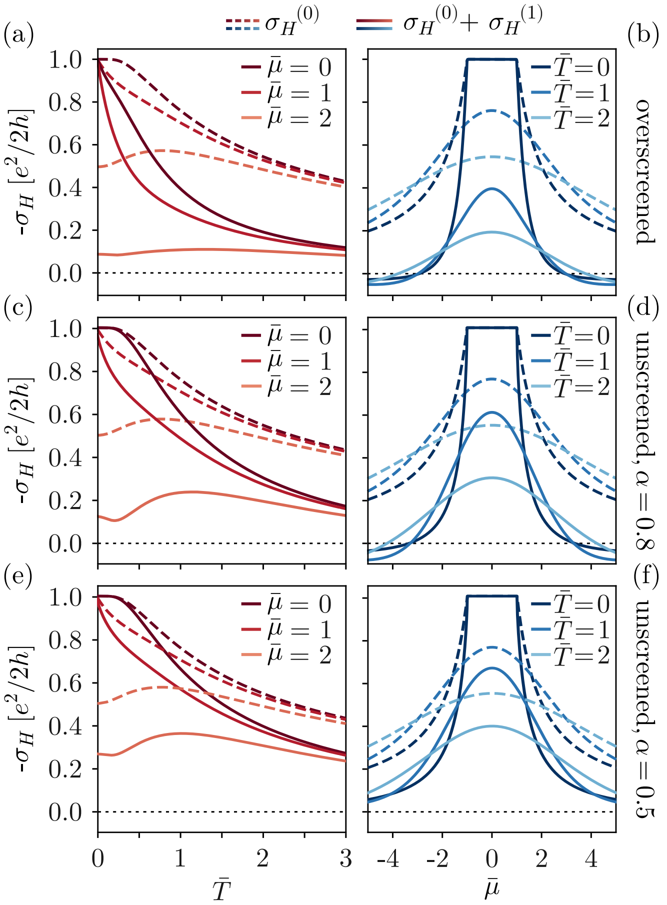


FIG. 3. A comparison between the noninteracting ($\sigma_H^{(0)}$, dashed lines) and full ($\sigma_H = \sigma_H^{(0)} + \sigma_H^{(1)}$, solid lines) Hall conductivities. Note that an overall minus sign was introduced. (a) Curves are shown as a function of temperature for fixed values of the chemical potential for a heavily screened Coulomb interaction. (b) Curves are shown as a function of chemical potential for fixed values of temperature for a heavily screened Coulomb interaction. [(c),(d)] Same as (a) and (b), respectively, for an unscreened Coulomb interaction with an effective fine structure constant $\alpha = 0.8$. [(e),(f)] Same as (a) and (b), respectively, for an unscreened Coulomb interaction with $\alpha = 0.5$.

B. Overscreened Coulomb interactions

When the density of states is large, the interaction is heavily screened. For a static Thomas–Fermi screening [22], the contact potential becomes $V_c \approx \nu^{-1}$, where $\nu(\mu, T) = \partial_\mu n(\mu, T)$ is the thermodynamic density of states. The effective coupling constant $N_0 V_c$ becomes

$$N_0 V_c = 4\bar{T} \left[\int_{|x|>1} dx |x| \cosh^{-2} \left(\frac{x - \bar{\mu}}{2\bar{T}} \right) \right]^{-1} . \quad (14)$$

The Hall conductivity to first order in interaction, $\sigma_H = \sigma_H^{(0)} + \sigma_H^{(1)}$, is shown with solid lines in Figs. 3(a) and 3(b). Comparing it with the noninteracting result (dashed lines), we see that e-e interactions in the high screening regime can be strong enough to offset the noninteracting Hall response: σ_H vanishes quicker than $\sigma_H^{(0)}$ with increasing chemical potential or temperature. This effect is similar to that of white-noise disorder, which strongly suppresses the anomalous Hall conductivity at high electron density [10].

Figure 3(b) also shows that the conductivity can become negative at large enough chemical potentials. This feature is most likely an artifact of the truncation to first-order in perturbation theory, since the choice $V_c \approx \nu^{-1}$ does not allow the potential to be arbitrarily small. We thus expect that summing higher order terms would remedy this feature. Note that even though Eq. (14) is formally valid only when the carrier density is large, we still continued the dark-blue-solid line ($\bar{T} = 0$) into the interval $|\bar{\mu}| \leq 1$ in Fig. 3(b) for completeness. This can be done because, due to the Coleman–Hill theorem, $\sigma_H^{(1)}$ is always zero in that region, irrespective of the exact form of V_c. A similar reasoning holds for the dark-red-solid line ($\bar{\mu} = 0$) in Fig. 3(a) at very low temperatures.

C. Unscreened Coulomb interaction

As before, UV divergences are isolated and absorbed into their corresponding counterterms [Eqs. (D1)–(D3)]

$$\frac{\delta\mu^{\mathrm{CT}}}{\mu} = -\frac{\alpha}{2}\frac{\Delta}{\mu}\left(\frac{v_{\mathrm{F}}\Lambda}{\Delta\,\mathrm{sinc}(\pi/n)}\right), \tag{15}$$

$$\frac{\delta v_{\mathrm{F}}^{\mathrm{CT}}}{v_{\mathrm{F}}} = -\frac{\alpha}{4}\left[\ln\left(\frac{v_{\mathrm{F}}\Lambda}{\Delta}\right) + \ln 2 + 1\right], \tag{16}$$

$$\frac{\delta\Delta^{\mathrm{CT}}}{\Delta} = -\frac{\alpha}{2}\left[\ln\left(\frac{v_{\mathrm{F}}\Lambda}{\Delta}\right) + \ln 2\right], \tag{17}$$

with $\alpha = e^2/(\kappa\hbar v_{\mathrm{F}})$ the effective fine structure constant.

The Hall conductivity as a function of temperature and chemical potential is shown in Figs. 3(c)–3(f). The exact expression for σ_H can be found in Appendix D.

In order to compare these results with those obtained for the high screening regime, we have chosen the value of the effective fine structure constant to be $\alpha = 0.8$ [panels (c)–(d)] and $\alpha = 0.5$ [panels (e)–(f)]. Similar to Figs. 3(a)–3(b), in both cases σ_H (solid lines) decays faster than $\sigma_H^{(0)}$ (dashed lines) as the temperature and the chemical potential increase. In Figs. 3(c) and 3(d) we can see that choosing $\alpha = 0.8$ gives a result of striking similarity with the high-screening regime. For $|\bar{\mu}| \gtrsim 3$, the Hall conductivity becomes negative, thus exhibiting the same artifact as Fig. 3(b).

Picking a smaller value of the coupling constant, e.g. $\alpha = 0.5$, moves the spurious zero of σ_H to larger values of the chemical potential, as can be seen in Fig. 3(f). This indicates that Hall conductivity becoming negative is most likely an artifact of truncation to first order.

V. DISCUSSION AND CONCLUSIONS

We calculated the first-order correction to the Hall conductivity due to e-e interactions. Firstly, we showed that the topological part of the Hall conductivity is robust against interactions, in agreement with the Coleman–Hill theorem. Secondly, we found that the effect of interactions can be strong enough to offset the noninteracting contribution. As a result, the total Hall conductivity goes to zero faster than the noninteracting one when either the temperature or the chemical potential increase. These features are present for both heavily screened and unscreened e-e interactions.

Note that this calculation pertains to the continuum model of a single Dirac cone. In a lattice model, massive Dirac fermions can be found as low-energy quasiparticle excitations of, e.g., a two-dimensional hexagonal lattice with broken sub-lattice symmetry. In this case, massive Dirac fermions of masses Δ_K and $\Delta_{K'}$ and opposite chiralities would appear in two distinct points of the Brillouin zone, K and K' respectively. Then the total Hall conductivity is the sum of the Hall conductivities of the two fermion species. This implies that, if the system is invariant under time-reversal symmetry, such that $\Delta_K = \Delta_{K'}$, then $\sigma_H = \sigma_H^{(0)} + \sigma_H^{(1)}$ exactly cancels between the two valleys. However, if time-reversal symmetry is broken but spatial inversion symmetry is preserved, such that $\Delta_K = -\Delta_{K'}$, then they add up to a finite value.

It is important to point out that the two valleys can be treated separately. This is because intervalley Coulomb backscattering is suppressed by a large intervalley momentum transfer $|\boldsymbol{K} - \boldsymbol{K}'|$ and therefore can be safely neglected. This remains true also for the contact interaction which should be interpreted as a heavily screened Coulomb potential. In fact, the interaction appears as contact only at length scales larger than the screening length. On the other hand, on smaller scales it recovers the original unscreened Coulomb interaction. We stress that this is consistent with our choice of regularization scheme, which forces the potential to decay when the momentum transfer exceeds the cut-off value Λ, which is taken to be much smaller than $|\boldsymbol{K} - \boldsymbol{K}'|$. Thus, intervalley backscattering can always be neglected.

Possible experimental platforms where these results can be tested include any gapped Dirac material with broken time-reversal symmetry. For instance, a gap can be opened in graphene via substrate effects [23], while creating an imbalance in the population of the opposite valleys by exposing it to circularly polarized light [24]. Moreover, a gap can also be induced by placing graphene in a chiral cavity [25] or by proximity coupling it to a ferromagnet [26–28]. We note that if the gap originates from the coupling to a ferromagnet, then the magnitude of the gap can decrease with temperature together with the magnetization in the substrate, which may make the Hall conductivity decay even faster (depending on the value of the Curie temperature) than simply due to ee interactions. As an alternative to graphene, the effect could also be observed in certain transition-metal dichalcogenides [29, 30], which have the advantage of a pre-existing gap due to spin-orbit coupling, and at the surfaces of 3D topological insulators [31, 32].

We stress that our results are applicable to other effects beyond anomalous Hall transport. We expect similar corrections to exist in, e.g., spin-Hall [33, 34] or orbital-Hall [35, 36] conductivities of sufficiently clean materials, although more work is needed to clarify the impact of e-e interactions on these effects.

ACKNOWLEDGMENTS

A.D.D.-I. and D.-A.D. acknowledge support from the Engineering and Physical Sciences Research Council, Grant No. EP/T517823/1. A.P. and A.E.K. acknowledge support from the Leverhulme Trust under the Grant Agreement No. RPG-2019-363. The authors also acknowledge support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (Project No. 873028 HYDROTRONICS).

Appendix A: Zeroth-order diagram

In the reverse order of the arrows in the diagram from Fig. 1(a), we use the finite-temperature Feynman rules given in Sec. 6.4 in Ref. [12][37] to calculate the noninteracting current-current response function

$$\begin{aligned}\chi^{(0)}_{j_x j_y}(\boldsymbol{q}, i\omega) =& \frac{1}{\beta}\sum_{i\omega_m}\int_{\boldsymbol{k}} \mathrm{Tr}[G^{(0)}(\boldsymbol{k}, i\omega_m) j_x \\ &\times G^{(0)}(\boldsymbol{k}+\boldsymbol{q}, i\omega_m + i\omega) j_y] \\ &= -\omega \frac{v_{\mathrm{F}}^2 \Delta}{2} \int_{\boldsymbol{k}} \frac{f_e(-\varepsilon_k) - f_e(\varepsilon_k)}{\varepsilon_k^3}. \end{aligned} \tag{A1}$$

In this equation we used the shorthand notation $\int_{\boldsymbol{k}} = \int d\boldsymbol{k}/(2\pi)^2$ for the integral, and $\sum_{i\omega_m}$ for the sum over Matsubara frequencies. The key advantage for evaluating this sum lies in the fact that the Fermi–Dirac distribution function $f_e(z) = [e^{\beta z} + 1]^{-1}$ has its poles at fermionic Matsubara frequencies $i\omega_m = i(2m+1)\pi/\beta$, where m is an integer, and the corresponding residues are $\mathrm{Res}[f_e(z)] = -1/\beta$. For a function of known simple poles, $g(z) = \Pi_j (z - z_j)^{-1}$,

$$\frac{1}{\beta}\sum_{i\omega_m} g(i\omega_m) = \sum_j \operatorname*{Res}_{z=z_j} \left[g(z)\right] f_e(z_j). \tag{A2}$$

After performing the Wick rotation to real frequencies $i\omega \to \omega + i0^+$ and using the Kubo formula given in Eq. (2), we obtain the Hall conductivity

$$\sigma_H^{(0)}(\mu, T) \;=\; -\frac{e^2\Delta}{4\pi}\int_{\Delta}^{\infty} d\varepsilon_k \frac{f_e(-\varepsilon_k) - f_e(\varepsilon_k)}{\varepsilon_k^2}. \tag{A3}$$

Using the Fermi–Dirac-like integral defined before in Eq. (4) and reintroducing the factors of $\hbar$, we finally get

$$\sigma_H^{(0)}(\bar{\mu}, \bar{T}) \;=\; -\frac{e^2}{2h}\left(1 - \mathcal{F}^{+}_{-2}(\bar{\mu}, \bar{T})\right) . \tag{A4}$$

At zero temperature, we can take the limit of $\mathcal{F}_n^{\pm}$,

$$\mathcal{F}_n^{+}(\bar{\mu}, 0) = \int_1^{|\bar{\mu}|} dx\, x^n = \frac{|\bar{\mu}|^{n+1} - 1}{n+1} , \tag{A5}$$

$$\mathcal{F}_n^{-}(\bar{\mu}, 0) = \mathrm{sign}(\bar{\mu}) \int_1^{|\bar{\mu}|} dx\, x^n = \mathrm{sign}(\bar{\mu}) \frac{|\bar{\mu}|^{n+1} - 1}{n+1} , \tag{A6}$$

which are true for $|\bar{\mu}| > 1$, but $\mathcal{F}_n^{\pm}(|\bar{\mu}| \leq 1, 0) = 0$ when the chemical potential lies inside the gap.

It can easily be seen that in this limit we get $\sigma_H^{(0)} = -e^2/2h$ when $|\bar{\mu}| \leq 1$ and $\sigma_H^{(0)}(\bar{\mu}, 0) = -(e^2/2h)(1/|\bar{\mu}|)$ otherwise. This recovers the well-known result for the anomalous Hall conductivity from Ref. [14].

Appendix B: First-order diagrams

Using the same Feynman rules for the other three diagrams, i.e. Figs. 1(b)–1(d), we calculate the response functions

$$\chi^{\mathrm{EX}}_{j_x j_y}(0, i\omega) = -\frac{1}{\beta^2}\sum_{i\omega_n, i\omega'_n}\int_{\boldsymbol{k}}\int_{\boldsymbol{k}'} V(\boldsymbol{k} - \boldsymbol{k}')\, \mathrm{Tr}[j_x G^{(0)}(\boldsymbol{k}, i\omega_n + i\omega) G^{(0)}(\boldsymbol{k}', i\omega'_n + i\omega) j_y G^{(0)}(\boldsymbol{k}', i\omega'_n) G^{(0)}(\boldsymbol{k}, i\omega_n)], \tag{B1}$$

$$\chi^{\mathrm{SE1}}_{j_x j_y}(0, i\omega) = \frac{1}{\beta}\sum_{i\omega_n}\int_{\boldsymbol{k}} \mathrm{Tr}[j_x G^{(0)}(\boldsymbol{k}, i\omega_n + i\omega)\Sigma(\boldsymbol{k}) G^{(0)}(\boldsymbol{k}, i\omega_n + i\omega) j_y G^{(0)}(\boldsymbol{k}, i\omega_n)], \tag{B2}$$

$$\chi^{\mathrm{SE2}}_{j_x j_y}(0, i\omega) = \frac{1}{\beta}\sum_{i\omega_n}\int_{\boldsymbol{k}} \mathrm{Tr}[j_x G^{(0)}(\boldsymbol{k}, i\omega_n + i\omega) j_y G^{(0)}(\boldsymbol{k}, i\omega_n)\Sigma(\boldsymbol{k}) G^{(0)}(\boldsymbol{k}, i\omega_n)]. \tag{B3}$$

The first one is the easiest to evaluate, and it can be simplified by using a parity transformation $\boldsymbol{k} \to -\boldsymbol{k}$ and $\boldsymbol{k}' \to -\boldsymbol{k}'$. The latter two, however, have second-order poles, which need to be treated carefully. These poles give rise to derivatives of the occupation functions, and change Eq. (A2) into

$$\frac{1}{\beta}\sum_{i\omega_m} g(i\omega_m) = \sum_{j\in\{j_1,j_2\}} \operatorname*{Res}_{z=z_j}\left[g(z)\right] f_e(z_j) + \sum_{j_2}\left[\frac{df_e(z)}{dz}\Big((z-z_{j_2})^2 g(z)\Big)\right]_{z=z_{j_2}}, \tag{B4}$$

where j_1 and j_2 are the simple and double poles, respectively.

After summing over the Matsubara frequencies and performing the traces over the Pauli matrices, we find that the first-order current-current response functions are

$$\chi^{\mathrm{EX}}_{j_xj_y}(0,i\omega) = -\omega\frac{v_{\mathrm{F}}^2\Delta}{4}\int_{\boldsymbol{k}}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\left[\frac{\varepsilon_{k'}^2+\Delta^2}{\varepsilon_{k'}^3\varepsilon_k^3}+\frac{v_{\mathrm{F}}^2\boldsymbol{k}'\cdot\boldsymbol{k}}{\varepsilon_{k'}^3\varepsilon_k^3}\right]\Phi^-_{k'}\Phi^-_k, \tag{B5}$$

$$\begin{aligned}
\chi^{\mathrm{SE}}_{j_xj_y} &= \omega\frac{v_{\mathrm{F}}\Delta}{2}\int_{\boldsymbol{k}}\left\{\Sigma_0(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\varepsilon_k^3}\frac{\partial\Phi^+_k}{\partial\varepsilon_k}\right]+\boldsymbol{\Sigma}_\parallel(\boldsymbol{k})\cdot\frac{v_{\mathrm{F}}^2\boldsymbol{k}}{\varepsilon_k}\left[\frac{3}{\varepsilon_k^4}\Phi^-_k-\frac{1}{\varepsilon_k^3}\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right]+\Sigma_\perp(\boldsymbol{k})\frac{v_{\mathrm{F}}}{\Delta}\left[\frac{3\Delta^2-\varepsilon_k^2}{\varepsilon_k^5}\Phi^-_k-\frac{\Delta^2}{\varepsilon_k^4}\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right]\right\}\\
&= \omega\frac{v_{\mathrm{F}}\Delta}{2}\int_{\boldsymbol{k}}\left\{\Sigma_0(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\varepsilon_k^3}\frac{\partial\Phi^+_k}{\partial\varepsilon_k}\right]-\boldsymbol{\Sigma}_\parallel(\boldsymbol{k})\cdot\frac{v_{\mathrm{F}}^2\boldsymbol{k}}{\varepsilon_k}\frac{\partial}{\partial\varepsilon_k}\left(\frac{\Phi^-_k}{\varepsilon_k^3}\right)+\Sigma_\perp(\boldsymbol{k})\frac{v_{\mathrm{F}}}{\Delta}\left[\left(\frac{2}{\varepsilon_k^3}-\frac{3}{\varepsilon_k^4}\frac{v_{\mathrm{F}}^2k^2}{\varepsilon_k}\right)\Phi^-_k+\left(\frac{v_{\mathrm{F}}^2k^2}{\varepsilon_k^4}-\frac{1}{\varepsilon_k^2}\right)\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right]\right\}\\
&= \omega\frac{v_{\mathrm{F}}\Delta}{2}\int_{\boldsymbol{k}}\left\{\Sigma_0(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\varepsilon_k^3}\frac{\partial\Phi^+_k}{\partial\varepsilon_k}\right]-\boldsymbol{\Sigma}_\parallel(\boldsymbol{k})\cdot\left[\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\left(\frac{\Phi^-_k}{\varepsilon_k^3}\right)\right]+\Sigma_\perp(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\Delta}\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\cdot\left(\boldsymbol{k}\frac{\Phi^-_k}{\varepsilon_k^3}\right)-\frac{v_{\mathrm{F}}}{\varepsilon_k^2\Delta}\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right]\right\},
\end{aligned}\tag{B6}$$

where Σ_0 is given in Eq. (6) and we split $\boldsymbol{\Sigma}$ from Eq. (7) into $\boldsymbol{\Sigma}_\parallel = (\Sigma_x,\Sigma_y)$ and $\Sigma_\perp = \Sigma_z$. When integrating the above by parts, $\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}$ gives both vanishing boundary terms and the 2D divergence (gradient) of $\boldsymbol{\Sigma}_\parallel(\Sigma_\perp)$,

$$\begin{aligned}
\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\cdot\boldsymbol{\Sigma}_\parallel(\boldsymbol{k}) &= \boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\cdot\left(\frac{1}{2}\int_{\boldsymbol{k}'}\frac{\boldsymbol{d}_\parallel(\boldsymbol{k}')}{\varepsilon_{k'}}V(\boldsymbol{k}-\boldsymbol{k}')\Phi^-_{k'}\right)\\
&= \frac{1}{2}\int_{\boldsymbol{k}'}\left[\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\cdot\left(\frac{\boldsymbol{d}_\parallel(\boldsymbol{k}')}{\varepsilon_{k'}}\right)\Phi^-_{k'}+\frac{\boldsymbol{d}_\parallel(\boldsymbol{k}')}{\varepsilon_{k'}}\cdot\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\left(\Phi^-_{k'}\right)\right]V(\boldsymbol{k}-\boldsymbol{k}')\\
&= \frac{v_{\mathrm{F}}}{2}\int_{\boldsymbol{k}'}\left[\left(\frac{\varepsilon_{k'}^2+\Delta^2}{\varepsilon_{k'}^3}\right)\Phi^-_{k'}+\frac{\boldsymbol{k}'}{\varepsilon_{k'}}\cdot\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\left(\Phi^-_{k'}\right)\right]V(\boldsymbol{k}-\boldsymbol{k}'),
\end{aligned}\tag{B7}$$

$$\begin{aligned}
\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\Sigma_\perp(\boldsymbol{k}) &= \boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\left(\frac{1}{2}\int_{\boldsymbol{k}'}\frac{\Delta}{\varepsilon_{k'}}V(\boldsymbol{k}-\boldsymbol{k}')\Phi^-_{k'}\right)\\
&= \frac{\Delta}{2}\int_{\boldsymbol{k}'}\left[\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\left(\frac{1}{\varepsilon_{k'}}\right)\Phi^-_{k'}+\frac{1}{\varepsilon_{k'}}\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\left(\Phi^-_{k'}\right)\right]V(\boldsymbol{k}-\boldsymbol{k}')\\
&= \frac{\Delta}{2}\int_{\boldsymbol{k}'}\left[\left(-\frac{v_{\mathrm{F}}^2\boldsymbol{k}'}{\varepsilon_{k'}^3}\right)\Phi^-_{k'}+\frac{1}{\varepsilon_{k'}}\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}\left(\Phi^-_{k'}\right)\right]V(\boldsymbol{k}-\boldsymbol{k}').
\end{aligned}\tag{B8}$$

Plugging these expressions back into Eq. (B6) gives an exact cancellation with the exchange response function

$$\begin{aligned}
\chi^{\mathrm{SE}}_{j_xj_y}(0,i\omega) =& -\omega\frac{v_{\mathrm{F}}\Delta}{2}\int_{\boldsymbol{k}}\left\{-\Sigma_0(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\varepsilon_k^3}\frac{\partial\Phi^+_k}{\partial\varepsilon_k}\right]-\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\cdot\boldsymbol{\Sigma}_\parallel(\boldsymbol{k})\left(\frac{\Phi^-_k}{\varepsilon_k^3}\right)+\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}}\Sigma_\perp(\boldsymbol{k})\cdot\left(\frac{v_{\mathrm{F}}\boldsymbol{k}}{\Delta}\frac{\Phi^-_k}{\varepsilon_k^3}\right)+\Sigma_\perp(\boldsymbol{k})\frac{v_{\mathrm{F}}}{\varepsilon_k^2\Delta}\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right\}\\
=& -\omega\frac{v_{\mathrm{F}}\Delta}{2}\int_{\boldsymbol{k}}\left\{-\Sigma_0(\boldsymbol{k})\left[\frac{v_{\mathrm{F}}}{\varepsilon_k^3}\frac{\partial\Phi^+_k}{\partial\varepsilon_k}\right]+\left[\Sigma_\perp(\boldsymbol{k})\frac{v_{\mathrm{F}}}{\varepsilon_k^2\Delta}\frac{\partial\Phi^-_k}{\partial\varepsilon_k}\right]-\frac{1}{2}\int_{\boldsymbol{k}'}V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_{\mathrm{F}}\boldsymbol{k}'}{\varepsilon_k^3\varepsilon_{k'}}\cdot\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}(\Phi^-_{k'})\Phi^-_k\right.\\
&\left.+\frac{1}{2}\int_{\boldsymbol{k}'}V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_{\mathrm{F}}\boldsymbol{k}}{\varepsilon_k^3\varepsilon_{k'}}\cdot\boldsymbol{\nabla}^{\mathrm{2D}}_{\boldsymbol{k}'}(\Phi^-_{k'})\Phi^-_k\right\}+\underbrace{\omega\frac{v_{\mathrm{F}}^2\Delta}{4}\int_{\boldsymbol{k}}\int_{\boldsymbol{k}'}V(\boldsymbol{k}-\boldsymbol{k}')\left[\frac{\varepsilon_{k'}^2+\Delta^2+v_{\mathrm{F}}^2\boldsymbol{k}\cdot\boldsymbol{k}'}{\varepsilon_k^3\varepsilon_{k'}^3}\right]\Phi^-_k\Phi^-_{k'}}_{=-\chi^{\mathrm{EX}}_{j_xj_y}(0,i\omega)},
\end{aligned}\tag{B9}$$

which is in accordance with the Ward identity. Summing the two leaves us with the result previously given in Eq. (8).

We isolate the divergences in Σ_0, $\boldsymbol{\Sigma}_\parallel$, and $\Sigma_\perp$ by separating the contribution of the holes (in the lower band) from that of the Fermi sea, *i.e.*, using the fact that $\Phi^\pm_k = f_e(-\varepsilon_k) \pm f_e(\varepsilon_k) = 1 - (f_h(\varepsilon_k) \mp f_e(\varepsilon_k))$. Since both $f_e(\varepsilon_k)$ and $f_h(\varepsilon_k)$ are exponentially suppressed at large momentum, only the terms containing the unity are UV divergent

$$\begin{aligned}\Sigma_0(\boldsymbol{k}) &= -\frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\Phi^+_{k'} = -\frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')(1-f_h(\varepsilon_{k'})+f_e(\varepsilon_{k'}))\\ &= -\frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}') - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')(f_e(\varepsilon_{k'})-f_h(\varepsilon_{k'})),\end{aligned}\tag{B10}$$

$$\begin{aligned}\boldsymbol{\Sigma}_\parallel(\boldsymbol{k}) &= \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\boldsymbol{d}_\parallel(\boldsymbol{k}')}{\varepsilon_{k'}}\Phi^-_{k'} = \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(1-f_h(\varepsilon_{k'})-f_e(\varepsilon_{k'}))\\ &= \frac{v_\mathrm{F}}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\boldsymbol{k}'}{\varepsilon_{k'}} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})),\end{aligned}\tag{B11}$$

$$\begin{aligned}\Sigma_\perp(\boldsymbol{k}) &= \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\Delta}{\varepsilon_{k'}}\Phi^-_{k'} = \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\Delta}{\varepsilon_{k'}}(1-f_h(\varepsilon_{k'})-f_e(\varepsilon_{k'}))\\ &= \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{1}{\varepsilon_{k'}} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\Delta}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})).\end{aligned}\tag{B12}$$

The divergences introduced by Σ_0, $\boldsymbol{\Sigma}_\parallel$, and $\Sigma_\perp$ become finite after the renormalization of the chemical potential, Fermi velocity and energy gap, respectively. Since these divergences depend on the form of the interaction, we treat the case of a constant contact potential $V(\boldsymbol{k}) = V_c$ and that of a Coulomb potential $V(\boldsymbol{k}) = 2\pi e^2/(\kappa k)$ separately.

Appendix C: Constant contact potential

We consider a regularised potential of the form $V(\boldsymbol{k}) \to V_c\left[1+(|\boldsymbol{k}|/\Lambda)^n\right]^{-1}$, where n is an integer. Introducing the counterterms in the previous equations leaves us with

$$\begin{aligned}\Sigma_0(\boldsymbol{k}) - \delta\mu^\mathrm{CT} &= -\frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}') - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')(f_e(\varepsilon_{k'})-f_h(\varepsilon_{k'})) - \delta\mu^\mathrm{CT}\\ &= -\frac{V_c\Lambda^2}{8\pi\,\mathrm{sinc}(2\pi/n)} - \delta\mu^\mathrm{CT} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')(f_e(\varepsilon_{k'})-f_h(\varepsilon_{k'}))\\ &= -\mu\underbrace{\left[\frac{N_0V_c}{4\,\mathrm{sinc}(2\pi/n)}\frac{\Delta}{\mu}\left(\frac{v_\mathrm{F}\Lambda}{\Delta}\right)^2 + \frac{\delta\mu^\mathrm{CT}}{\mu}\right]}_{=0} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')(f_e(\varepsilon_{k'})-f_h(\varepsilon_{k'})),\end{aligned}\tag{C1}$$

$$\begin{aligned}\boldsymbol{\Sigma}_\parallel(\boldsymbol{k}) + \delta v_\mathrm{F}^\mathrm{CT}\boldsymbol{k} &= \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}')\frac{v_\mathrm{F}(\boldsymbol{k}+\boldsymbol{k}')}{\varepsilon_{\boldsymbol{k}+\boldsymbol{k}'}} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})) + \delta v_\mathrm{F}^\mathrm{CT}\boldsymbol{k}\\ &= \frac{v_\mathrm{F}}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}')\left(\frac{\cancel{\boldsymbol{k}'}}{\cancel{\varepsilon_{k'}}} + \boldsymbol{k}\frac{\varepsilon_{k'}^2+\Delta^2}{2\varepsilon_{k'}^3} + \mathcal{O}(k^2)\right) - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})) + \delta v_\mathrm{F}^\mathrm{CT}\boldsymbol{k}\\ &= \boldsymbol{k}\left(\frac{V_c\Lambda}{8\pi\,\mathrm{sinc}(\pi/n)} + \delta v_\mathrm{F}^\mathrm{CT}\right) - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'}))\\ &= v_\mathrm{F}\boldsymbol{k}\underbrace{\left[\frac{N_0V_c}{4\,\mathrm{sinc}(\pi/n)}\frac{v_\mathrm{F}\Lambda}{\Delta} + \frac{\delta v_\mathrm{F}^\mathrm{CT}}{v_\mathrm{F}}\right]}_{=0} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{v_\mathrm{F}\boldsymbol{k}'}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})),\end{aligned}\tag{C2}$$

$$\begin{aligned}\Sigma_\perp(\boldsymbol{k}) + \delta\Delta^\mathrm{CT} &= \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}')\frac{1}{\varepsilon_{\boldsymbol{k}+\boldsymbol{k}'}} - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\Delta}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})) + \delta\Delta^\mathrm{CT}\\ &= \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}')\left(\frac{1}{\varepsilon_{k'}} - \boldsymbol{k}\cdot\frac{v_\mathrm{F}^2\cancel{\boldsymbol{k}'}}{\cancel{\varepsilon_{k'}^3}} + \mathcal{O}(k^2)\right) - \frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\Delta}{\varepsilon_{k'}}(f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})) + \delta\Delta^\mathrm{CT}\end{aligned}$$

$$
\begin{aligned}
&= \frac{V_c\Lambda}{4\pi\,\mathrm{sinc}(\pi/n)}\frac{\Delta}{v_\mathrm{F}} - \frac{V_c}{4\pi}\frac{\Delta^2}{v_\mathrm{F}^2} + \delta\Delta^\mathrm{CT} - \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})}{\varepsilon_{k'}} \\
&= \underbrace{\Delta\left[\frac{N_0V_c}{2\,\mathrm{sinc}(\pi/n)}\frac{v_\mathrm{F}\Lambda}{\Delta} - \frac{N_0V_c}{2} + \frac{\delta\Delta^\mathrm{CT}}{\Delta}\right]}_{=0} - \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{f_e(\varepsilon_{k'})+f_h(\varepsilon_{k'})}{\varepsilon_{k'}},
\end{aligned}
\tag{C3}
$$

where the crossed terms vanish by rotational symmetry and integrals of all $\mathcal{O}(k^2)$ terms vanish identically in the limit $\Lambda \to \infty$.

Note that in addition to the divergence itself, we also absorbed a finite constant term into $\delta\Delta^\mathrm{CT}$ in the last line. This is required to ensure that, after the renormalization, Δ is the "physical mass", which is defined as the pole of the dressed Green's function at zero momentum (at zero electron density). When $\mu = 0$ and $T = 0$, all the Fermi–Dirac distribution functions above vanish. Thus this requirement translates to $\Sigma_\perp(\boldsymbol{k}=0) + \delta\Delta^\mathrm{CT} = 0$. Choosing not to absorb this constant in the counterterm would make the renormalized mass related to the physical mass as $\Delta = \Delta^\mathrm{ph}(1 + N_0V_c/2)^{-1}$.

Plugging the (now finite) expressions for Σ_0 and $\Sigma_\perp$ from Eqs. (C1) and (C3) into Eq. (8), we can perform the substitution $i\omega \to \omega + i0^+$, and use the Kubo formula to recover the Hall conductivity given in Eq. (13).

When the temperature is zero and the chemical potential lies inside the gap $|\bar\mu| \le 1$, all functions $\mathcal{F}_n^\pm$ are identically zero. However, when $|\bar\mu| > 1$, we use the zero-temperature limits of $\mathcal{F}_n^\pm$ given in Eqs. (A5) and (A5)(A) and the fact that $f_h(\Delta) + f_e(\Delta) \xrightarrow{T\to 0} 1$ and $f_h(\Delta) - f_e(\Delta) \xrightarrow{T\to 0} -\mathrm{sign}(\bar\mu)$, to find the limit

$$
\sigma_H^{(1)} \xrightarrow{T\to 0} \frac{5|\bar\mu|^2 - 2|\bar\mu| - 3}{4|\bar\mu|^2}\bar\sigma_H,
\tag{C4}
$$

which at large doping converges to a finite value, i.e., $\sigma_H^{(1)} \to 5/4\,\bar\sigma_H$, as shown in Fig. 2(c).

For the high-temperature regime, we evaluate the following limits separately:

$$
\mathcal{F}_0^+ = \int_1^\infty dx\left(\frac{1}{e^{(x-\bar\mu)/\bar T}+1} + \frac{1}{e^{(x+\bar\mu)/\bar T}+1}\right) = \bar T\int_{1/\bar T}^\infty dx\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} + \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \xrightarrow{T\to\infty} 2\bar T\ln 2,
\tag{C5}
$$

$$
\begin{aligned}
\mathcal{F}_{-2}^+ &= \int_1^\infty dx\frac{1}{x^2}\left(\frac{1}{e^{(x-\bar\mu)/\bar T}+1} + \frac{1}{e^{(x+\bar\mu)/\bar T}+1}\right) = \frac{1}{\bar T}\int_{1/\bar T}^\infty dx\frac{d}{dx}\left(-\frac{1}{x}\right)\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} + \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \\
&= \left[\frac{1}{e^{(1-\bar\mu)/\bar T}+1} + \frac{1}{e^{(1+\bar\mu)/\bar T}+1}\right] + \frac{1}{\bar T}\int_{1/\bar T}^\infty dx\frac{d}{dx}(\ln x)\frac{d}{dx}\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} + \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \\
&= 1 - \frac{1}{2\bar T} + \mathcal{O}(\bar T^{-3}) + \frac{1}{\bar T}\left[\ln\bar T\left(-\frac{1}{2} + \mathcal{O}(\bar T^{-2})\right)\right] - \frac{1}{\bar T}\underbrace{\int_{1/\bar T}^\infty dx\ln x\frac{d^2}{dx^2}\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} + \frac{1}{e^{x+\bar\mu/\bar T}+1}\right)}_{\mathcal{C}=6\log(A)-0.5\gamma-0.5-0.66\ln(2)\approx 0.24} \\
&\xrightarrow{T\to\infty} 1 - \frac{1}{2\bar T}\ln\bar T - \frac{1}{\bar T}(\mathcal{C}+0.5),
\end{aligned}
\tag{C6}
$$

$$
\begin{aligned}
\mathcal{F}_1^- &= \int_1^\infty dx\,x\left(\frac{1}{e^{(x-\bar\mu)/\bar T}+1} - \frac{1}{e^{(x+\bar\mu)/\bar T}+1}\right) = \bar T^2\int_{1/\bar T}^\infty dx\,x\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} - \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \\
&= \bar T^2\frac{1}{\bar T}\ln\left(\frac{e^{(\bar\mu-1)/\bar T}+1}{e^{-(\bar\mu+1)/\bar T}+1}\right) - \frac{1}{\bar T}(\mathrm{Li}_2(-e^{-(1-\bar\mu)/\bar T}) - \mathrm{Li}_2(-e^{-(\bar\mu+1)/\bar T})) \xrightarrow{T\to\infty} 2\bar T(\bar\mu\ln 2),
\end{aligned}
\tag{C7}
$$

$$
\begin{aligned}
\mathcal{F}_{-3}^- &= \int_1^\infty dx\frac{1}{x^3}\left(\frac{1}{e^{(x-\bar\mu)/\bar T}+1} - \frac{1}{e^{(x+\bar\mu)/\bar T}+1}\right) = \frac{1}{\bar T^2}\int_{1/\bar T}^\infty dx\frac{d}{dx}\left(-\frac{1}{2x^2}\right)\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} - \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \\
&= -\frac{1}{2\bar T^2}\left[-\frac{\bar\mu\bar T}{2} + \mathcal{O}(\bar T^{-1})\right] + \frac{1}{2\bar T^2}\int_{1/\bar T}^\infty dx\frac{d}{dx}\left(-\frac{1}{x}\right)\frac{d}{dx}\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} - \frac{1}{e^{x+\bar\mu/\bar T}+1}\right) \\
&= \frac{\bar\mu}{4\bar T} - \frac{1}{2\bar T}\left[\frac{\bar\mu}{4\bar T} + \mathcal{O}(\bar T^{-3})\right] + \underbrace{\frac{1}{2\bar T^2}\int_{1/\bar T}^\infty dx\frac{1}{x}\frac{d^2}{dx^2}\left(\frac{1}{e^{x-\bar\mu/\bar T}+1} - \frac{1}{e^{x+\bar\mu/\bar T}+1}\right)}_{\mathcal{O}(\bar T^{-3})} \xrightarrow{T\to\infty} \frac{\bar\mu}{4\bar T},
\end{aligned}
\tag{C8}
$$

where we used square brackets for the boundary terms, γ for Euler's constant, A for Glaisher's constant [for the constant $\mathcal{C}$ in Eq. (C6)], and neglected terms of order $\mathcal{O}(\bar{T}^{-2})$. The latter is because the highest temperature divergence in these terms was linear and this can only be multiplied by terms of order $\mathcal{O}(\bar{T}^{0})$ and $\mathcal{O}(\bar{T}^{-1})$ to give a nonvanishing contribution to the expansion of $\sigma_H^{(1)}$ as $\bar{T} \to \infty$. Also note that in this limit $f_e(\Delta) + f_h(\Delta) \to 1 - 1/(2\bar{T})$ and $f_e(\Delta) - f_h(\Delta) \to \bar{\mu}/(2\bar{T})$. Imposing these limits to Eq. (13), we finally arrive at

$$\sigma_H^{(1)} \xrightarrow{T\to\infty} \ln 2(\ln \bar{T} + 2\mathcal{C} + 0.5)\bar{\sigma}_H \approx \ln 2(\ln \bar{T} + 0.98)\bar{\sigma}_H . \tag{C9}$$

Appendix D: Coulomb potential

In the case of a Coulomb potential, $\mathcal{O}(k^2)$ terms no longer vanish in the limit $\Lambda \to \infty$, but instead they give a finite contribution. Since the terms containing sums and differences of Fermi–Dirac distributions remain the same as in Eqs. (B10)–(B12), we only look at the contribution from the Fermi sea when introducing the counterterms

$$\Sigma_0^{\rm FS}(\boldsymbol{k}) - \delta\mu^{\rm CT} = -\frac{1}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}') - \delta\mu^{\rm CT} = -\frac{1}{2}\int_{\boldsymbol{k}'} \frac{2\pi e^2}{\kappa k'(1+|\boldsymbol{k}'|^n/\Lambda^n)} - \delta\mu^{\rm CT} = -\mu \underbrace{\left[\frac{\alpha}{2}\frac{\Delta}{\mu}\frac{v_{\rm F}\Lambda}{\Delta\,\mathrm{sinc}(\pi/n)} + \frac{\delta\mu^{\rm CT}}{\mu}\right]}_{=0} \tag{D1}$$

$$\begin{aligned}
\boldsymbol{\Sigma}_{\parallel}^{\rm FS}(\boldsymbol{k}) + \boldsymbol{k}\delta v_{\rm F}^{\rm CT} &= \frac{v_{\rm F}}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{\boldsymbol{k}'}{\varepsilon_{k'}} + \boldsymbol{k}\delta v_{\rm F}^{\rm CT} = \frac{v_{\rm F}}{2}\int_{\boldsymbol{k}'} \frac{2\pi e^2}{\kappa|\boldsymbol{k}-\boldsymbol{k}'|\big(1+|\boldsymbol{k}-\boldsymbol{k}'|^n/\Lambda^n\big)}\frac{\boldsymbol{k}'}{\varepsilon_{k'}} + \boldsymbol{k}\delta v_{\rm F}^{\rm CT} \\
&= \frac{v_{\rm F}^2 e^2}{\kappa}\int_0^1 \frac{dx}{\sqrt{x(1-x)}}\int_{\boldsymbol{k}'} \frac{\boldsymbol{k}'\big(1+|\boldsymbol{k}-\boldsymbol{k}'|^n/\Lambda^n\big)^{-1}}{x v_{\rm F}^2(\boldsymbol{k}-\boldsymbol{k}')^2 + (1-x)(v_{\rm F}^2\boldsymbol{k}'^2+\Delta^2)} + \boldsymbol{k}\delta v_{\rm F}^{\rm CT} \\
&= \frac{v_{\rm F}^2 e^2}{\kappa}\int_0^1 \frac{dx}{\sqrt{x(1-x)}}\int_{\boldsymbol{k}'} \frac{(\boldsymbol{k}'+x\boldsymbol{k})\big(1+|\boldsymbol{k}'|^n/\Lambda^n\big)^{-1}}{v_{\rm F}^2 k'^2 + (1-x)(x v_{\rm F}^2 k^2+\Delta^2)} + \boldsymbol{k}\frac{e^2}{8\kappa} + \boldsymbol{k}\delta v_{\rm F}^{\rm CT} \\
&= v_{\rm F}\boldsymbol{k}\underbrace{\left[\frac{\alpha}{4}\ln\left(\frac{v_{\rm F}\Lambda}{\Delta}\right) + \frac{\alpha}{4}(1+\ln 2) + \frac{\delta v_{\rm F}^{\rm CT}}{v_{\rm F}}\right]}_{=0} - v_{\rm F}\boldsymbol{k}\frac{\alpha}{4}\left[\frac{\varepsilon_k-\Delta}{2(\varepsilon_k+\Delta)} + \ln\left(\frac{\varepsilon_k+\Delta}{2\Delta}\right)\right],
\end{aligned} \tag{D2}$$

$$\begin{aligned}
\Sigma_{\perp}^{\rm FS}(\boldsymbol{k}) + \delta\Delta^{\rm CT} &= \frac{\Delta}{2}\int_{\boldsymbol{k}'} V(\boldsymbol{k}-\boldsymbol{k}')\frac{1}{\varepsilon_{k'}} + \delta\Delta^{\rm CT} = \frac{\Delta}{2}\int_{\boldsymbol{k}'} \frac{2\pi e^2}{\kappa|\boldsymbol{k}-\boldsymbol{k}'|\big(1+|\boldsymbol{k}-\boldsymbol{k}'|^n/\Lambda^n\big)}\frac{1}{\varepsilon_{k'}} + \delta\Delta^{\rm CT} \\
&= \frac{\Delta e^2}{\kappa}\int_0^1 \frac{dx}{\sqrt{x(1-x)}}\int_{\boldsymbol{k}'} \frac{v_{\rm F}\big(1+|\boldsymbol{k}-\boldsymbol{k}'|^n/\Lambda^n\big)^{-1}}{x v_{\rm F}^2(\boldsymbol{k}-\boldsymbol{k}')^2 + (1-x)(v_{\rm F}^2\boldsymbol{k}'^2+\Delta^2)} + \delta\Delta^{\rm CT} \\
&= \frac{\Delta v_{\rm F} e^2}{\kappa}\int_0^1 \frac{dx}{\sqrt{x(1-x)}}\int_{\boldsymbol{k}'} \frac{\big(1+|\boldsymbol{k}'|^n/\Lambda^n\big)^{-1}}{v_{\rm F}^2 k'^2 + (1-x)(x v_{\rm F}^2 k^2+\Delta^2)} + \delta\Delta^{\rm CT} \\
&= \Delta\underbrace{\left[\frac{\alpha}{2}\ln\left(\frac{v_{\rm F}\Lambda}{\Delta}\right) + \frac{\alpha}{2}\ln 2 + \frac{\delta\Delta^{\rm CT}}{\Delta}\right]}_{=0} - \frac{\alpha}{2}\Delta\ln\left(\frac{\varepsilon_k+\Delta}{2\Delta}\right),
\end{aligned} \tag{D3}$$

which we simplified by employing rotational symmetry. To get to the second lines of Eqs. (D2) and (D3) we used the Feynman parametrization formula

$$\frac{1}{A^\alpha B^\beta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 dx \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\big(xA+(1-x)B\big)^{\alpha+\beta}}, \tag{D4}$$

with terms $A = v_{\rm F}^2(\boldsymbol{k}-\boldsymbol{k}')^2$ and $B = \varepsilon_k^2 = v_{\rm F}^2 k'^2 + \Delta^2$ and exponents $\alpha = \beta = 1/2$.

As discussed in the previous subsection, we choose to absorb a finite constant term into $\delta\Delta^{\rm CT}$ to ensure that our definition of the energy gap corresponds to the "physical mass". However, unlike the case of the contact potential, there is an extra $\boldsymbol{k}$-dependent piece in Eq. (D3) introduced by the self-energy, but it vanishes when $\boldsymbol{k} = 0$.

The regularized potential used above should be read as $V(\boldsymbol{k}-\boldsymbol{k}') = \big(2\pi e^2/(\kappa|\boldsymbol{k}-\boldsymbol{k}'|)\big)\big(1+|\boldsymbol{k}-\boldsymbol{k}'|^n/\Lambda^n\big)^{-1}$, such that when changing the variable $\boldsymbol{k}' \to \boldsymbol{k}' + x\boldsymbol{k}$ in the third line of Eq. (D2), we are left with an extra term

$\boldsymbol{k}e^2/(8\kappa)$. Note that there is a finite constant absorbed in the counterterm $\delta v_{\mathrm{F}}^{\mathrm{CT}}$ similar to the one from the mass renormalization, but this does not influence our final result, as the latter does not depend on the way the Fermi velocity is renormalized. This can be better understood by noting that $\partial_{v_{\mathrm{F}}}\sigma_H^{(0)} = 0$ back in Eq. (A3), which means that there are no counterterm diagrams introduced by the renormalization of v_{F}.

Note that in addition to the contributions from the Fermi sea given in Eqs. (D1) and (D3), Σ_0 and $\Sigma_\perp$ also contain temperature and chemical potential dependent pieces. The latter are integrals over Fermi–Dirac distribution functions $f_e(\varepsilon_{k'}) \pm f_h(\varepsilon_{k'})$ and were given previously in Eqs. (B10) and (B12). Once we plug these together with the renormalized Fermi sea contributions given by Eqs. (D1) and (D3) back into Eq. (8), we notice that the resulting integrand nontrivially depends on the angle φ between $\boldsymbol{k}$ and $\boldsymbol{k}'$, unlike the case of the contact potential. This appears in the form $\int d\varphi V(\boldsymbol{k}-\boldsymbol{k}')$ and $\int d\varphi V(\boldsymbol{k}-\boldsymbol{k}')\ \boldsymbol{k}\cdot\boldsymbol{k}'$, which multiplied by $\kappa/(2\pi e^2)$ can be written as

$$\begin{aligned}
\int_0^{2\pi} d\varphi \frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|} &= \int_0^{2\pi} d\varphi \frac{1}{\sqrt{k^2+k'^2-2kk'\cos\varphi}} = \int_0^{2\pi} d\varphi \frac{1}{\sqrt{k^2+k'^2-2kk'\cos(\varphi+\pi)}} \\
&= 2\int_0^{\pi} d\varphi \frac{1}{\sqrt{k^2+k'^2+2kk'\cos\varphi}} = 2\int_0^{\pi} d\varphi \frac{1}{\sqrt{(k+k')^2-4kk'\sin^2(\varphi/2)}} \\
&= \frac{4}{k+k'}\int_0^{\pi/2} d\varphi \frac{1}{\sqrt{1-\frac{4kk'}{(k+k')^2}\sin^2\varphi}} = \frac{4}{k+k'} K\left(\frac{2\sqrt{kk'}}{k+k'}\right),
\end{aligned} \tag{D5}$$

$$\begin{aligned}
\int_0^{2\pi} d\varphi \frac{\boldsymbol{k}\cdot\boldsymbol{k}'}{|\boldsymbol{k}-\boldsymbol{k}'|} &= 2\int_0^{\pi} d\varphi \frac{kk'\cos\varphi}{\sqrt{k^2+k'^2-2kk'\cos\varphi}} = -2\int_0^{\pi} d\varphi \frac{kk'\left(1-2\sin^2(\varphi/2)\right)}{\sqrt{k^2+k'^2+2kk'\left(1-2\sin^2(\varphi/2)\right)}} \\
&= -4kk'\int_0^{\pi/2} d\varphi \frac{1}{\sqrt{(k+k')^2-4kk'\sin^2\varphi}} + 8kk'\int_0^{\pi/2} d\varphi \frac{\sin^2\varphi}{\sqrt{(k+k')^2-4kk'\sin^2\varphi}} \\
&= -\frac{4kk'}{k+k'} K\left(\frac{2\sqrt{kk'}}{k+k'}\right) + 2(k+k')\int_0^{\pi/2} d\varphi \frac{\frac{4kk'}{(k+k')^2}\sin^2\varphi}{\sqrt{1-\frac{4kk'}{(k+k')^2}\sin^2\varphi}} \\
&= -\frac{4kk'}{k+k'} K\left(\frac{2\sqrt{kk'}}{k+k'}\right) + 2(k+k')\left(-E\left(\frac{2\sqrt{kk'}}{k+k'}\right) + K\left(\frac{2\sqrt{kk'}}{k+k'}\right)\right) \\
&= \frac{2(k^2+k'^2)}{k+k'} K\left(\frac{2\sqrt{kk'}}{k+k'}\right) - 2(k+k') E\left(\frac{2\sqrt{kk'}}{k+k'}\right),
\end{aligned} \tag{D6}$$

where K and E are complete elliptic integrals of the first and second kind, respectively [38]. Plugging these integrals into Eq. (8), we find the Hall conductivity

$$\begin{aligned}
\sigma_H^{(1)} = -\frac{e^2}{2h}\alpha\frac{1}{4\pi}\int_0^\infty d\bar{k}\,\bar{k}\int_0^\infty d\bar{k}'\,\bar{k}'\Bigg\{ K\left(\frac{2\sqrt{\bar{k}\bar{k}'}}{\bar{k}+\bar{k}'}\right)\Bigg[\frac{4}{\bar{k}+\bar{k}'}\left(\frac{1}{\bar{\varepsilon}_k^3}\frac{\partial\Phi_k^+}{\partial\bar{\varepsilon}_k}(\Phi_{k'}^+-1) + \frac{1}{\bar{\varepsilon}_k^2\bar{\varepsilon}_{k'}}\frac{\partial\Phi_k^-}{\partial\bar{\varepsilon}_k}(\Phi_{k'}^- -1)\right) \\
- \frac{2(\bar{k}-\bar{k}')}{\bar{\varepsilon}_k^2\bar{\varepsilon}_{k'}^3}\frac{\partial\Phi_k^-}{\partial\bar{\varepsilon}_k}\Phi_{k'}^-\Bigg] - E\left(\frac{2\sqrt{\bar{k}\bar{k}'}}{\bar{k}+\bar{k}'}\right)\frac{2(\bar{k}+\bar{k}')}{\bar{\varepsilon}_k^2\bar{\varepsilon}_{k'}^3}\frac{\partial\Phi_k^-}{\partial\bar{\varepsilon}_k}\Phi_{k'}^-\Bigg\} + \frac{e^2}{2h}\frac{\alpha}{2}\int_0^\infty d\bar{k}\frac{\bar{k}}{\bar{\varepsilon}_k^2}\frac{\partial\Phi_k^-}{\partial\bar{\varepsilon}_k}\ln\left(\frac{\bar{\varepsilon}_k+1}{2}\right),
\end{aligned} \tag{D7}$$

where all the barred quantities are dimensionless $\bar{k} = v_{\mathrm{F}}k/\Delta$, $\bar{\varepsilon}_k = \varepsilon_k/\Delta$. This expression can be evaluated numerically as a function of both temperature and chemical potential, and its contribution added to $\sigma_H^{(0)}$ is shown in Figs. 3(c)–3(f) for different values of α.

[1] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. denNijs, Phys. Rev. Lett. **49**, 405 (1982).

[2] D. Xiao, M.-C. Chang, and Q. Niu, Reviews of modern physics **82**, 1959 (2010).

[3] S. Coleman and B. Hill, Physics Letters B **159**, 184 (1985).

[4] C. Zhang and M. Zubkov, Journal of Physics A: Mathematical and Theoretical **53**, 195002 (2020).

[5] F. D. M. Haldane, Physical review letters **61**, 2015 (1988).

[6] H.-H. Lai and H.-H. Hung, Physical Review B **89**, 165135 (2014).

[7] E. M. Hankiewicz and G. Vignale, Phys. Rev. B **73**, 115339 (2006).

[8] D.-H. Lee, Physical Review Letters **107**, 166806 (2011).

[9] J.-X. Zhang and W. Chen, Quantum theory of phonon induced anomalous hall effect in 2d massive dirac metals (2024), arXiv:2402.09049 [cond-mat.mes-hall].

[10] I. A. Ado, I. A. Dmitriev, P. M. Ostrovsky, and M. Titov, Phys. Rev. B **96**, 235148 (2017).

[11] I. Ado, I. Dmitriev, P. Ostrovsky, and M. Titov, Europhysics Letters **111**, 37004 (2015).

[12] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005).

[13] Since our calculation pertains to a single valley, it only recovers *half of* the result for the Haldane model.

[14] N. A. Sinitsyn, J. E. Hill, H. Min, J. Sinova, and A. H. MacDonald, Phys. Rev. Lett. **97**, 106804 (2006).

[15] L. Leonforte, D. Valenti, B. Spagnolo, A. A. Dubkov, and A. Carollo, Journal of Statistical Mechanics: Theory and Experiment **2019**, 094001 (2019).

[16] Note that there are three diagrams, which we omitted in the main text, namely the RPA and the Hartree self-energy diagrams. The first only contributes to the screening of the external field, while the latter two are canceled by the uniform positive background.

[17] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (CRC Press, Boca Raton, FL, 2005).

[18] We note that dimensional regularization could be used as well. In the case of a contact potential none of the divergences would appear, as they are all positive powers of the cutoff.

[19] J. González, F. Guinea, and M. Vozmediano, Nuclear Physics B **424**, 595 (1994).

[20] V. N. Kotov, V. M. Pereira, and B. Uchoa, Physical Review B **78**, 075433 (2008).

[21] This is because the Coulomb interaction is mediated by photons, which are much faster than the interacting electrons, thus making the interactions “instantaneous”.

[22] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Rev. Mod. Phys. **83**, 407 (2011).

[23] S. Y. Zhou, G.-H. Gweon, A. Fedorov, d. First, PN, W. De Heer, D.-H. Lee, F. Guinea, A. Castro Neto, and A. Lanzara, Nature materials **6**, 770 (2007).

[24] J. W. McIver, B. Schulte, F.-U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, Nature physics **16**, 38 (2020).

[25] X. Wang, E. Ronca, and M. A. Sentef, Physical Review B **99**, 235156 (2019).

[26] Z. Wang, C. Tang, R. Sachs, Y. Barlas, and J. Shi, Physical review letters **114**, 016603 (2015).

[27] P. Högl, T. Frank, K. Zollner, D. Kochan, M. Gmitra, and J. Fabian, Phys. Rev. Lett. **124**, 136403 (2020).

[28] K. Zollner, M. Gmitra, T. Frank, and J. Fabian, Phys. Rev. B **94**, 155441 (2016).

[29] T. Habe and M. Koshino, Phys. Rev. B **96**, 085411 (2017).

[30] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shi, Q. Niu, P. Tan, E. Wang, B. Liu, *et al.*, Nature communications **3**, 887 (2012).

[31] Y. L. Chen, J.-H. Chu, J. G. Analytis, Z. K. Liu, K. Igarashi, H.-H. Kuo, X. L. Qi, S. K. Mo, R. G. Moore, D. H. Lu, M. Hashimoto, T. Sasagawa, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z. X. Shen, Science **329**, 659 (2010).

[32] S. Bhattacharyya, G. Akhgar, M. Gebert, J. Karel, M. T. Edmonds, and M. S. Fuhrer, Advanced Materials **33**, 2007795 (2021).

[33] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, Rev. Mod. Phys. **87**, 1213 (2015).

[34] C. L. Kane and E. J. Mele, Physical review letters **95**, 226801 (2005).

[35] Y.-G. Choi, D. Jo, K.-H. Ko, D. Go, K.-H. Kim, H. G. Park, C. Kim, B.-C. Min, G.-M. Choi, and H.-W. Lee, Nature **619**, 52 (2023).

[36] S. Bhowal and G. Vignale, Physical Review B **103**, 195309 (2021).

[37] Note that there is a wrong minus sign in Rule (vii), which was later corrected in the errata of this book.

[38] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 2014).

Non-conservation of the valley density and its implications for the observation of the valley Hall effect

Alexander Kazantsev,[1,*] Amelia Mills,[1] Eoin O'Neill,[1] Hao Sun,[2] Giovanni Vignale,[2] and Alessandro Principi[1,†]

[1]*Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK*
[2]*The Institute for Functional Intelligent Materials (I-FIM),*
National University of Singapore, 4 Science Drive 2, Singapore 117544

We show that the conservation of the valley density in multi-valley and time-reversal-invariant insulators is broken in an unexpected way by the electric field that drives the valley Hall effect. This implies that fully-gapped insulators can support a valley Hall current in the bulk and yet show no valley density accumulation on the edges. Thus, the valley Hall effect cannot be observed in such systems. If the system is not fully gapped then valley density accumulation at the edges is possible and can result in a net generation of valley density. The accumulation has no contribution from undergap states and can be expressed as a Fermi surface average, for which we derive an explicit formula. We demonstrate the theory by calculating the valley density accumulations in an archetypical valley-Hall insulator: a gapped graphene nanoribbon. Surprisingly, we discover that a net valley density polarization is dynamically generated for some types of edge terminations.

Introduction—The valley Hall effect (VHE) in non-topological systems has recently stirred considerable controversy [1–9]. When the band structure features two valleys with a non-vanishing distribution of Berry curvature, electrons skew in the direction orthogonal to the applied electric field, even in the absence of magnetic field. However, since the system is not topological, electrons originating from one valley skew in the opposite direction of those from the other valley giving rise to a zero (charge) Hall current but to a finite *valley* Hall current $\boldsymbol{j}_v(\boldsymbol{r},t)$. This is defined as the difference between charge currents of electrons originating in opposite valleys. When this current hits the edge of the system, a valley density $n_v(\boldsymbol{r},t)$ (or, more physically, a density of orbital magnetic moment [10]), is expected to accumulate at its boundaries. This assumes that the valley density obeys a standard continuity equation [5, 6]. This seems a reasonable assumption: the two valleys are well-separated in momentum space, up to the point that they could ideally be taken as completely disconnected.

Some authors [1, 6] went further and claimed that even a fully-gapped non-topological insulator such as graphene aligned with hexagonal boron nitride (hBN) [3, 4] can exhibit nonlocal charge transport mediated by transverse undergap valley currents flowing in the bulk of the material. The authors of Ref. [6] argued that, at finite temperature, the valley-density accumulation could drive a "squeezed edge current" (parallel to the edges) in apparent agreement with experimental observations [2]. However, other authors [7–9] found from microscopic calculations that there is no valley density accumulation and no edge current in the simple graphene/hBN model. They proposed that the observed nonlocal resistances are caused by substrate-induced edge states crossing the Fermi level [7] or by substrate-induced valley-dependent scattering [9]. In the case of a fully gapped insulator this leaves us with the following puzzle: on one hand, the electric field drives a finite dissipationless valley Hall current in the bulk; on the other hand, time reversal symmetry implies that a valley density accumulation—a time-reversal-odd quantity—cannot appear in response to an electric field, unless there is dissipation, which is impossible if there are no states at the Fermi level. So where did the valley current go?

In this paper we solve the puzzle by observing that the valley density does not satisfy the conventional continuity equation when an electric field is present. This includes the field applied in order to drive the valley Hall current. The reason is that the electric field breaks the conservation of crystal momentum and therefore of valley number, which depends explicitly on crystal momentum. As a result, the bulk valley current is internally short-circuited as electrons flow from one valley to the other (and thus switch the sign of the Berry curvature) under the action of the very same electric field that drives the valley Hall current in the first instance. The process is schematically illustrated in Fig. 1.

Our results have profound implications for the observation of the VHE [11]. In a fully gapped time-reversal invariant insulator the undergap valley current is incapable of producing a valley density accumulation on the edge. This makes observing the VHE impossible in such systems, unless, e.g., the valley degeneracy is lifted (see, for example, Refs. [12–14]) or carriers are selectively injected into a single valley via circularly polarized light [15]. In these cases, however, an anomalous Hall effect "in disguise" is measured. This also means that the non-local resistance detection in Refs. [1, 16–19] must have been caused by partially occupied bands or edge states.

In metallic systems, which support a Fermi surface, our predictions for the valley density accumulation are quite different from those of the conventional theory which treats the entire bulk valley Hall current as the source of the accumulation. In particular, the value of the accumulation depends on the form of the electronic wave functions near the edge. The length over which it occurs

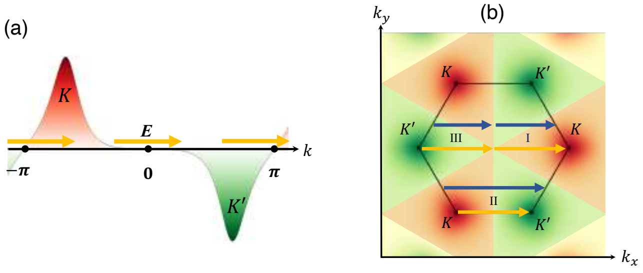


FIG. 1. Cyclic flow of the Bloch wave vector under the action of an electric field in the x direction. In both panels, the valley charge is -1 in the red regions and $+1$ in the green regions. Also shown are the Berry curvature hot spots with positive value near K and negative value near K'. Panel (a) shows the flow in a one-dimensional ribbon where the valley charge is non-conserved (changes sign) when k crosses the origin of momentum space. At the same time the Berry curvature, and hence the anomalous velocity also changes sign ensuring the existence of a steady valley Hall current. In a fully gapped insulator the flow does not alter the occupation of the states (i.e., a full band remains full) so there is no change in the valley density. Panel (b) shows examples of flows in a two-dimensional periodic system. In both cases the assumption of time-reversal symmetry rules out the possibility of a valley density response arising from the magneto-electric polarizability of the Bloch states.

is not related to the carrier diffusion length as in, e.g., Ref. [5], but reflects the much shorter localization length of edge states, as observed in some experiments [20], or the Fermi wavelength of bulk states. Perhaps the most important result of this study is that the valley density in the VHE is not simply transported from one edge to the other: it can be simultaneously generated on both edges by processes that involve the electric field in the bulk of the material.

Summary of main ideas—We consider a generic system in the shape of a strip of finite width which is indefinitely extended along the x axis. As we show below, the continuity equation satisfied by the valley density is

$$\partial_t n_v(y,t) + \partial_y j_v^y(y,t) = -e^2 E(t) \sum_k S(k) \partial_k f_k(y)\,, \quad (1)$$

where the electric field is in the x direction, which is parallel to the edge, and the valley current is in the y direction, perpendicular to the edge [21]. The system is assumed to be macroscopically homogeneous along x so that the valley density and current depend only on y. The electronic states (in the absence of the electric field) are taken to be of the form $\psi_{k,n}(x,y) = e^{ikx}u_{k,n}(x,y)/\sqrt{2\pi}$ where k is the x-component of the Bloch wave vector and n is the band index. The sum over k in Eq. (1) stands for $\int dk/(2\pi)$. The mixed electronic distribution $f_k(y) = a^{-1}\int_0^a dx \sum_n f_{k,n}\left|u_{k,n}(x,y)\right|^2$ is defined in terms of the electronic wave functions and the occupations of the corresponding states $f_{k,n}$, with the integral taken over one period a in the x direction. $S(k)$ is a "valley charge" function (odd under time reversal), which is a smooth periodic function of k in the Brillouin zone. We assume that the band structure features only two valleys, thus $S(k)$ assigns number $+1$ to states around one valley and -1 to states around the other valley. The valley density operator is $\hat{n}_v(\boldsymbol{r}) \equiv -(e/2)\sum_j \left\{S(\hat{k}_j), \delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_j)\right\}$ where $\hat{\boldsymbol{r}}_j$ and $\hat{k}_j$ are the position and Bloch momentum operator (along the edge) of the j-th electron, respectively, and $\{\hat{A},\hat{B}\} \equiv \left(\hat{A}\hat{B}+\hat{B}\hat{A}\right)$[22]. The valley current density is $\hat{\boldsymbol{j}}_v(\boldsymbol{r}) \equiv -(e/4)\sum_j \left\{S(\hat{k}_j), \left\{\hat{\boldsymbol{v}}_j, \delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_j)\right\}\right\}$, where $\hat{\boldsymbol{v}}_j$ is the velocity operator of the j-th electron. Because $\hat{k}$ is a constant of motion, $\hat{n}_v$ and $\hat{\boldsymbol{j}}_v(\boldsymbol{r})$ obey a conventional continuity equation in the absence of the electric field.

As we show below, in a fully gapped time-reversal invariant insulator, in which no edge or bulk state crosses the Fermi level, and at zero temperature, the right-hand side of Eq. (1) completely cancels the nearly-quantized contribution due to the second term on the left hand side. In this case, therefore, the valley density accumulation vanishes, even though there is a finite valley current in the bulk. In all other cases the cancellation is not exact. The correct equation for the density accumulation in the absence of relaxation processes is then $\partial_t n_v(y,t) = -Q_s(y)$, where the source term

$$Q_s(y) = \frac{e^2 E}{a}\int_0^a dx \sum_{k,n} \left(\partial_k f_{k,n}\right) S(k) \left|u_{k,n}(x,y)\right|^2, \quad (2)$$

is a Fermi surface property. Note that $Q_s(y)$ cannot be written, in general, as the divergence of a current. In fact, this is only possible if its integral over the whole strip vanishes, which implies that density accumulates at one edge and depletes at the other [23]. However, if the width of the strip is macroscopically large, the source term is localized on the edges. One can then define the "effective current" I_s, obtained by integrating Eq. (2) across a given edge, that feeds the valley number accumulation thereat. Since valley density is not conserved, the sum of the effective currents associated with the two edges does not have to be zero. I_s can be split as $I_s = I_s^e + I_s^b$, where $I_s^e = e^2 E \sum_{k,e}\left(\partial_k f_{k,e}\right) S(k)$ is the contribution of edge states. Here, the sum over e is that over edge states. The calculation of the contribution of bulk states, I_s^b, is complicated by the fact that the integral over y cannot be extended to infinity before performing the sum over n and k: the result would diverge. Nevertheless, a closed expression can be obtained in terms of the probability amplitude for Bloch waves to scatter off the edge [see Eq. (6) below]. Once I_s is known, the valley number accumulation can be estimated as $I_s\tau_{tr}$, where τ_{tr} is the intra- or inter-valley momentum relaxation time for the bulk or edge states' contribution, respectively.

Anomalous continuity equation—We consider a 2D crystal periodic in the x direction with period $a = 1$

and with the edges positioned at $y = 0$ and $y = -W$. A uniform electric field of magnitude E oscillating at frequency ω is applied along the x direction. For the sake of conciseness, hereafter we set $\hbar = 1$. Thus the conductance quantum e^2/h is equal to $e^2/(2\pi)$, where e is the electron charge. From the Kubo formula [24, 25], the y component of the valley current (averaged over x) is [26]

$$\begin{aligned} j_v^y(y,\omega) &= iEe^2 \sum_{k,n,n'} \int_0^y dy'(\varepsilon_{k,n}-\varepsilon_{k,n'}) \\ &\times\ S(k)\mathcal{L}_{k,nn'}(\omega)\mathcal{W}_{k,nn'}(y')\mathcal{A}_{k,n'n}, \end{aligned} \tag{3}$$

and the valley density (also averaged over x)

$$\begin{aligned} n_v(y,\omega) = &-\frac{iEe^2}{\omega+i0}\sum_{k,n}\partial_k f_{k,n}\, S(k)\, \mathcal{W}_{k,nn}(y) \\ &-Ee^2\sum_{k,n,n'} S(k)\mathcal{L}_{k,nn'}(\omega)\mathcal{W}_{k,nn'}(y)\mathcal{A}_{k,n'n}, \end{aligned} \tag{4}$$

where $\mathcal{L}_{k,nn'}(\omega) \equiv (f_{k,n} - f_{k,n'})/(\omega + \varepsilon_{k,n} - \varepsilon_{k,n'} + i0)$ is the usual Lindhard factor [25], $\mathcal{W}_{k,nn'}(y) \equiv \int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)$, and $\mathcal{A}_{k,n'n} = \int_0^1 dx \int_{-W}^0 dy\, u^\dagger_{k,n'}(x,y)i\partial_k u_{k,n}(x,y)$ is the Berry connection. The Fourier transform of Eq. (1) follows directly [26] from Eqs. (3) and (4):

$$-i\omega n_v(y,\omega) + \partial_y j_v^y(y,\omega) = -e^2E\sum_k S(k)\partial_k f_k(y). \tag{5}$$

The vanishing of valley density accumulation—Let us first assume that the system is a fully gapped time-reversal invariant insulator. The first term on the right hand side of Eq. (4) vanishes because $\partial_k f_{k,n} = 0$, since there are no bands that cross the Fermi level. In [26] we show that, due to time-reversal symmetry, the second line on the right hand side of Eq. (4) is proportional to ω, so the valley density accumulation vanishes in the limit of static electric field. This result implies that $\partial_y j_v^y(y)$ can be different from zero—as it must necessarily be, since the valley Hall current is finite in the bulk but vanishes at the edges—yet this finite divergence does not cause any density change at the edge or anywhere else. The resolution of this apparent paradox is provided by the anomalous term on the right hand side of Eq. (1) which exactly matches the divergence term on the left-hand side when the system is gapped. The undergap current does not produce a density accumulation.

The source of valley density—Let us now consider the case in which some energy levels cross the Fermi level. The first term on the right hand side of Eq. (4) causes the density to grow at a constant rate, leading to a breakdown of linear response theory unless a limiting momentum relaxation mechanism, such as intra- or inter-valley scattering, is taken into account. The Fermi surface term, obtained by multiplying Eq. (4) by $-i\omega$ and taking the $\omega \to 0$ limit, is the "source term" $Q_s(y)$ in Eq. (2). As discussed above, the integral of $Q_s(y)$ over y across a single edge can be interpreted as an effective current I_s that feeds the density accumulation thereat. I_s has contributions from both edge and bulk states that cross the Fermi level. The latter give (for the edge at $y = 0$)

$$I_s^b = -2e^2E\sum_{\lambda,k,p>0}\partial_k f^\lambda_{k,p}\,\mathfrak{Im}\left[\frac{[v^\lambda_{k,p}]^\dagger v^\lambda_{k,-p}R_\lambda(k,p)}{p+i0}\right], \tag{6}$$

where momentum integration is restricted to the valley with valley number +1, p is momentum in the y direction measured from the valley bottom, $v^\lambda_{k,p}$ are envelope amplitudes of propagating stationary states, labelled by index λ, $R_\lambda(k,p)$ is the reflection probability amplitude ($|R_\lambda(k,p)| = 1$) (see [26] for details).

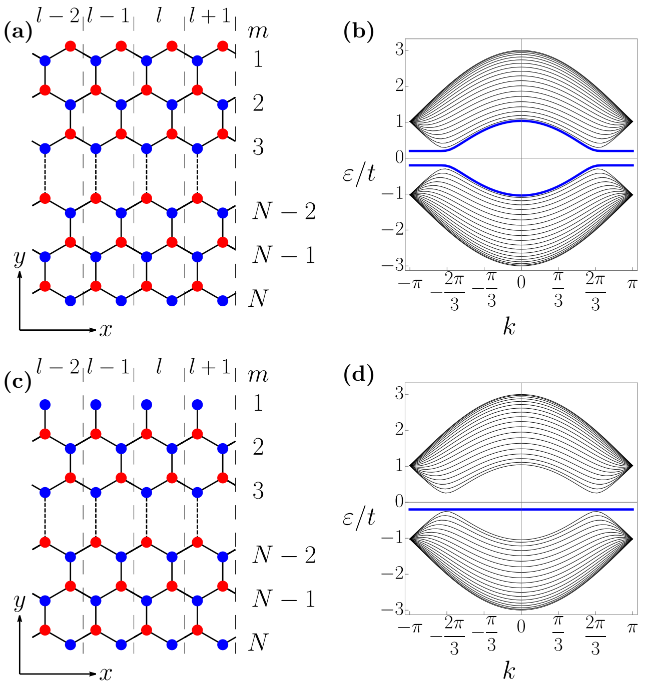


FIG. 2. Panel (a) and (c): Gapped graphene nanoribbon with zig-zag boundaries on both edges and a zig-zag and a bearded edge (at the top), respectively. Red (blue) discs signify atoms of the $\sigma = A$ ($\sigma = B$) sublattice, l labels the unit cells, m the rows in each unit cell. Panel (b) and (d): Band structures of nanoribbons of panels (a) and (c), respectively, for $N = 20$ and $\Delta = 0.2t$. Non-topological edge states are depicted by blue lines.

Example: "gapped graphene"—To demonstrate the main features of the general theory developed above, we calculate the valley Hall current and the valley density accumulation rate for a nanoribbon of "gapped graphene"—a model system that captures some aspects of monolayer graphene on a gap-inducing hBN substrate. Lattice sites are labelled with a unit cell number l and a composite index (m, σ), where $m = 1, \ldots, N$ denotes the row, while $\sigma = A, B$ distinguishes the sublattice within a given row. The y coordinate will be assumed to take integer values to mark the position within a row and half-integer values to mark the position in between the rows. The two sublattices, A and B, have different on-site potentials $\pm\Delta$. Electrons are assumed to hop only between nearest neighbors. We neglect spin-orbit interaction and therefore consider spinless electrons.

For the nanoribbon we consider two terminations: a) zig-zag boundaries on both edges [Fig. 2(a)] and b) a zig-zag and a bearded edge [Fig. 2(c)]. These lattice terminations ensure that the valley number is conserved by the unperturbed Hamiltonian. Each unit cell consists of N horizontal rows, with two atoms in each row as shown in Fig. 2(a), except the edge rows, where one atom may be missing as shown in Fig. 2(c).

The band structures for the two terminations, shown in Figs. 2(b) and (d), respectively, feature two bands separated by a gap equal to $2|\Delta|$ with minima at $k = \pm 2\pi/3$. These points define the two valleys in the one-dimensional Brillouin zone. When the lattice is terminated with zig-zag boundaries on both edges, two dispersionless bands of edge states [blue lines in Fig. 2(b)] connect the two valleys and become bulk states for $|k| < 2\pi/3$. The upper (lower) band of edge states resides on the upper (lower) edge in Fig. 2 (a).

Our main results are presented in Fig. 3. For a Fermi energy in the gap ($\varepsilon_{\mathrm{F}} = 0$) and at zero temperature, we find that $n_v(m, 0) = 0$ for either termination, consistent with the fact that there are no states at the Fermi level. At the same time $j^y_v(m+1/2, 0) = -Ee^2 \cdot \mathrm{sign}(\Delta)/(2\pi) + O(\Delta/t)$ for $1 \leq m \leq N-1$ as shown in panel (a), blue line: this is the undergap current associated with the nearly-quantized Hall conductance (the actual value -0.9 deviates from the ideal quantized value -1 due to the finite bandwidth of the model) [26].

When the system is doped with electrons ($\varepsilon_{\mathrm{F}} = 0.3t$), the current distributions differ dramatically for the two terminations, as shown by the red lines in Figs. 3 (a) and (b). In the case of the double zig-zag termination the current shows a linear variation across the ribbon (red line in (a)), changing sign about the center of the ribbon. This behavior is completely at odds with our intuition, which would lead us to expect an approximately constant current in the bulk, but not entirely unexpected, because there is no scattering and electrons propagate ballistically. Of greater physical interest, however, is the valley density accumulation rate which is shown in Fig. 3 (c). There is a significant cancellation between $-\partial_y j^y_v$ (green line) and the non-conservation term (black line) at the edges. The sum of the two results in a density accumulation rate which displays oscillations (red dots) on the scale of half the Fermi wavelength and two spikes of equal signs at the edges. These are the result of interference between the electronic waves incident on and reflected off the edge. The fact that the accumulation rate does not integrate to zero is the result of the anomaly on the right-hand side of Eq. (5): valley number is pumped from one valley into the other via a partially filled band of edge states connecting the two (upper blue line in Fig. 2 (b)). This opens the way to an intriguing possibility of generating a net valley density polarization by purely electrical means, as opposed to the standard optical methods. Notice, however, that the form of valley density accumulation rate cannot be predicted from the valley Hall current alone and depends on the boundary conditions.

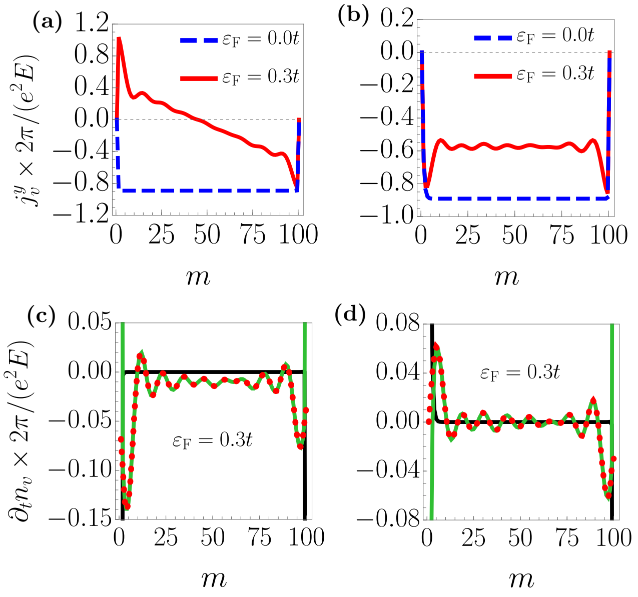


FIG. 3. Panel (a): Gapped graphene nanoribbon with zig-zag edges: the blue dashed line shows the valley Hall current as a function of position at Fermi energy $\varepsilon_{\mathrm{F}} = 0$ and the red solid line shows the same quantity at $\varepsilon_{\mathrm{F}} = 0.3t$ Panel (b): Same as in (a) for a nanoribbon with one edge zig-zag and the other bearded. Panel (c): The green line shows the valley density accumulation rate contributed by the valley Hall current in the doped nanoribbon with zig-zag edges. The black line shows the contribution of the non-conservation term, and the red dotted line is the sum of the two, *i.e.,* the total accumulation rate. Panel (d): Same as in (c) for the nanoribbon with one zig-zag and one bearded edge. In all plots $N = 100$, $\Delta = 0.1t$. In plots (c) and (d) $\varepsilon_{\mathrm{F}} = 0.3t$.

The zig-zag+bearded termination presents us with a more familiar scenario. Panel (b) of Fig. 3 shows that the valley Hall current is approximately constant (-0.55 in units of $e^2E/(2\pi)$ at $\varepsilon_{\mathrm{F}} = 0.3t$) in the bulk. At the same time the valley density accumulation rate, presented in panel (d) has spikes of opposite signs on the two edges. This suggests a more conventional picture of valley density being transported from one edge to the other. The reason for the overall valley number conservation is, in contrast to the previous example, absence of any partially filled bands that would connect the two valleys.

Conclusion—The modified continuity equation (1) allows us to explain how a non-vanishing undergap valley current can coexist with a vanishing valley density accumulation in a fully gapped non-topological time-reversal-invariant system with perfectly degenerate valleys. Any valley density accumulation requires the existence of states at the Fermi level and furthermore it is a dissipative process which requires a scattering mechanism to reach a steady state. We have provided closed expressions for calculating valley density accumulation rates on the edges of a two-dimensional material and we have applied them to the gapped graphene model: these formulas show that the connection between bulk currents and measurable edge accumulations is much more complex than previously suspected. This, in particular, leads us to surmise that any physical system in which evidence of the VHE has been found either by Kerr rotation microscopy [11] or by non-local resistance measurements [1, 16–19] cannot be a true insulator but must have partially populated bulk or edge states.

Acknowledgments—A.P. acknowledges support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS). A.P. and A.K. acknowledge support from the Leverhulme Trust under the grant RPG-2019-363. H.S. and G.V. were supported by the Ministry of Education, Singapore, under its Research Centre of Excellence award to the Institute for Functional Intelligent Materials (I-FIM, project No. EDUNC-33-18-279-V12).

---

[*] alexander.kazantsev@manchester.ac.uk
[†] alessandro.principi@manchester.ac.uk

[1] R. V. Gorbachev, J. C. W. Song, G. L. Yu, A. V. Kretinin, F. Withers, Y. Cao, A. Mishchenko, I. V. Grigorieva, K. S. Novoselov, L. S. Levitov, and A. K. Geim, Science **346**, 448 (2014).

[2] M. J. Zhu, A. V. Kretinin, M. D. Thompson, D. A. Bandurin, S. Hu, G. L. Yu, J. Birkbeck, A. Mishchenko, I. J. Vera-Marun, K. Watanabe, T. Taniguchi, M. Polini, J. R. Prance, K. S. Novoselov, A. K. Geim, and M. Ben Shalom, Nature Communications **8**, 14552 (2017).

[3] Y. D. Lensky, J. C. W. Song, P. Samutpraphoot, and L. S. Levitov, Phys. Rev. Lett. **114**, 256601 (2015).

[4] D. Xiao, W. Yao, and Q. Niu, Phys. Rev. Lett. **99**, 236809 (2007).

[5] M. Beconcini, F. Taddei, and M. Polini, Phys. Rev. B **94**, 121408 (2016).

[6] J. C. W. Song and G. Vignale, Phys. Rev. B **99**, 235405 (2019).

[7] J. M. Marmolejo-Tejada, J. H. García, M. D. Petrović, P.-H. Chang, X.-L. Sheng, A. Cresti, P. Plecháč, S. Roche, and B. K. Nikolić, Journal of Physics: Materials **1**, 015006 (2018).

[8] S. Roche, S. R. Power, B. K. Nikolić, J. H. García, and A.-P. Jauho, Journal of Physics: Materials **5**, 021001 (2022).

[9] T. Aktor, J. H. Garcia, S. Roche, A.-P. Jauho, and S. R. Power, Phys. Rev. B **103**, 115406 (2021).

[10] S. Bhowal and G. Vignale, Phys. Rev. B **103**, 195309 (2021).

[11] J. Lee, K. F. Mak, and J. Shan, Nature Nanotechnology **11**, 421 (2016).

[12] W. Du, R. Peng, Z. He, Y. Dai, B. Huang, and Y. Ma, npj 2D Materials and Applications **6**, 11 (2022).

[13] X. Ma, X. Shao, Y. Fan, J. Liu, X. Feng, L. Sun, and M. Zhao, J. Mater. Chem. C **8**, 14895 (2020).

[14] B. Zhou, Z. Li, J. Wang, X. Niu, and C. Luan, Nanoscale **11**, 13567 (2019).

[15] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, Science **344**, 1489 (2014).

[16] M. Sui, G. Chen, L. Ma, W.-Y. Shan, D. Tian, K. Watanabe, T. Taniguchi, X. Jin, W. Yao, D. Xiao, and Y. Zhang, Nature Physics **11**, 1027 (2015).

[17] K. Endo, K. Komatsu, T. Iwasaki, E. Watanabe, D. Tsuya, K. Watanabe, T. Taniguchi, Y. Noguchi, Y. Wakayama, Y. Morita, and S. Moriyama, Applied Physics Letters **114**, 243105 (2019), https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.5094456/13307549/243105_1_online.pdf.

[18] Y. Shimazaki, M. Yamamoto, I. V. Borzenets, K. Watanabe, T. Taniguchi, and S. Tarucha, Nature Physics **11**, 1032 (2015).

[19] E. Arrighi, V. H. Nguyen, M. D. Luca, G. Maffione, Y. Hong, L. Farrar, K. Watanabe, T. Taniguchi, D. Mailly, J. C. Charlier, and R. Ribeiro-Palau, Non-identical moiré twins in bilayer graphene (2023), arXiv:2205.01760 [cond-mat.mes-hall].

[20] I. Lyalin, S. Alikhah, M. Berritta, P. M. Oppeneer, and R. K. Kawakami, APS Bulletin (2023).

[21] Edges are chosen to be parallel to the separation of the valleys in momentum space in such a way that the valley index is a good quantum number in the absence of the electric field and impurities.

[22] The Bloch momentum (modular momentum) operator in a lattice is defined by the identity $e^{i\hat{k}na} = e^{i\hat{p}na}$, where $\hat{p}$ is the ordinary momentum operator and n an integer. The operator $S(\hat{k})$ is well defined because $S(k)$, a periodic function of k over the Brillouin zone, is expandable in a series of plane waves of the form e^{ikna}.

[23] A similar situation was discussed in Ref. [**?**] for the spin current in the spin Hall effect, with the crucial difference that the non-conservation of the spin density was caused, in that case, by the intrinsic spin-orbit torque, while in our case it is caused by the very same electric field that drives the valley Hall effect.

[24] R. Kubo, Journal of the Physical Society of Japan **12**, 570 (1957).

[25] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[26] See the supplemental online material for more details.

Supplemental Material

Everywhere below $\hbar = 1$.

I. THE SETUP

We consider a 2D system periodic in the x direction and of finite width W in the y direction. We will take the period in the x direction to be equal to one. We divide the system into unit cells, labelled by index l. We assume that the periodic direction is chosen such that the valley index remains a good quantum number, *i.e.,* edges are chosen to be parallel to the separation between the valleys in momentum space (one of the edges will be taken to be at $y = 0$ and the other at $y = -W$). The Bloch eigenstates are denoted as $\psi_{k,n}(x, y) = u_{k,n}(x, y)e^{ikx}/\sqrt{2\pi}$, where k is the wavevector in the x direction and n is the band index. To take account of the spin-orbit interaction, we will assume that both $\psi_{k,n}$ and $u_{k,n}$ are two-dimensional spinors. We apply the electric field of magnitude E in the direction parallel to the edges (*i.e.,* the x direction), *i.e.,* the Hamiltonian is perturbed by a potential term $eE\hat{x}$, where $-e$ is the electron charge and $\hat{x}$ the position operator.

For graphene nanoribbon with zig-zag and bearded edges (which preserve the valley number), the role of coordinates (x, y) is played by unit cell number l and combined index (m, σ), where $m = 1 \dots N$ denotes the two-atom horizontal row in each unit cell and σ the atom (A or B) within it (see Fig. 1 (a), (c) in the main text). Note that a row may miss an atom of either sublattice, as in Fig. 1 (c), where it misses an A atom in row 1 of each unit cell. Everywhere below when displaying the results for densities and currents in the nanoribbon, the position on the y axis will only be resolved down to the row number m. Therefore, the y coordinate will be replaced by a discrete index that will take integer values marking the position in a certain row and half-integer values marking the position (half-way) between the rows.

II. POSITION OPERATOR IN BLOCH STATE BASIS

The derivation of the position operator representation in the Bloch state basis here follows closely Ref. [1]. Consider a wavepacket $\chi(x, y)$ that is a superposition of Bloch eigenstates with coefficients $\chi_{k,n}$

$$\chi(x,y) = \sum_n \int dk \chi_{k,n}\psi_{k,n}(x,y). \tag{1}$$

The application of the position operator $\hat{x}$ to $\chi(x, y)$ gives

$$\begin{aligned}\hat{x}\chi(x,y) &= \sum_n \int dk\, \chi_{k,n}\, x\psi_{k,n}(x,y) \\ &= \sum_n \int dk \chi_{k,n}\Big(-i\partial_k\psi_{k,n}(x,y) \\ &\qquad + ie^{ikx}\partial_k u_{k,n}(x,y)/\sqrt{2\pi}\Big).\end{aligned} \tag{2}$$

We integrate by parts in the first term, while discarding the boundary contribution. Furthermore, we expand $i\partial_k u_{k,n}$ in the series of functions $u_{k,n}$ so that

$$i\partial_k u_{k,n}(x,y) = \sum_{n'} u_{k,n'}(x,y)\mathcal{A}_{k,n'n}, \tag{3}$$

with $\mathcal{A}_{k,n'n} = \int_0^1 dx \int_{-W}^0 dy\, u^\dagger_{k,n'}(x,y) i\partial_k u_{k,n}(x,y)$. As a result, we obtain (lightening the notation by suppressing x and y dependence)

$$\begin{aligned}\hat{x}\chi &= \sum_n \int dk \Big[(i\partial_k\chi_{k,n})\psi_{k,n} + \chi_{k,n}\mathcal{A}_{k,n'n}\psi_{k,n'}\Big] \\ &= \sum_{nn'} \int dk dk'\ \psi_{k',n'}\ x_{k'n';kn}\ \chi_{k,n},\end{aligned} \tag{4}$$

where

$$x_{k'n';kn} = -i\partial_k\delta(k-k')\delta_{nn'} + \delta(k-k')\mathcal{A}_{k,n'n}. \tag{5}$$

III. VALLEY NUMBER AND VALLEY HALL CURRENT OPERATORS

Here everything refers to just one electron, the generalization to many electrons is trivial. The operator that gives valley density at position $\boldsymbol{x} = (x, y)$ has the form

$$\hat{n}_v(x,y) = -\frac{e}{2}\{\hat{n}(x,y), S(\hat{k})\}, \tag{6}$$

where $\hat{n}(x, y) = |x, y\rangle\langle x, y|$ is the particle density operator at position (x, y) (with any spin polarization) and the curly brackets stand for an anticommutator.

Note that the particle density operator satisfies the continuity equation

$$\partial_t\hat{n}(x,y,t) + \boldsymbol{\nabla}\hat{\boldsymbol{j}}(x,y,t) = 0, \tag{7}$$

with a particle number current

$$\hat{\boldsymbol{j}}(x,y) = \frac{1}{2}\{\hat{\boldsymbol{v}}, n(x,y)\}, \tag{8}$$

where $\hat{\boldsymbol{v}}$ is the velocity operator. Operator $S(\hat{k})$, on the other hand, is a constant of motion. Therefore, taking the time derivative of Eq. (6) at time t, we obtain

$$\partial_t\hat{n}_v(x,y,t) + \boldsymbol{\nabla}\hat{\boldsymbol{j}}_v(x,y,t) = 0, \tag{9}$$

where

$$\begin{aligned}\hat{\boldsymbol{j}}_v(x,y) &= -\frac{e}{2}\{\hat{\boldsymbol{j}}(x,y), S(\hat{k})\} \\ &= -\frac{e}{4}\big\{\{\hat{\boldsymbol{v}}, \hat{n}(x,y)\}, S(\hat{k})\big\}. \end{aligned} \tag{10}$$

Note also that $\hat{n}(x,y) = \delta(\boldsymbol{x}-\hat{\boldsymbol{r}})$, where $\hat{\boldsymbol{r}}$ is the electron's position operator.

IV. KUBO FORMULA FOR VALLEY NUMBER AND VALLEY HALL CURRENT

The Kubo formula [2, 3] allows one to calculate the change in the ensemble average of an observable caused by a perturbation to first order in its strength. If the Hamiltonian is perturbed by an operator $\hat{B}\exp(-i\omega t)$ then, long after the perturbation was turned on, the ensemble average of the observable $\hat{A}$ will be oscillating at the frequency ω with the amplitude

$$A(\omega) = \sum_{\alpha\beta} \frac{f_\alpha - f_\beta}{\omega + \varepsilon_\alpha - \varepsilon_\beta + i0} A_{\alpha\beta} B_{\beta\alpha}. \tag{11}$$

Here, α and β label eigenstates of the unperturbed Hamiltonian, whose eigenvalues are ε_α and ε_β, respectively. Furthermore, $A_{\alpha\beta}$ and $B_{\beta\alpha}$ are matrix elements of operators $\hat{A}$ and $\hat{B}$ between eigenstates α and β, while $f_{\alpha,\beta} = (\exp[(\varepsilon_{\alpha,\beta} - \varepsilon_{\mathrm{F}})/(k_B T)] + 1)^{-1}$ are the occupation numbers of such states.

For the nanoribbon, the unperturbed stationary states are labelled by the value of (quasi)momentum k and by the band index n. In the case of a harmonic electric field applied parallel to the edge, $\hat{B} = eE\hat{x}$. Using Eq. (5), the matrix elements of the perturbation are then given by the equation

$$\begin{aligned} B_{k'n';kn} &= eEx_{k'n';kn} \\ &= eE\big(-i\partial_k \delta(k-k')\delta_{nn'} + \delta(k-k')\mathcal{A}_{k,n'n}\big). \end{aligned} \tag{12}$$

The role of observable $\hat{A}$ is played by either $\hat{n}_v(x,y)$ or $\hat{\boldsymbol{j}}_v(x,y)$, depending on the response function under consideration. Matrix elements of operator $\hat{n}_v(x,y)$ between stationary states are

$$\begin{aligned} [\hat{n}_v(x,y)]_{kn;k'n'} &= -\frac{e}{2}\langle k,n|\{\hat{n}(x,y), S(\hat{k})\}|k',n'\rangle \\ &= -\frac{e}{2}\big(S(k)+S(k')\big)\langle k,n|\hat{n}(x,y)|k',n'\rangle \\ &= -\frac{e}{4\pi}\big(S(k)+S(k')\big)e^{-i(k-k')x} \\ &\qquad \times u^\dagger_{k,n}(x,y)u_{k',n'}(x,y), \end{aligned} \tag{13}$$

where we have used the fact that, by definition, $\hat{n}(x,y) = |x,y\rangle\langle x,y|$ and $\langle x,y|k,n\rangle = \psi_{k,n}(x,y)$. For $k = k'$, Eq. (13) becomes

$$[\hat{n}_v(x,y)]_{kn;kn'} = -\frac{e}{2\pi}u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)S(k). \tag{14}$$

Matrix elements of operator $\hat{\boldsymbol{j}}_v(x,y)$ are given by

$$[\hat{\boldsymbol{j}}_v(x,y)]_{kn;k'n'} = -\frac{e}{2}\big(S(k)+S(k')\big)[\hat{\boldsymbol{j}}(x,y)]_{kn;k'n'}. \tag{15}$$

At $k' = k$ we obtain

$$[\hat{\boldsymbol{j}}_v(x,y)]_{kn;kn'} = -eS(k)[\hat{\boldsymbol{j}}(x,y)]_{kn;kn'}. \tag{16}$$

Note that due to Eq. (7) the following identity is true

$$\begin{aligned} \boldsymbol{\nabla}\langle k,n|\hat{\boldsymbol{j}}(x,y)|k,n'\rangle &= -i\langle k,n|[\hat{H}, \hat{n}(x,y)]|k,n'\rangle \\ &= i(\varepsilon_{k,n'} - \varepsilon_{k,n})\langle k,n|\hat{n}(x,y)|k,n'\rangle. \end{aligned} \tag{17}$$

Plugging into this equation $\hat{n}(x,y) = |x,y\rangle\langle x,y|$ and $\psi_{k,n}(x,y) = u_{k,n}(x,y)/\sqrt{2\pi}$ will give

$$\boldsymbol{\nabla}\big[\hat{\boldsymbol{j}}(x,y)\big]_{kn,kn'} = \frac{i(\varepsilon_{k,n'} - \varepsilon_{k,n})}{2\pi}u^\dagger_{k,n}(x,y)u_{k,n'}(x,y). \tag{18}$$

Let us integrate this equation along the length of one unit cell and using the periodicity of $[\boldsymbol{j}(x,y)]_{kn,kn'}$ in the x direction discard the boundary terms. The result will be

$$\begin{aligned} &\partial_y \int_0^1 dx\, \big[\hat{j}^y(x,y)\big]_{kn,kn'} \\ &\qquad = \frac{i(\varepsilon_{k,n'} - \varepsilon_{k,n})}{2\pi}\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y). \end{aligned} \tag{19}$$

Now using the fact that $\hat{j}^y$ must vanish at the boundary of the strip at $y = 0$, we can integrate this equation with respect to y to obtain

$$\begin{aligned} \int_0^1 dx\, \big[\hat{j}^y(x,y)\big]_{kn,kn'} &= \frac{i(\varepsilon_{k,n'} - \varepsilon_{k,n})}{2\pi}\int_0^y dy' \int_0^1 dx' \\ &\qquad \times u^\dagger_{k,n}(x',y')u_{k,n'}(x',y'). \end{aligned} \tag{20}$$

From Eqs. (16) and (20) we can now obtain that

$$\int_0^1 dx\, \big[\hat{j}^y_v(x,y)\big]_{kn,kn'} = -\frac{ieS(k)}{2\pi}\big(\varepsilon_{k,n'} - \varepsilon_{k,n}\big)\int_0^y dy' \int_0^1 dx'\, u^\dagger_{k,n}(x',y')u_{k,n'}(x',y'). \tag{21}$$

In calculating the linear response, we will average over the length of the unit cell in the x direction, so that Eq. (21) will turn out to be useful. The linear response formula for the valley Hall current, averaged over the length of the unit cell, has the form

$$j_v^y(y,\omega) \;=\; eE\int_0^1 dx\sum_{nn'}\int dk dk'\frac{(f_{k,n}-f_{k',n'})}{\omega+\varepsilon_{k,n}-\varepsilon_{k',n'}+i0}[j_v^y(x,y)]_{kn;k'n'}(-i\partial_k\delta(k-k')\delta_{nn'}+\delta(k-k')\mathcal{A}_{k,n'n}).\quad(22)$$

Integrating by parts the first term in the round brackets and then evaluating the integral with respect to k' will give

$$j_v^y(y,\omega) = \frac{ieE}{\omega+i0}\int_0^1 dx\sum_n\int dk\,\partial_k f_{k,n}\,[j_v^y(x,y)]_{kn;kn} + eE\int_0^1 dx\sum_{nn'}\int dk\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}[j_v^y(x,y)]_{kn;kn'}\mathcal{A}_{k,n'n}.\quad(23)$$

The first term vanishes because the unit cell average $\int_0^1 dx\,[j_v^y(x,y)]_{kn,kn}$ in that term vanishes, see Eq. (21) at $n=n'$. Plugging Eq. (21) into the second line, we obtain

$$j_v^y(y,\omega) = ie^2E\int_0^y dy'\sum_{nn'}\int\frac{dk}{2\pi}\frac{(f_{k,n}-f_{k,n'})(\varepsilon_{k,n}-\varepsilon_{k,n'})}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}\int_0^1 dx'\,u^\dagger_{k,n}(x',y')u_{k,n'}(x',y')S(k)\mathcal{A}_{k,n'n},\quad(24)$$

which gives Eq. (3) of the main text. Next, let us calculate the valley number response. According to Eq. (11), it is given by the equation, averaged across the length of the unit cell,

$$n_v(y,\omega) = eE\int_0^1 dx\sum_{nn'}\int dkdk'\frac{f_{k,n}-f_{k',n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k',n'}+i0}[n_v(x,y)]_{kn,k'n'}(-i\partial_k\delta(k-k')\delta_{nn'}+\delta(k-k')\mathcal{A}_{k,n'n}).\quad(25)$$

Integrating by parts the first term in the round brackets and using Eq. (14), we obtain

$$n_v(y,\omega) \;=\; -\frac{ie^2E}{\omega+i0}\sum_n\int\frac{dk}{2\pi}\partial_k f_{k,n}S(k)\int_0^1 dx\,u^\dagger_{k,n}(x,y)u_{k,n}(x,y) - e^2E\sum_{nn'}\int\frac{dk}{2\pi}\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}S(k)\int_0^1 dx\,u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n},\quad(26)$$

which is Eq. (4) in the main text. In the limit of zero frequency, if there is a partially filled band such that $\partial_k f_{k,n}$ does not vanish, the first term gives the dominant contribution. One can demonstrate that if the system is gapped and time-reversal symmetry is not broken, the second term on the right-hand side of Eq. (26) is of order $O(\omega)$ and, therefore, vanishes. This fact, which is demonstrated in Sect. VIII below, was used in the main text to discard its contribution to the valley density accumulation.

V. THE VALLEY DENSITY RESPONSE IS GAUGE INVARIANT

Let us demonstrate in this section that the valley density response, Eq. (4) in the main text, is gauge invariant. The first line of Eq. (4) is obviously gauge invariant so let us turn to the second line. This is proportional to

$$P(\omega+i0) = \sum_{n,n'}\int\frac{dk}{2\pi}S(k)\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}\times\int_0^1 dx\,u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n}\quad(27)$$

Let us change phases of all the stationary states as follows

$$u_{k,n}(x,y)\to e^{-i\phi_{k,n}}u_{k,n}(x,y).\quad(28)$$

After this the Berry connection $\mathcal{A}_{k,n'n}$ changes as

$$\mathcal{A}_{k,n'n} \to \int_0^1 dx \int_{-W}^0 dy\, u^\dagger_{k,n'}(x,y)\Big(i\partial_k u_{k,n}(x,y)\Big)e^{i(\phi_{k,n'}-\phi_{k,n})}$$
$$+\int_0^1 dx \int_{-W}^0 dy\, u^\dagger_{k,n'}(x,y)u_{k,n}(x,y)(\partial_k\phi_{k,n})e^{i(\phi_{k,n'}-\phi_{k,n})}. \qquad (29)$$

Using the normalization condition this can be rewritten as

$$\mathcal{A}_{k,n'n} \to \mathcal{A}_{k,n'n}e^{i(\phi_{k,n'}-\phi_{k,n})} + \delta_{n,n'}(\partial_k\phi_{k,n}). \qquad (30)$$

Let us now plug the modified states (28) into Eq. (27). This will result in the following change

$$P(\omega+i0) \;\to\; \sum_{n,n'}\int\frac{dk}{2\pi}S(k)\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)e^{i(\phi_{k,n}-\phi_{k,n'})}$$
$$\times\Big(\mathcal{A}_{k,n'n}e^{i(\phi_{k,n'}-\phi_{k,n})} + \delta_{n,n'}(\partial_k\phi_{k,n})\Big). \qquad (31)$$

The second term in the round brackets multiplies $f_{k,n}-f_{k,n'}$ and disappears. The phases in the rest of the equation cancel each other. Therefore, $P(\omega+i0)$ will not change, which means that it is gauge invariant.

VI. PROOF OF THE CONTINUITY EQUATION

In this section we demonstrate the validity of Eq. (1) in the main text. Let us multiply Eq. (26) by $-i\omega$. By representing ω as $\omega = \omega+\varepsilon_{k,n}-\varepsilon_{k,n'}-(\varepsilon_{k,n}-\varepsilon_{k,n'})$, we can rewrite the result as

$$-i\omega n_v(y,\omega) \;=\; -Ee^2\sum_n\int\frac{dk}{2\pi}\partial_k f_{k,n}S(k)\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n}(x,y)$$
$$+\,iEe^2\sum_{nn'}\int\frac{dk}{2\pi}(f_{k,n}-f_{k,n'})S(k)\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n}$$
$$-\,iEe^2\sum_{nn'}\int\frac{dk}{2\pi}\frac{(f_{k,n}-f_{k,n'})(\varepsilon_{k,n}-\varepsilon_{k,n'})}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}S(k)\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n}. \qquad (32)$$

Let us work on the sum in the second line of this equation,

$$I(y) \equiv \sum_{nn'}\int\frac{dk}{2\pi}(f_{k,n}-f_{k,n'})S(k)\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n}. \qquad (33)$$

Let us divide this sum in two across the minus sign and relabel n to n' and vica versa in the second sum. The result reads

$$I(y) = \sum_{nn'}\int\frac{dk}{2\pi}f_{k,n}S(k)\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n} - \sum_{nn'}\int\frac{dk}{2\pi}f_{k,n}S(k)\int_0^1 dx\, u^\dagger_{k,n'}(x,y)u_{k,n}(x,y)\mathcal{A}_{k,nn'}. \qquad (34)$$

Now let us plug into this equation the definition for $\mathcal{A}_{k,nn'}$

$$\mathcal{A}_{k,nn'} \equiv \int_0^1 dx'\int_{-W}^0 dy'\, u^\dagger_{k,n}(x',y')\Big(i\partial_k u_{k,n'}(x',y')\Big) = -\int_0^1 dx'\int_{-W}^0 dy'\Big(i\partial_k u^\dagger_{k,n}(x',y')\Big)u_{k,n'}(x',y'). \qquad (35)$$

The result will read

$$I(y) \;=\; \sum_{nn'}\int_0^1 dx\int_0^1 dx'\int_{-W}^0 dy'\int\frac{dk}{2\pi}f_{k,n}S(k)u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)u^\dagger_{k,n'}(x',y')\Big(i\partial_k u_{k,n}(x',y')\Big)$$
$$+\sum_{nn'}\int_0^1 dx\int_0^1 dx'\int_{-W}^0 dy'\int\frac{dk}{2\pi}f_{k,n}S(k)u^\dagger_{k,n'}(x,y)u_{k,n}(x,y)\Big(i\partial_k u^\dagger_{k,n}(x',y')\Big)u_{k,n'}(x',y'). \qquad (36)$$

Let us first perform summation over n' by using the completeness relation

$$\sum_{n'} u_{k,n'}(x,y)u^{\dagger}_{k,n'}(x',y') = 1_2 \sum_{l} \delta(y-y')\delta(x-x'-l), \tag{37}$$

where 1_2 is the identity operator in the spin space, and then perform integration over x' and y'. This will lead to

$$I(y) = \sum_n \int \frac{dk}{2\pi} f_{k,n} S(k) \int_0^1 dx\, u^{\dagger}_{k,n}(x,y)\Big(i\partial_k u_{k,n}(x,y)\Big) + \sum_n \int \frac{dk}{2\pi} f_{k,n} S(k) \int_0^1 dx \Big(i\partial_k u^{\dagger}_{k,n}(x,y)\Big) u_{k,n}(x,y). \tag{38}$$

Joining the two sums together we obtain

$$I(y) = \sum_n \int_0^1 dx \int \frac{dk}{2\pi} f_{k,n} S(k) i\partial_k \Big(u^{\dagger}_{k,n}(x,y) u_{k,n}(x,y)\Big). \tag{39}$$

Plugging this into Eq. (32), we obtain

$$\begin{aligned} -i\omega n_v(y,\omega) = &- Ee^2 \sum_n \int \frac{dk}{2\pi} \partial_k f_{k,n} S(k) \int_0^1 dx\, u^{\dagger}_{k,n}(x,y) u_{k,n}(x,y) \\ &- Ee^2 \sum_n \int \frac{dk}{2\pi} f_{k,n} S(k) \int_0^1 dx\, \partial_k \Big(u^{\dagger}_{k,n}(x,y) u_{k,n}(x,y)\Big) \\ &- iEe^2 \sum_{n,n'} \int \frac{dk}{2\pi} \frac{(f_{k,n}-f_{k,n'})(\varepsilon_{k,n}-\varepsilon_{k,n'})}{\omega + \varepsilon_{k,n} - \varepsilon_{k,n'} + i0} S(k) \int_0^1 dx\, u^{\dagger}_{k,n}(x,y) u_{k,n'}(x,y) \mathcal{A}_{k,n'n}. \end{aligned} \tag{40}$$

We combine the first two lines to obtain

$$\begin{aligned} -i\omega n_v(y,\omega) = &- Ee^2 \sum_n \int \frac{dk}{2\pi} \int_0^1 dx\, \partial_k \Big(f_{k,n} u^{\dagger}_{k,n}(x,y) u_{k,n}(x,y)\Big) S(k) \\ &- iEe^2 \sum_{n,n'} \int \frac{dk}{2\pi} \frac{(f_{k,n}-f_{k,n'})(\varepsilon_{k,n}-\varepsilon_{k,n'})}{\omega + \varepsilon_{k,n} - \varepsilon_{k,n'} + i0} S(k) \int_0^1 dx\, u^{\dagger}_{k,n}(x,y) u_{k,n'}(x,y) \mathcal{A}_{k,n'n}. \end{aligned} \tag{41}$$

From Eq. (24), the divergence of the current is obtained to be

$$\partial_y j^y_v(y,\omega) = iEe^2 \sum_{nn'} \int \frac{dk}{2\pi} \frac{(f_{k,n}-f_{k,n'})(\varepsilon_{k,n}-\varepsilon_{k,n'})}{\omega + \varepsilon_{k,n} - \varepsilon_{k,n'} + i0} S(k) \int_0^1 dx\, u^{\dagger}_{k,n}(x,y) u_{k,n'}(x,y) \mathcal{A}_{k,n'n}, \tag{42}$$

which is the term on the second line of Eq. (41). Combining Eq. (41) and (42), we obtain Eq. (1) in the main text.

VII. LIGHTNING SPEED DERIVATION OF THE CONTINUITY EQUATION

The continuity equation derived in the previous section can be easily derived using Heisenberg equations of motion. Everywhere below operators are given in the Heisenberg picture. The equation of motion for the valley density has the form

$$\partial_t \hat{n}_v(\boldsymbol{r},t) = i\big[\hat{H}(t) + eE(t)\hat{x}(t), \hat{n}_v(\boldsymbol{r},t)\big]. \tag{43}$$

Recall that in the absence of the electric field valley density satisfies a continuity equation, *i.e.*,

$$i\big[\hat{H}(t), \hat{n}_v(\boldsymbol{r},t)\big] = -\boldsymbol{\nabla}\hat{\boldsymbol{j}}_v(\boldsymbol{r},t), \tag{44}$$

with the current defined as in Eq. (10). Now using definition (6) and the fact that

$$[\hat{x}, S(\hat{k})] = iS'(\hat{k}), \tag{45}$$

where the prime indicates a derivative, one can obtain

$$[\hat{x}(t), \hat{n}_v(\boldsymbol{r},t)] = -\frac{ie}{2}\big\{S'\big(\hat{k}(t)\big), \hat{n}(\boldsymbol{r},t)\big\}. \tag{46}$$

Thus the equation of motion for the valley density has the form

$$\partial_t \hat{n}_v(\boldsymbol{r},t) + \boldsymbol{\nabla}\hat{\boldsymbol{j}}_v(\boldsymbol{r},t) = \frac{e^2 E(t)}{2}\big\{S'\big(\hat{k}(t)\big), \hat{n}(\boldsymbol{r},t)\big\}. \tag{47}$$

Replacing all the operators in this equation with their many-body versions and then calculating the ensemble average to first order in E will give Eq. (1) in the main text.

VIII. TIME REVERSAL SYMMETRY IMPLIES NO STATIC VALLEY POLARIZATION

Consider Eq. (4) in the main text and take the second line from that equation. It is proportional to

$$P(\omega+i0)=\sum_{n,n'}\int\frac{dk}{2\pi}S(k)\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}\int_0^1 dx\, u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)\mathcal{A}_{k,n'n}. \tag{48}$$

Assume presence of time-reversal symmetry which acts on the electron's wavefunction as

$$T\psi(x,y)=-i\sigma_y\psi^*(x,y). \tag{49}$$

For the Bloch states this implies that

$$(-i\sigma_y)u^*_{k,n}(x,y)=\alpha_{k,n}u_{-k,\tilde{n}}(x,y), \tag{50}$$

or

$$u^*_{k,n}(x,y)=\alpha_{k,n}i\sigma_y u_{-k,\tilde{n}}(x,y), \tag{51}$$

where $\tilde{n}$ labels another stationary state which satisfies $\varepsilon_{k,n}=\varepsilon_{-k,\tilde{n}}$ and $\alpha_{k,n}$ is a possible phase factor. Plugging Eq. (51) into the connection $\mathcal{A}_{k,n'n}$ in Eq. (48), we obtain

$$\begin{aligned}\mathcal{A}_{k,n'n}=i\int_0^1 dx\int_{-W}^0 dy\, u^t_{-k,\tilde{n}'}(x,y)\Big(\partial_k u^*_{-k,\tilde{n}}(x,y)\Big)\alpha_{k,n'}\alpha^*_{k,n}\\ +i\int_0^1 dx\int_{-W}^0 dy\,\alpha_{k,n'}\Big(\partial_k\alpha^*_{k,n}\Big)u^t_{-k,\tilde{n}'}(x,y)u^*_{-k,\tilde{n}}(x,y),\end{aligned} \tag{52}$$

where superscript t stands for 'transposed'. Interchanging the electron wavefunctions in the bilinear products and using the normalization condition, this equation can be rewritten as

$$\begin{aligned}\mathcal{A}_{k,n'n} &= i\int_0^1 dx\int_{-W}^0 dy\, u^\dagger_{-k,\tilde{n}}(x,y)\Big(-\partial_k u_{-k,\tilde{n}'}(x,y)\Big)\alpha_{k,n'}\alpha^*_{k,n}+i\alpha_{k,n}\Big(\partial_k\alpha^*_{k,n}\Big)\delta_{n,n'}\\ &= \mathcal{A}_{-k,\tilde{n}\tilde{n}'}\alpha_{k,n'}\alpha^*_{k,n}+i\alpha_{k,n}\Big(\partial_k\alpha^*_{k,n}\Big)\delta_{n,n'}.\end{aligned} \tag{53}$$

Analogously,

$$u^\dagger_{k,n}(x,y)u_{k,n'}(x,y)=\alpha_{k,n}\alpha^*_{k,n'}u^\dagger_{-k,\tilde{n}'}(x,y)u_{-k,\tilde{n}}(x,y). \tag{54}$$

Let us now substitute Eqs. (53) and (54) into Eq. (48). We obtain

$$\begin{aligned}P(\omega+i0) &= \sum_{n,n'}\int\frac{dk}{2\pi}S(k)\frac{f_{k,n}-f_{k,n'}}{\omega+\varepsilon_{k,n}-\varepsilon_{k,n'}+i0}\int_0^1 dx\, u^\dagger_{-k,\tilde{n}'}(x,y)u_{-k,\tilde{n}}(x,y)\alpha_{k,n}\alpha^*_{k,n'}\\ &\qquad\times\Big(\mathcal{A}_{-k,\tilde{n}\tilde{n}'}\alpha_{k,n'}\alpha^*_{k,n}+i\alpha_{k,n}\Big(\partial_k\alpha^*_{k,n}\Big)\delta_{n,n'}\Big).\end{aligned} \tag{55}$$

The term in the round brackets proportional to $\delta_{n,n'}$ gives a vanishing contribution to the sum and overall all the phases disappear. Let us also take into account that $\varepsilon_{k,n}=\varepsilon_{-k,\tilde{n}}$, $\varepsilon_{k,n'}=\varepsilon_{-k,\tilde{n}'}$ and, consequently, $f_{k,n}=f_{-k,\tilde{n}}$ and $f_{k,n'}=f_{-k,\tilde{n}'}$. Let us make use of these identities to make the summand only depend on $\tilde{n}$ and $\tilde{n}'$ and then also note that summation over n and n' is the same as summation over $\tilde{n}$ and $\tilde{n}'$. Changing the summation variables from the former to the latter and then relabelling these back to n and n', we obtain

$$P(\omega+i0)=\sum_{n,n'}\int\frac{dk}{2\pi}S(k)\frac{f_{-k,n}-f_{-k,n'}}{\omega+\varepsilon_{-k,n}-\varepsilon_{-k,n'}+i0}\int_0^1 dx\, u^\dagger_{-k,n'}(x,y)u_{-k,n}(x,y)\mathcal{A}_{-k,nn'}. \tag{56}$$

Changing the integration variable from k to $-k$, relabelling the summation indices from n to n' and vice versa, and using the fact that $S(-k) = -S(k)$, we obtain

$$P(\omega + i0) = -\sum_{n,n'} \int \frac{dk}{2\pi} S(k) \frac{f_{k,n} - f_{k,n'}}{-\omega + \varepsilon_{k,n} - \varepsilon_{k,n'} - i0} \int_0^1 dx\, u^\dagger_{k,n}(x,y) u_{k,n'}(x,y) \mathcal{A}_{k,n'n} = -P(-\omega - i0). \tag{57}$$

If the system is fully gapped, at small frequency the $i0$ prescription is irrelevant and can be neglected so that Eq. (57) implies that $P(\omega) = -P(-\omega)$, which means that the second line of Eq. (4) is of order $O(\omega)$ and can be neglected at zero frequency.

IX. VALLEY HALL CURRENT AT VANISHING FREQUENCY

At vanishing frequency, the following relation holds (compare Eqs. (24) and (33)):

$$j^y_v(y,0) = ie^2 E \int_0^y dy'\, I(y'). \tag{58}$$

Plugging the result (39) into this equation, we obtain

$$j^y_v(y,0) = -e^2 E \int_0^y dy' \int_0^1 dx' \sum_n \int \frac{dk}{2\pi} f_{k,n} S(k) \times \partial_k \Big(u^\dagger_{k,n}(x',y') u_{k,n}(x',y') \Big). \tag{59}$$

X. EFFECTIVE VALLEY CURRENT

Let us try to evaluate the integral

$$Q_s(y) = e^2 E \sum_n \int \frac{dk}{2\pi} \Big(\partial_k f_{k,n} \Big) S(k) \times \int_0^1 dx\, u^\dagger_{k,n}(x,y) u_{k,n}(x,y) \tag{60}$$

in the thermodynamic limit, *i.e.,* for the ribbon width $W \to \infty$. Assume that the chemical potential is in the conduction or valence band. Due to the factor $\partial_k f_{k,n}$ the integral over k and sum over n are restricted to the Fermi surface which consists of two disjoints parts, one in each valley. Due to time reversal symmetry the two parts give equal contributions to the integral, so we can calculate just one and multiply the result by two, *i.e.,*

$$Q_s(y) = 2e^2 E \sum_n \int_{S(k)=1} \frac{dk}{2\pi} \Big(\partial_k f_{k,n} \Big) \times \int_0^1 dx\, u^\dagger_{k,n}(x,y) u_{k,n}(x,y). \tag{61}$$

Everything from now on refers to the valley with valley number $+1$. Assuming that in the sum over n and integral over k we never go too far away from the bottom of the valley, we can use the envelope wave function description for the stationary states. Suppose energy eigenstates for a system without boundaries are described by a set of multicomponent valence (conduction) band envelope amplitudes $v^\lambda_{k,p}$ with energies $\varepsilon_\lambda(k,p)$, where p is the component of momentum along y measured from the bottom of the valley and λ is a discrete label counting the stationary states. Now let us introduce the boundaries at $y = 0$ and $y = -W$. Assume that the scattering off the boundaries does not mix eigenstates with different values of λ and that $\varepsilon_\lambda(k,-p) = \varepsilon_\lambda(k,+p)$. Then the valence (conduction) band envelope wave-functions for the system with boundaries will have the form

$$u^\lambda_{k,m}(y) = \mathcal{N}_\lambda(k,p_m) \Big(v^\lambda_{k,p_m} e^{ip_m y} + R_\lambda(k,p_m) v^\lambda_{k,-p_m} e^{-ip_m y} \Big), \tag{62}$$

which is nothing but a sum of an incident and a reflected wave. Here $R_\lambda(k,p)$ is the probability amplitude for scattering off the boundary at $y = 0$, momenta p_m take discrete positive values (labelled by m), with the distance between them approaching π/W as $W \to \infty$, and an overall factor $\mathcal{N}_\lambda(k,p_m)$ is chosen such that

$$\int_{-W}^0 dy\, [u^\lambda_{k,m}]^\dagger(y) u^\lambda_{k,m}(y) = 1. \tag{63}$$

Note that at large W

$$|\mathcal{N}_\lambda(k,p)|^2 = \frac{1}{2W} + O\Big(\frac{1}{W^2}\Big). \tag{64}$$

In terms of the envelope wave-functions the expression for $Q_s(y)$ takes the form

$$Q_s(y) = 2e^2 E \sum_{m,\lambda} \int_{S(k)=1} \frac{dk}{2\pi} \Big(\partial_k f^\lambda_{k,m} \Big) [u^\lambda_{k,m}]^\dagger(y) u^\lambda_{k,m}(y), \tag{65}$$

Plugging Eq. (62) into this equation, we obtain

$$Q_s(y) = 4e^2 E \sum_{m,\lambda} \int_{S(k)=1} \frac{dk}{2\pi} \Big(\partial_k f^\lambda_{k,m} \Big) \times \Big(1 + \mathfrak{Re} \Big([v^\lambda_{k,p_m}]^\dagger v^\lambda_{k,-p_m} R_\lambda(k,p_m) e^{-2ip_m y} \Big) \Big), \tag{66}$$

where we used the normalization condition $|v^\lambda_{k,p}|^2 = 1$ and the fact that $|R_\lambda(k,p)| = 1$. Let us consider values of y such that $|y| \ll W$, which means that we are keeping very close to the edge at $y = 0$. In this case each term in the sum over m is not much different from the next one and we can exchange the sum over m for an integral with respect to p. More precisely, to evaluate the sum over m we can use the Euler–Maclaurin formula, which will give us an expansion in powers of $1/W$ and the leading-order term is obtained by simply exchanging the sum over m for an integral over m. This in turn can be exchanged for an integral over p via $p = \pi m/W$. Taking into account Eq. (64), this will lead to

$$\begin{aligned} Q_s(y) = 4e^2 E \sum_\lambda \int_{p>0} \frac{dp}{2\pi} \int_{S(k)=1} \frac{dk}{2\pi} \left(\partial_k f^\lambda_{k,p}\right) \\ \times \left(1 + \mathfrak{Re}\left([v^\lambda_{k,p}]^\dagger v^\lambda_{k,-p} R_\lambda(k,p) e^{-2ipy}\right)\right), \end{aligned} \tag{67}$$

where we neglected the terms of order $O(1/W)$. Contribution of the first term in the outer round brackets in the second line to the integral is equal to zero as $f^\lambda_{k,p} = \theta(\varepsilon_\mathrm{F} - \varepsilon_\lambda(k,p))$ vanishes at the values of k far enough from the bottom of the valley. Therefore we are left with

$$\begin{aligned} Q_s(y) = 4e^2 E \sum_\lambda \int_{p>0} \frac{dp}{2\pi} \int_{S(k)=1} \frac{dk}{2\pi} \left(\partial_k f^\lambda_{k,p}\right) \\ \times \mathfrak{Re}\left([v^\lambda_{k,p}]^\dagger v^\lambda_{k,-p} R_\lambda(k,p) e^{-2ipy}\right). \end{aligned} \tag{68}$$

By Riemann–Lebesgue lemma, $Q_s(y) \to 0$ (up to terms of order $O(1/W)$ that we discarded) as $y \to \infty$. Because integration in Eq. (68) is restricted to the Fermi surface, it is clear that $Q_s(y)$ oscillates in space at a wavelength corresponding to double the Fermi momentum. We can now introduce an "effective current" feeding valley number accumulation at the edge

$$\begin{aligned} I^b_s &= \int_{-\infty}^{0} dy\, Q_s(y) \\ &= -2e^2 E \sum_\lambda \int_{S(k)=1} \frac{dk}{2\pi} \int_{p>0} \frac{dp}{2\pi} \left(\partial_k f^\lambda_{k,p}\right) \\ &\quad \times \mathfrak{Im}\left[\frac{1}{p+i0} [v^\lambda_{k,p}]^\dagger v^\lambda_{k,-p} R_\lambda(k,p)\right]. \end{aligned} \tag{69}$$

XI. GRAPHENE NANORIBBON. DESCRIPTION OF THE STATIONARY STATES

In this section and the next we prove analytically that in the completely gapped state the valley Hall current for a graphene nanoribbon with zig-zag edges is non-zero and quantized. The treatment here closely follows that of Ref. [4]. Because the spin-orbit coupling in graphene is small, we will neglect it completely. In this case spin polarization is a good quantum number and its only effect is the introduction of a degeneracy factor of 2 in all linear response formulae. We will avoid that by considering a spinless electron from now on. For a spinful electron the results given here will have to be multiplied by a factor of 2.

As promised in Section I, the pair of coordinates (x, y) will be replaced by unit cell number l and combined index (m, σ), where $m = 1 \ldots N$ and $\sigma = A, B$. Integrals over x' and y' within the unit cell will be replaced as follows

$$\int_0^1 dx' \int_0^y dy' \to -\sum_{m'=1}^{m} \sum_{\sigma=A,B} . \tag{70}$$

Furthermore, in these expressions, the current that flows between rows m and $m+1$ is denoted as $j^y_v(m+1/2, \omega)$.

The tight binding Hamiltonian for graphene nanoribbon with zig-zag edges in second quantized form reads

$$\begin{aligned} \hat{H} = -t \sum_l \Bigg[\sum_{m=1}^{N} \hat{a}^\dagger_l(m) \hat{b}_l(m) + \sum_{m=1}^{N-1} \hat{a}^\dagger_l(m+1) \hat{b}_l(m) \\ + \sum_{m\in even} \hat{a}^\dagger_l(m) \hat{b}_{l-1}(m) \\ + \sum_{m\in odd} \hat{a}^\dagger_l(m) \hat{b}_{l+1}(m) + \text{h.c.} \Bigg] \\ + \Delta \sum_l \sum_{m=1}^{N} \left[\hat{a}^\dagger_l(m) \hat{a}_l(m) - \hat{b}^\dagger_l(m) \hat{b}_l(m) \right], \end{aligned} \tag{71}$$

where t is the nearest neighbour hopping parameter and $\hat{a}_l(m)$ and $\hat{b}_l(m)$ destroy an electron on sublattice A or B, respectively, in a two-atom row m of unit cell l. They satisfy the usual anticommutation relations

$$\{\hat{a}_l(m), \hat{a}^\dagger_{l'}(m')\} = \{\hat{b}_l(m), \hat{b}^\dagger_{l'}(m')\} = \delta_{ll'} \delta_{mm'} . \tag{72}$$

We introduce the Fourier-transformed operators $\hat{\alpha}_k(m)$ and $\hat{\beta}_k(m)$ from the relations

$$\hat{a}_l(m) = \int \frac{dk}{\sqrt{2\pi}} \hat{\alpha}_k(m) e^{ikx_{l,m,A}}, \tag{73}$$

$$\hat{b}_l(m) = \int \frac{dk}{\sqrt{2\pi}} \hat{\beta}_k(m) e^{ikx_{l,m,B}}, \tag{74}$$

where $x_{l,m,A}$ and $x_{l,m,B}$ are the positions of the atoms of the A and B sublattices in the direction along the ribbon and k takes values in the interval $(-\pi, \pi]$. The Fourier-transformed creation and annihilation operators satisfy the anticommutation relations

$$\begin{aligned} \{\hat{\alpha}_k(m), \hat{\alpha}^\dagger_{k'}(m')\} &= \{\hat{\beta}_k(m), \hat{\beta}^\dagger_{k'}(m')\} \\ &= \delta(k-k') \delta_{mm'} . \end{aligned} \tag{75}$$

In terms of Fourier transformed creation and annihilation operators, the Hamiltonian reads

$$\begin{aligned}\hat{H} = & \int dk\Big[-t\sum_{m=1}^{N}\hat{\alpha}_k^\dagger(m)\hat{\beta}_k(m)g_k + \text{h.c.} \\ & -t\sum_{m=1}^{N-1}\hat{\alpha}_k^\dagger(m+1)\hat{\beta}_k(m) + \text{h.c.} \\ & +\Delta\sum_{m=1}^{N}\Big(\hat{\alpha}_k^\dagger(m)\hat{\alpha}_k(m) - \hat{\beta}_k^\dagger(m)\hat{\beta}_k(m)\Big)\Big], \end{aligned} \quad (76)$$

where $g_k = 2\cos(k/2)$. From now on, we will set $t = 1$. The one-particle eigenstates of the Hamiltonian are Bloch states $\psi_{k,n}(l,m,\sigma) = u_{k,n}(m,\sigma)\exp(ikx_{l,m,\sigma})/\sqrt{2\pi}$. As is customary, we will combine the amplitudes $u_{k,n}(m,\sigma)$ for $\sigma = A, B$ into a sublattice pseudospinor $u_{k,n}(m) = (u_{k,n}(m,A),\ u_{k,n}(m,B))^t$. For bulk states, this takes the form

$$u_{kps}(m) = \mathcal{N}_{kps}\begin{pmatrix}(\varepsilon_{kps}+\Delta)\sin\big[p(N+1-m)\big] \\ (-1)^j\sqrt{\varepsilon_{kps}^2-\Delta^2}\sin(pm)\end{pmatrix}, \quad (77)$$

where $\varepsilon_{kps} = s\sqrt{\Delta^2 + g_k^2 + 2g_k\cos(p) + 1}$ and the role of the band index n is played by the combination ps, with $s = \pm$ and $p \in (0,\pi)$ the solution of the equation

$$pN + \arccos\left(\frac{1+g_k\cos p}{\sqrt{g_k^2+2g_k\cos p+1}}\right) = \pi j. \quad (78)$$

In Eqs. (77)–(78), $j = 1, 2, \ldots, N$ for $g_k > N/(N+1)$ and $j = 1, 2, \ldots, N-1$ for $g_k < N/(N+1)$.

In Eq. (77) the normalization constant equals

$$\begin{aligned}\mathcal{N}_{kps} = \Bigg[& (\varepsilon_{kps}+\Delta)\varepsilon_{kps} \\ & \times\left(N - \frac{\sin(pN)\cos\big[p(N+1)\big]}{\sin(p)}\right)\Bigg]^{-1/2}.\end{aligned} \quad (79)$$

We note that, for $g_k < N/(N+1)$, apart from the states described by Eqs. (77)–(78), there is also an edge state

$$\begin{aligned} & u_{kes} \\ & = \mathcal{N}_{kes}(-1)^m\begin{pmatrix}(\varepsilon_{kes}+\Delta)\sinh\big[(N+1-m)\eta_k\big] \\ -\sqrt{\varepsilon_{kes}^2-\Delta^2}\sinh(\eta_k m)\end{pmatrix},\end{aligned} \quad (80)$$

where $\varepsilon_{kes} = s\sqrt{\Delta^2 + g_k^2 - 2g_k\cosh(\eta_k) + 1}$, with $s = \pm$, and $\eta_k > 0$ is the solution of the equation

$$e^{2\eta_k(N+1)} = \frac{g_k - e^{\eta_k}}{g_k - e^{-\eta_k}}. \quad (81)$$

The role of the band index n here is played by the combination es, where e stands for 'edge' and s is described above. The normalization factor equals

$$\begin{aligned}\mathcal{N}_{kes} = \Bigg[& \varepsilon_{kes}(\varepsilon_{kes}+\Delta) \\ & \times\left(\frac{\cosh\big[(N+1)\eta_k\big]\sinh(N\eta_k)}{\sinh(\eta_k)} - N\right)\Bigg]^{-1/2}.\end{aligned} \quad (82)$$

XII. CURRENT IS QUANTIZED FOR THE TOTALLY GAPPED GRAPHENE NANORIBBON

Let us calculate the valley Hall current for the totally gapped graphene nanoribbon with zig-zag edges in the limit of vanishing frequency. Using Eq. (59) (and neglecting spin) we obtain

$$\begin{aligned} j_v^y(m+1/2,0) \; = \; & e^2E\sum_n\int\frac{dk}{2\pi}f_{k,n}\,S(k) \\ & \times\sum_{m'=1}^{m}\partial_k(u_{k,n}^\dagger(m')u_{k,n}(m')), \end{aligned} \quad (83)$$

where n stands for either ps or es, as explained in Sect. XI. Note that explicit summation over σ here is replaced by matrix multiplication of the hermitian conjugate of a pseudospinor with itself. When the chemical potential is in the gap, $f_{kps} = f_{kes}$ and equals 1 for $s = -1$ and 0 for $s = +1$. For the evaluation of the integral with respect to k note that for any n holds $u_{k,n} = u_{-k,n}$ and we can take $S(k)$ equal to 1 for $0 < k < \pi$ and -1 for $-\pi < k < 0$. Thus keeping only the occupied states in the sum over n and evaluating the integral with respect to k, we obtain

$$\begin{aligned} j_v^y(m+1/2,0) \; = \; & \frac{e^2E}{\pi}\sum_{m'=1}^{m}\Bigg(\sum_p u_{kp-}^\dagger(m')u_{kp-}(m') \\ & + \; u_{ke-}^\dagger(m')u_{ke-}(m')\Bigg)\Bigg|_{k=0}^{k=\pi}. \end{aligned} \quad (84)$$

The states $u_{k,n}$ form a complete set, therefore

$$\begin{aligned} & \sum_p\Big(u_{kp+}(m')u_{kp+}^\dagger(m') + u_{kp-}(m')u_{kp-}^\dagger(m')\Big) \\ & + u_{ke+}(m')u_{ke+}^\dagger(m') + u_{ke-}(m')u_{ke-}^\dagger(m') = 1_2, \end{aligned} \quad (85)$$

where 1_2 is the 2×2 identity matrix in the sublattice space. Using the explicit form of the wavefunctions [see Eq. (77)] one can find that

$$\begin{aligned} & u_{kp-}^\dagger(m')u_{kp-}(m') - u_{kp+}^\dagger(m')u_{kp+}(m') \\ & \quad = \frac{2\Delta}{\varepsilon_{kp-}}\cdot\frac{\sin\big[p(N+1-2m')\big]\sin\big[p(N+1)\big]}{N - \sin(pN)\cos\big[p(N+1)\big]/\sin(p)}. \end{aligned} \quad (86)$$

Using Eqs. (85) and (86) one can rewrite the sum in the round brackets in Eq. (84) in the form

$$\sum_p u^\dagger_{kp-}(m')u_{kp-}(m') + u^\dagger_{ke-}(m')u_{ke-}(m') = 1 \;+\; \frac{1}{2}\Big(u^\dagger_{ke-}(m')u_{ke-}(m') - u^\dagger_{ke+}(m')u_{ke+}(m')\Big) + \sum_p \frac{\Delta}{\varepsilon_{kp-}}\cdot\frac{\sin\big[p(N+1-2m')\big]\sin\big[p(N+1)\big]}{N-\sin(pN)\cos\big[p(N+1)\big]/\sin(p)}. \tag{87}$$

Plugging this into Eq. (84) and taking into account the fact that there are no edge states at $k=0$ we obtain

$$j^y_v(m+1/2,0) \;=\; \sum_{m'=1}^{m}\frac{e^2E}{2\pi}\Big(u^\dagger_{\pi e-}(m')u_{\pi e-}(m') - u^\dagger_{\pi e+}(m')u_{\pi e+}(m')\Big) + \frac{e^2E}{\pi}\sum_p \frac{\Delta}{\varepsilon_{kp-}}\cdot\frac{\sin(pm)\sin\big[p(N-m)\big]\sin\big[p(N+1)\big]/\sin(p)}{N-\sin(pN)\cos\big[p(N+1)\big]/\sin(p)}\bigg|_{k=0}^{k=\pi}. \tag{88}$$

Let us estimate the sum over p in Eq. (88). Since by Eq. (78), the factors

$$\frac{\sin\big[p(N+1)\big]}{\sin(p)} = \frac{(-1)^j}{\sqrt{g_k^2+2g_k\cos(p)+1}}, \tag{89}$$

$$\frac{\sin(pN)}{\sin(p)} = \frac{(-1)^{j-1}g_k}{\sqrt{g_k^2+2g_k\cos(p)+1}}, \tag{90}$$

remain of order $O(N^0)$ at $k=0$ or $k=\pi$, one can see that each term in the sum in the second line of Eq. (88) is of order Δ/N. In making this estimate we also took into account that ε_{kp-} is of order one [in units of t] at $k=0$ and $k=\pi$. Therefore the whole sum over p in Eq. (88) is of order Δ. Hence the current density response equals (restoring the hopping parameter t)

$$j^y_v(m+1/2,0) = \sum_{m'=1}^{m}\frac{e^2E}{2\pi}\Big(u^\dagger_{\pi e-}(m')u_{\pi e-}(m') - u^\dagger_{\pi e+}(m')u_{\pi e+}(m')\Big) + O(\Delta/t). \tag{91}$$

For edge states, as k approaches π, $g_k\to 0$ while η_k approaches positive infinity as $\eta_k\propto -\ln(g_k)$. In this limit $\varepsilon_{kes}\to s|\Delta|$ and the probability distribution for the edge states has the form

$$u^\dagger_{\pi es}(m')u_{\pi es}(m') \;=\; \frac{1}{2}\bigg(\Big[1+s\cdot\mathrm{sign}(\Delta)\Big]\delta_{m',1} + \Big[1-s\cdot\mathrm{sign}(\Delta)\Big]\delta_{m',N}\bigg). \tag{92}$$

Plugging this into Eq. (91), for $1\le m\le N-1$ we obtain

$$j^y_v(m+1/2,0) = -\frac{e^2E}{2\pi}\mathrm{sign}(\Delta) + O(\Delta/t). \tag{93}$$

Let us also, for future reference, quote here the result for the divergence of the current, $\mathrm{div}\, j^y_v(m,0) = j^y_v(m-1/2,0) - j^y_v(m+1/2,0)$, which follows from this equation

$$\mathrm{div}\, j^y_v(m,0) = \frac{e^2E}{2\pi}\mathrm{sign}(\Delta)(\delta_{m,1}-\delta_{m,N}) + O(\Delta/t). \tag{94}$$

XIII. QUALITATIVE EXPLANATION OF THE VALLEY HALL CURRENT PROFILE FOR DIFFERENT BOUNDARY CONDITIONS

Throughout this section we assume that the Fermi level lies in the conduction band and takes a value in the interval $|\Delta| < \varepsilon_{\mathrm{F}} < t$. Let us first consider the zig-zag nanoribbon. Expression for the current, Eq. (83), can be written in an alternative form as $j^y_v(m+1/2,0) = Ee^2(2\pi)^{-1}\big[T(m+1/2)+G(m+1/2)\big]$, where for $1\le m\le N-1$

$$T(m+1/2) = \sum_n\sum_{m'=1}^{m}\int dk S(k)\partial_k\Big(f_{k,n}u^\dagger_{k,n}(m')u_{k,n}(m')\Big) \tag{95}$$

and

$$G(m+1/2) = -\sum_n\sum_{m'=1}^{m}\int dk S(k)(\partial_k f_{k,n})u^\dagger_{k,n}(m')u_{k,n}(m') \tag{96}$$

and we set $T(1/2)=G(1/2)=T(N+1/2)=G(N+1/2)=0$. Let us consider T first. Consider $\mathrm{div}\, T(m) = T(m-1/2)-T(m+1/2)$, which is given by the equation

$$\mathrm{div}\, T(m) = -\sum_n\int dk S(k)\partial_k\Big(f_{k,n}u^\dagger_{k,n}(m)u_{k,n}(m)\Big). \tag{97}$$

The integral with respect to k can be easily calculated to produce

$$\mathrm{div}\, T(m) = -2\sum_n f_{k,n}u^\dagger_{k,n}(m)u_{k,n}(m)\Big|_{k=0}^{k=\pi}, \tag{98}$$

where a factor of 2 appeared because the contribution of the left valley to the integral equals the contribution of the right valley, so we left only the latter and multiplied it by two. Plugging in the occupation numbers, we obtain

$$\begin{aligned}\operatorname{div} T(m) = &-2\, u^{\dagger}_{\pi e+}(m) u_{\pi e+}(m) \\ &-2\bigg(\sum_p u^{\dagger}_{kp-}(m) u_{kp-}(m) \\ &\qquad + u^{\dagger}_{ke-}(m) u_{ke-}(m)\bigg)\bigg|_{k=0}^{k=\pi}. \end{aligned} \quad (99)$$

The sum in the second and third lines on the right hand side of this equation (including the factor of -2) has already been calculated in the previous section, compare Eq. (84). It is given by whatever multiplies $e^2E/(2\pi)$ in Eq. (94). The term in the first line is the contribution of the upper band of edge states, which has appeared because we raised the Fermi energy and this band became occupied. So, using Eq. (92) and Eq. (94), we obtain (ignoring terms of order $O(\Delta/t)$)

$$\operatorname{div} T(m) = -\delta_{m,1} - \delta_{m,N}. \quad (100)$$

Note that $\operatorname{div} T$ changes extremely fast on the boundaries (it goes from -1 to zero on the scale of one inter-atomic distance) and does not change at all inside the ribbon. This can be traced back to the fact that ultimately the spatial behavior of $T(m)$ is governed by the localized edge states (see Eq. (91) and Eq. (99)).

Consider now $\operatorname{div} G(m)$, which is given by the equation

$$\operatorname{div} G(m) = \sum_n \int dk S(k)(\partial_k f_{k,n}) u^{\dagger}_{k,n}(m) u_{k,n}(m). \quad (101)$$

Let us point out that $\operatorname{div} G(m)$, as opposed to $\operatorname{div} T(m)$, changes slowly in space, because its behavior is governed by the states on the Fermi surface. Therefore it varies significantly on the length scale defined by the inverse Fermi momentum $(\sqrt{\varepsilon_{\mathrm{F}}^2 - \Delta^2})^{-1} \gg 1$ for ε_{F} close enough to the bottom of the conduction band. Therefore the roughest (but only the roughest) estimate of $\operatorname{div} G(m)$ can be given by just the spatial average $\langle \operatorname{div} G(m)\rangle_{\mathrm{sp}} = (1/N)\sum_m \operatorname{div} G(m)$. This is not zero. Indeed, using $\sum_m u^{\dagger}_{k,n}(m) u_{k,n}(m) = 1$ we obtain

$$\langle \operatorname{div} G(m)\rangle_{\mathrm{sp}} = \frac{1}{N}\sum_n \int dk S(k)(\partial_k f_{k,n}), \quad (102)$$

which is non-zero because of the left-mover–right-mover imbalance in each valley. Indeed, using $f_{k,n} = \theta(\varepsilon_{\mathrm{F}} - \varepsilon_{k,n})$ we obtain

$$\sum_n \partial_k f_{k,n} = -\sum_i \delta(k - k_i)\operatorname{sign}(v_i), \quad (103)$$

where the sum runs over all values of k at which the Fermi level crosses an energy band and v_i is the group velocity of the band crossed at a point k_i. Because there is one more right-mover than there are left-movers in the left valley and one more left-mover than there are right-movers in the right valley, see Fig. 1(b) (the upper blue line) in the main text, we obtain

$$\langle \operatorname{div} G(m)\rangle_{\mathrm{sp}} = \frac{2}{N}. \quad (104)$$

To get $j^y_v(m+1/2, 0)$ we need to integrate (*i.e.*, sum over m) Eq. (100) and Eq. (104) with the boundary conditions that the current vanishes outside the ribbon. It is not difficult to observe that Eq. (100) determines the one-sided limiting values of the current on the boundaries as approached from within the ribbon and Eq. (104) its overall slope as a function of position. It then follows that the current $j^y_v(m+1/2, 0)$ is approximately equal to $Ee^2/(2\pi)$ at $m = 1$, to $-Ee^2/(2\pi)$ at $m = N-1$ and those two values are connected roughly by a straight line with the slope $-[e^2E/(2\pi)](2/N)$.

Now let us briefly discuss the case of the nanoribbon with a bearded edge. Because there is only one band of edge states whose occupation number does not change as we raise ε_{F}, we see that $\operatorname{div} T$ does not change compared to the undoped case. This means that the values of the current on the boundaries will stay roughly the same as in the undoped case. Next, because there is no left-mover–right-mover imbalance in the valleys, the overall average slope will be zero. This very rough analysis is confirmed by the numerical results, see Fig. 2 (a), (b) in the main text.

XIV. VALLEY HALL CURRENT AS FUNCTION OF THE FERMI ENERGY

In this section we provide numerical results for the valley Hall current as a function of the Fermi energy in graphene nanoribbon with one edge zigzag and the other

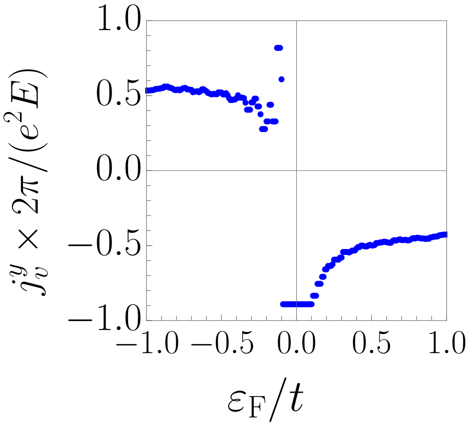


FIG. 1. Valley Hall current as a function of the Fermi energy in graphene nanoribbon with one edge zig-zag and the other bearded. The value is taken in the middle of the ribbon, the width $N = 100$, $\Delta = 0.1t$.

bearded, see Fig. 1. In an infinite system this is predicted to be fixed and quantized when the chemical potential is in the band gap and to go down to zero as $-\Delta/|\varepsilon_{\mathrm{F}}|$ (in units of $e^2 E/(2\pi)$) for $|\varepsilon_{\mathrm{F}}| > |\Delta|$. For the nanoribbon the behavior of the curve is a bit different, see Fig. 1. It is indeed fixed and quantized (up to corrections of order $O(\Delta/t)$) when the Fermi energy is in the gap but outside the gap it does not conform to the $\Delta/|\varepsilon_{\mathrm{F}}|$ law and settles on a value of around $\pm 1/2$ for high enough hole or electron doping.

At $\varepsilon_{\mathrm{F}} = -|\Delta|$ the valley Hall current has a discontinuity due to all the edge states suddenly changing their occupation numbers. The oscillations below $\varepsilon_{\mathrm{F}} = -|\Delta|$ are a finite size effect due to small values of the Fermi momentum.

[1] E. Lifshitz and L. P. Pitaevskii, *Statistical Physics. Theory of the Condensed State* (Butterworth-Heinemann, Oxford, 1980).

[2] R. Kubo, Journal of the Physical Society of Japan **12**, 570 (1957).

[3] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[4] K. Wakabayashi, K. Sasaki, T. Nakanishi, and T. Enoki, Science and Technology of Advanced Materials **11**, 054504 (2010), pMID: 27877361, https://doi.org/10.1088/1468-6996/11/5/054504.

Nonconserved Density Accumulations in Orbital Hall Transport: Insights from Linear Response Theory

Hao Sun,[1,*] Alexander Kazantsev,[2] Alessandro Principi,[2] and Giovanni Vignale[1,†]
[1]*The Institute for Functional Intelligent Materials (I-FIM),*
National University of Singapore, 4 Science Drive 2, Singapore 117544
[2]*Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK*

(Dated: November 19, 2024)

We present a linear response theory for stationary density accumulations in anomalous transport phenomena, such as the orbital Hall effect, where the transported density is odd under time reversal and the underlying charge is not conserved. Our framework applies to both metals and insulators, topologically trivial or nontrivial, and distinguishes between contributions from bulk and edge states, as well as undergap and dissipative currents. In time-reversal invariant systems, we prove a microscopic reciprocity theorem showing that only dissipative currents at the Fermi level contribute to density accumulation, while undergap currents do not. In contrast, in non-time-reversal invariant systems, non-dissipative density accumulations, such as magnetoelectric polarization, can appear in both the bulk and edges. Importantly, we find that the net density accumulation does not always vanish, pointing to a global non-conservation that implies the existence of a non-vanishing integrated "net torque" in addition to a "distributed torque", which has zero spatial average. We show that the distributed torque can be absorbed in the divergence of a redefined current that satisfies Onsager reciprocity, while the net torque must be explicitly accounted for. Finally, we apply our theory to two-dimensional models with edge terminations.

I. INTRODUCTION

Recent years have witnessed a revolution in the theory of electronic transport in crystalline materials [1–5]. Traditional classifications of "metals" and "insulators" have given way to a more nuanced classification, recognizing that some materials can behave as insulators in the bulk while exhibiting metallic properties on their surfaces or edges. At the same time, quantum geometric features of the band structure have been shown to profoundly influence transport phenomena [6–8]. In particular, "Berry curvature" has emerged as a crucial actor, serving as a momentum space analogue of the magnetic field. Following this, many conducting materials have been found to support transverse currents (perpendicular to the electric field), which are termed "anomalous" because they arise in the absence of an external magnetic field [6, 7, 9–13].

While the anomalous Hall effect is the best-known example in this class of phenomena, our focus in this paper is on effects that entail the generation of transverse currents of *non-conserved* quantities, such as spin, valley, and orbital magnetic moment (OMM), all of which are electrically neutral and odd under time reversal [14–17]. Although these currents are difficult to observe directly, they generally lead to accumulations of the corresponding densities on the surfaces or edges of the systems in which they flow. The practical importance of these "density accumulations" cannot be overstated since they are a direct and observable manifestation of bulk currents, and provide a natural way to connect these currents to external devices [18–22]. However, the process by which density accumulations are established is surprisingly complex and not completely understood, even in the well-studied case of the electric anomalous Hall effect (AHE), where the accumulation involves the conserved electric charge (this deceptively simpler case will be discussed in the concluding section).

Our primary aim in this paper is to establish a versatile linear response formalism for calculating not just the currents (as it is customary) but also the density accumulations in systems with edges or surfaces. Interest in this problem is stimulated by a flurry of experimental papers reporting observations of orbital and valley density accumulations in transition metals and 2D layered materials [22–24]. In particular, the orbital Hall effect (OHE) is known to be stronger than the spin Hall effect (SHE) in principle, as it does not require spin-orbit coupling. While extensive theoretical work has been done on calculating orbital and spin Hall conductivities using linear response theory [25, 26], our work distinguishes itself by focusing on observable OMM density accumulations and on the complications arising from the fact that the underlying "charge" (OMM in this case) is not conserved. In this context, the linear response theory is free of certain ambiguities that plague the "modern theory" of semiclassical transport [27–30]. For example, modern theory views the Hall effect as the manifestation of an "anomalous velocity", inviting the question whether this anomalous velocity should also be included in the calculation of the orbital magnetic moments and torques. But in linear response theory, there is no "anomalous velocity" and the definition of the key observables is unambiguous (see Section II).

The model system we focus on in this paper is illustrated in Fig. 1(a). A periodic system with edge or surface terminations that cut across one or two directions (y, z) is subjected to an electric field parallel to the open direction (x) so that an electric charge current may or may not flow depending on whether the system is a metal, a trivial insulator, or a topological insulator [31]. Typical band structures of our model system (practically realized, for instance, in nanoribbons of honeycomb lattices) are shown in Fig. 1(b)-(d).

An important topological distinction emerges according to whether the edge states connect states on the same side of

---

[*] sun.hao@nus.edu.sg
[†] vgnl.g@nus.edu.sg

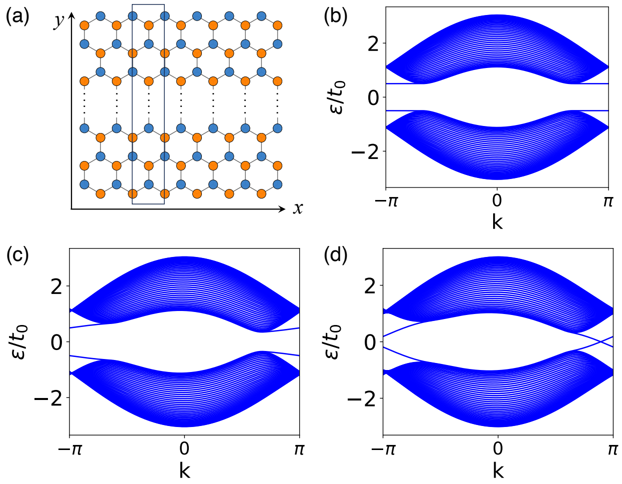


FIG. 1. (a) Nanoribbon model for the OMM density accumulation. (b) Band structure for graphene nanoribbon with staggered sublattice potentials. (c) Band structure for the topologically trivial phase of the Haldane model nanoribbon with staggered sublattice potentials. (d) Band structure for the topologically nontrivial phase of the Haldane model nanoribbon with staggered sublattice potentials. In these figures, the wave vector k is parallel to the edges and expressed in units of inverse lattice constant a^{-1}.

the gap or run across the gap between the bands. In the first case, illustrated in Fig. 1(b)-(c), the system is said to be topologically trivial and will be considered (i) an insulator, (ii) an "edge metal", or (iii) a metal, depending on whether the Fermi level lies (i) within the gap without crossing any edge states, (ii) crosses only edge states, or (iii) crosses bulk states. In the topologically nontrivial case, illustrated in Fig. 1(d), the system will be considered a topological insulator (Chern insulator) if the Fermi level crosses only edge states, or a metal, if the Fermi level lies in the continuum of bulk states.

A key feature of our linear response theory is that Fermi surface contributions to the density accumulations are clearly separated from "undergap" contributions which arise from occupied states below the Fermi energy. This separation allows us to answer the longstanding question of whether undergap currents contribute to edge density accumulations. In TR invariant systems, we find that undergap currents do not contribute to density accumulations. The latter can only occur in metallic or edge-metallic states and are inevitably accompanied by dissipation. We refer to this result, first derived in Ref.[32], as the *"no-dissipation no-accumulation theorem"*.

Notice that the case of quantum spin Hall insulators requires a more careful discussion. These systems are TR invariant, yet they support spin-density accumulations and persistent edge spin currents due to spin-momentum locking. Both effects arise from the presence of edge states crossing the Fermi level in agreement with the thesis of our theorem. However, our theorem also implies that such edge currents must be dissipative: indeed, while they are protected against elastic scattering from non-magnetic impurities, they are not protected against more general scattering processes and therefore are qualitatively different from undergap currents such as the ones that flow in the anomalous Hall effect (see further discussion in Sec. V).

In systems with broken TR symmetry, non-dissipative edge density accumulations can occur even in topologically trivial states, i.e., in the absence of edge states crossing the Fermi level. An example of this is the bulk magnetoelectric effect, which will be discussed in the concluding section [33–35].

Our theoretical analysis does not explicitly cover some extrinsic effects, such as skew-scattering and side jump, in which impurity scattering acts as a *source* of current rather than a limiting factor. These effects, however, are implicitly contained in the exact eigenstates formulation of Sections II and III, and the only technical problem is to perform disorder averages, which is usually done with the help of diagrammatic techniques [36–39]. It should also be noted that these extrinsic effects occur in metals and are necessarily dissipative, so their inclusion will not change qualitative conclusions such as the no-dissipation-no-accumulation theorem for insulators.

The remainder of this paper is structured as follows. Section II introduces the framework for studying generalized densities and current densities. We focus on electrically driven nonconserved density accumulations. The response of the generalized density and current is derived in Sec. III, and the no-dissipation no-accumulation theorem is proved making use of the microscopic reciprocity theorem of Appendix I. In Sec. IV we present the generalized continuity equation with a torque density term and discuss the potential advantages of introducing a "proper" current density that satisfies the macroscopic Onsager reciprocity relations [40, 41]. Using a torque dipole density, the conventional continuity equation is recovered in Sec. V. In Sec. VI we present the calculation of the OMM density response in model nanoribbon systems with TR symmetry (Sec. VI A) and without TR symmetry (Sec. VI B) – the latter case including both trivial and nontrivial topological phases of the Haldane model. We end with a summary and a brief discussion of the anomalous Hall effect in Sect. VII.

II. BASIC DEFINITIONS

A. Generalized densities, currents and torques

Consider an extensive physical quantity (the "charge") represented by an operator of the form

$$\hat{O} = \sum_i \hat{O}_i, \tag{1}$$

where the index i runs over the particles and $\hat{O}_i = \hat{O}\left(\hat{\mathbf{r}}_i, \hat{\mathbf{p}}_i, \hat{\sigma}_i \cdots\right)$ is a function of the position, momentum, spin... of the i-th particle. This can also be written in the second-quantized form:

$$\hat{O} = \sum_{\alpha\beta} \langle\alpha|\hat{O}|\beta\rangle\, \hat{c}^\dagger_\alpha \hat{c}_\beta, \tag{2}$$

where $|\alpha\rangle$ and $|\beta\rangle$ represent exact one-electron eigenstates. Notice that these are not necessarily Bloch states. Typical examples are the electric charge operator ($\hat{O}_i = -e\hat{1}_i$), the spin operator ($\hat{O}_i = \hbar\hat{\boldsymbol{\sigma}}_i/2$) and, our focus in this paper, the

OMM operator [42–44] $\hat{O}_i = (-e/4)\,(\hat{\mathbf{r}}_i \times \hat{\mathbf{v}}_i - \hat{\mathbf{v}}_i \times \hat{\mathbf{r}}_i)$ where $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{v}}_i$ are the position and velocity operators respectively. Here, the velocity operator is defined as the time derivative of the position operator: $\hat{\mathbf{v}}_i = \partial_t \hat{\mathbf{r}}_i = i\hbar^{-1}[\hat{H}_E, \hat{\mathbf{r}}_i]$, where $\hat{H}_E$ is the Hamiltonian of the system in the presence of the external field[1]:

$$\hat{H}_E = \hat{H}_0 + e \sum_i \mathbf{E} \cdot \hat{\mathbf{r}}_i . \tag{3}$$

Only the unperturbed Hamiltonian, $\hat{H}_0$, contributes to the commutator, so we can simply write

$$\hat{\mathbf{v}}_i = i\hbar^{-1}[\hat{H}_0, \hat{\mathbf{r}}_i] . \tag{4}$$

To the operators $\hat{O}_i$, we associate a one-body density operator $\hat{n}_O(\mathbf{r})$ defined as follows

$$\hat{n}_O(\mathbf{r}) = \sum_i \hat{O}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i), \tag{5}$$

where the $\star$ is the symmetrized (Hermitian) product $\hat{A} \star \hat{B} = (\hat{A}\hat{B} + \hat{B}\hat{A})/2 = \{\hat{A}, \hat{B}\}/2$. This is the observable whose expectation value determines the "density accumulation". Another important observable is the "density accumulation rate", i.e., the operator associated with the time derivative of the density, which is obtained by taking the commutator of the density with the Hamiltonian $\hat{H}_E$:

$$\partial_t \hat{n}_O(\mathbf{r}) = \sum_i (\partial_t \hat{O}_i) \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) - \nabla_{\mathbf{r}} \cdot \sum_i \hat{O}_i \star \hat{\mathbf{v}}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) , \tag{6}$$

where $\partial_t \hat{O}_i = i\hbar^{-1}[\hat{H}_E, \hat{O}_i]$. In writing this equation, we have used the fact that the single-particle density operator obeys the equation of motion

$$\partial_t \delta(\mathbf{r} - \hat{\mathbf{r}}_i) = -\nabla_{\mathbf{r}} \cdot \hat{\mathbf{v}}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) . \tag{7}$$

This suggests that we define the current density operator

$$\hat{\mathbf{J}}_O(\mathbf{r}) = \sum_i \hat{O}_i \star \hat{\mathbf{v}}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \tag{8}$$

and the torque density operator

$$\hat{\mathcal{T}}_O(\mathbf{r}) = \sum_i \hat{\mathcal{T}}_{O,i} \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) , \tag{9}$$

where

$$\hat{\mathcal{T}}_{O,i} \equiv i\hbar^{-1}[\hat{H}_E, \hat{O}_i] \tag{10}$$

is the torque acting on the moment of the i-th particle. Then, the expression for the density accumulation rate can be recast as a generalized continuity equation

$$\partial_t \hat{n}_O(\mathbf{r}) + \nabla_{\mathbf{r}} \cdot \hat{\mathbf{J}}_O(\mathbf{r}) = \hat{\mathcal{T}}_O(\mathbf{r}) . \tag{11}$$

The divergence of the current on the left-hand side of this equation describes the flow of "charge" that is physically transported in or out of an infinitesimal volume centered at $\mathbf{r}$, while the torque on the right-hand side describes the non-conservation of the "charge" within that volume. The latter vanishes when the "charge" is conserved: this is, of course, the case for the electric charge. But even in the familiar case of the spin density, the presence of spin-orbit interactions gives a non-zero contribution to the right-hand side of Eq. (11). In the case of the orbital moment, the situation is much more complex, since the electric field itself breaks the conservation of the orbital moment. In general, the total torque is the sum of two contributions: one, $i\hbar^{-1}[\hat{H}_0, \hat{O}]$, arising from the non-conservation of the charge density, due to internal interactions with lattice and impurities and one, $i\hbar^{-1}[e\mathbf{E} \cdot \hat{\mathbf{r}}, \hat{O}]$, arising from the action of the externally applied electric field.

B. Proper current and macroscopic Onsager reciprocity

It is natural at this point to ask whether the torque term on the right-hand side of Eq. (11) can be absorbed in a redefinition of the current density so that the standard continuity equation (with zero on the right-hand side) holds. This is the approach that was taken in Ref. [41] and recently in Ref. [45] to arrive at a "proper" definition of the spin current. The idea was to express the torque that appears on the right-hand side of Eq. (11) as the negative of the divergence of a torque dipole density

$$\hat{\mathcal{T}}_O(\mathbf{r}) = -\nabla_{\mathbf{r}} \cdot \sum_i \hat{\mathbf{r}}_i \star \hat{\mathcal{T}}_{O,i} \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) . \tag{12}$$

Moving this to the left-hand side of Eq. (11) and combining it with the divergence of the current led to a continuity equation of the standard form, i.e,

$$\partial_t \hat{n}_O + \nabla_{\mathbf{r}} \cdot \hat{\mathcal{J}}_O = 0 \tag{13}$$

with a "proper" current density

$$\hat{\mathcal{J}}_O = \sum_i \partial_t (\hat{O}_i \star \hat{\mathbf{r}}_i) \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) . \tag{14}$$

An appealing feature of this definition is that the coupling of the proper current to a spatially uniform but time-dependent vector potential $\mathbf{A}(t)$ takes the form

$$-\sum_i \partial_t (\hat{O}_i \star \hat{\mathbf{r}}_i) \cdot \mathbf{A}(t) , \tag{15}$$

as can be seen starting from a coupling of the form $-\hat{\mathbf{P}}_O \cdot \mathbf{E}(t)$, where $\hat{\mathbf{P}}_O \equiv \sum_i \hat{O}_i \star \hat{\mathbf{r}}_i$ is the macroscopic polarization associated with the charge $\hat{O}$ and $\mathbf{E}(t) = -\partial_t \mathbf{A}(t)$ is the electric field. The expression (15) is obtained by performing a gauge transformation of the form $\exp\left[i\hat{\mathbf{P}}_O \cdot \mathbf{A}(t)\right]$. Crucially, this allows one to establish a macroscopic (Onsager) reciprocity relation between the response of $\hat{\mathcal{J}}_O$ to an electric field that couples to the electric polarization $\hat{\mathbf{P}}$ and the response of the standard electric

---

[1] Notice that the time derivative of an operator is constructed within the Heisenberg picture of time evolution. After doing this, we return to the Schrödinger picture.

current to an electric field that couples to the generalized polarization $\hat{\mathbf{P}}_O$. We call this reciprocity "macroscopic" because it applies to the response of currents to spatially uniform electric field, whereas the microscopic reciprocity relation described in Section III and in Appendix I applies to all linear response functions.

The problem with this approach is that the representation of the torque density as the divergence of a torque dipole density is possible only when the net torque density, integrated over space, is zero. This is not generally the case [45, 46]. Under the action of an electric field, a net spin density or a net orbital moment may be generated. In particular, density accumulations on opposite boundaries of the finite system may have the same sign [47]. When this happens, thinking of the orbital moment as a quantity that is simply transported from one edge to another is no longer possible. Furthermore, even when the spatial average of the torque is zero, Eq. (12) is valid only in the limit of slowly varying density, as will be shown in detail in Appendix V.

Nevertheless we will see in Section IV that it is still possible and useful (at least in the limit of slow spatial variation of the density) to absorb the torque dipole density in a "proper" definition of the current, along the lines of Ref. [41], while still keeping the *net* torque on the right-hand side of the continuity equation. We will see that the proper current (unlike the conventional current) vanishes identically in a fully gapped time-reversal invariant insulator.

III. LINEAR RESPONSE THEORY FOR GENERALIZED DENSITIES

We start from the well-known formula [48] for the linear response of the expectation value of a single particle Hermitian operator $\hat{A}$ to a periodic external field F of frequency ω, which couples linearly to a single-particle Hermitian operator $\hat{B}$:

$$\delta A(\omega) = \chi_{A,B}(\omega) F(\omega)\,, \tag{16}$$

where the response function $\chi_{A,B}(\omega)$ is given by

$$\chi_{A,B}(\omega) = \sum_{\alpha\beta} \mathcal{L}^{\eta}_{\alpha\beta}(\omega) [\hat{A}]_{\alpha\beta} [\hat{B}]_{\beta\alpha}\,, \tag{17}$$

and

$$\mathcal{L}^{\eta}_{\alpha\beta}(\omega) \equiv \frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta + \omega + i\eta} \tag{18}$$

is the Lindhard factor, $f_\alpha = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)}+1}$ is the Fermi-Dirac average occupation of state-α at temperature T. We have introduced a compact notation for the matrix elements of any Hermitian operator $\hat{A}$:

$$[\hat{A}]_{\alpha\beta} \equiv \langle\alpha|\hat{A}|\beta\rangle = [\hat{A}]^*_{\beta\alpha}\,. \tag{19}$$

Equations (17) and (18) are valid for non-interacting systems, but our analysis is general and remains valid in fully interacting systems, as shown in Appendix I.

The density response is obtained by setting $\hat{A} = \hat{n}_O(\mathbf{r})$, $\hat{B} = \hat{\mathbf{r}}$, and $F(t) = e\mathbf{E}(t)$. Thus we have

$$\delta n_O(\mathbf{r},\omega) = e\chi_{n_O,\mathbf{r}}(\mathbf{r},\omega)\cdot\mathbf{E}(\omega)\,, \tag{20}$$

where

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},\omega) = \sum_{\alpha\beta} \mathcal{L}^{\eta}_{\alpha\beta}(\omega)[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\,. \tag{21}$$

A. Time reversal invariant systems: the no-accumulation-no-dissipation theorem

Let us assume that the system is TR invariant. This means that for each eigenstate $|\alpha\rangle$ there is a time-reversed partner $|\tilde{\alpha}\rangle$ with the same energy, so that the sum over α and β in Eq. (21) can be replaced by a sum over $\tilde{\alpha}$ and $\tilde{\beta}$. It is shown in Appendix I that TR invariance implies the *microscopic Onsager reciprocity relation*

$$\chi_{A,B}(\omega) = \chi_{\tilde{B},\tilde{A}}(\omega)\,, \tag{22}$$

where $\tilde{A} = \Theta\hat{A}^\dagger\Theta^{-1}$ is the transformation of $\hat{A}$ under TR. We consider Hermitian operators of definite parity under TR, i.e., such that $\tilde{A} = \lambda_A\hat{A}$ and $\tilde{B} = \lambda_B\hat{B}$ where $\lambda = +1$ for a TR-even operator such as $\hat{\mathbf{r}}$, and $\lambda = -1$ for a TR-odd operator such as the orbital moment density or the valley density. Then Eq. (22) tells us that

$$\begin{aligned}\chi_{A,B}(\omega) &= \lambda_A\lambda_B\chi_{B,A}(\omega)\\ &= \frac{1}{2}\left[\chi_{A,B}(\omega) + \lambda_A\lambda_B\chi_{B,A}(\omega)\right].\end{aligned} \tag{23}$$

Making use of Eq. (17) we readily find (see Appendix I)

$$\begin{aligned}\chi_{AB}(\omega) = {}& \frac{1+\lambda_A\lambda_B}{2}\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\,\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}]\\ &+ i\frac{1-\lambda_A\lambda_B}{2}\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\,\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}]\,.\end{aligned} \tag{24}$$

If the operators $\hat{A}$ and $\hat{B}$ have opposite parities under time reversal, as is the case when $\hat{n}_O(\mathbf{r})$ is the orbital moment density or the valley density or the spin density, then $\lambda_A\lambda_B = -1$ and the formula (21) for the linear response simplifies to

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},\omega) = \sum_{\alpha\beta} i\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\,\mathrm{Im}\{[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\}\,. \tag{25}$$

We separate the real and the imaginary parts of $\mathcal{L}^{\eta}_{\alpha\beta}(\omega)$ as follows

$$\mathcal{L}^{\eta}_{\alpha\beta}(\omega) = P\frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta + \omega} - i\pi(f_\alpha - f_\beta)\delta(\epsilon_\alpha - \epsilon_\beta + \omega)\,, \tag{26}$$

where P denotes the principal part: $P\frac{a}{x} \equiv \lim_{\eta\to 0}\frac{ax}{x^2+\eta^2}$. In the limit of zero frequency ($\omega \to 0$) the first term reduces to $P(f_\alpha - f_\beta)/(\epsilon_\alpha - \epsilon_\beta)$, which is symmetric under the interchange of α and β. At the same time, $\mathrm{Im}\{[\hat{n}_O]_{\alpha\beta}(\mathbf{r})[\hat{\mathbf{r}}]_{\beta\alpha}\}$ is antisymmetric under the interchange of α and β – a fact that follows immediately from the hermiticity relations (19). Therefore, the first term of Eq. (26) gives a vanishing contribution when summed over α and β. This leaves us with the simple result

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},0) = \lim_{\omega\to 0} \pi \sum_{\alpha\beta} \mathrm{Im}\{[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\}(f_\alpha - f_\beta)\delta(\epsilon_\alpha - \epsilon_\beta + \omega)\,. \tag{27}$$

It is evident that the response vanishes at zero temperature, unless there are pairs of states on opposite sides of the Fermi level ($f_\alpha - f_\beta \neq 0$), separated by arbitrarily small excitation energy ($\epsilon_\beta - \epsilon_\alpha = \omega \to 0$). The existence of these pairs of states is also a requirement for the occurrence of dissipation in a static electric field because these are the only pairs of states that can absorb energy from such a field. We conclude that in a TR-invariant system the accumulation of a TR-odd density must necessarily vanish if there are no states at the Fermi level that can absorb energy from the electric field. We have previously referred to this result as the "no dissipation-no accumulation theorem" [32]. For a more mathematical discussion of this result, we refer the reader to Appendix II. There we show that in the case of a disordered metal in the diffusive regime, under the relaxation time approximation, Eq. (27) can be cast in the form

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},0) = \sum_i N_i(\epsilon_F)\tau_i \left\langle \mathrm{Re}\{[\hat{n}_O(\mathbf{r})]_{\mathbf{k},\mathbf{k}}[\hat{\mathbf{v}}]_{\mathbf{k},\mathbf{k}}\}\right\rangle_{FS,i}\,, \tag{28}$$

where the sum runs over the sheets of the Fermi surface (denoted by subscripts FS, i), the angular brackets denote an average over the i-th sheet of the Fermi surface, $N_i(\epsilon_F)$ and τ_i are, respectively, the partial density of states and the momentum relaxation time on the i-th sheet of the Fermi surface. As an illustration of this formula, the calculation of the current-induced spin polarization [49–52] in a disordered Rashba metal is presented in Appendix II.

By contrast, consider the response of the generalized current $\mathbf{J}_O(\mathbf{r})$ to the electric field. The current associated with a TR-odd density is TR-even. Making use of Eq. (24) we then find (in the zero-frequency limit)

$$\chi_{J_O,\mathbf{r}}(\mathbf{r},0) = 2P\sum_{\alpha\beta} f_\alpha \frac{\mathrm{Re}\{[\hat{\mathbf{J}}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\}}{\epsilon_\alpha - \epsilon_\beta}\,, \tag{29}$$

where

$$[\hat{\mathbf{J}}_O(\mathbf{r})]_{\alpha\beta} = \sum_i \langle\alpha|\hat{O}_i \star (\partial_t\hat{\mathbf{r}}_i) \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i)|\beta\rangle \tag{30}$$

denotes the matrix element of the generalized current, The current response involves contributions from all occupied states, so it can differ from zero even in an insulator: this is known as an "undergap current". In Appendix III we show that Eq. (29) is equivalent to the familiar Berry curvature formulas for anomalous conductivities such as the orbital Hall conductivity.

B. Broken time-reversal invariance

Let us now discuss what happens in systems that are not TR invariant. Let us go back to the original expression (17) of the response function of a generalized density at zero frequency, which reads:

$$\begin{aligned}\chi_{n_O,\mathbf{r}}(\mathbf{r},0) &= \sum_{\alpha\beta}\mathcal{L}^{\eta}_{\alpha\beta}(0)[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\\ &= \sum_{\alpha\beta} P\frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta}[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\\ &-i\pi\sum_{\alpha\beta}(f_\alpha - f_\beta)\delta(\epsilon_\alpha - \epsilon_\beta)[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}.\end{aligned} \tag{31}$$

Broken TR symmetry allows the first term on the right-hand side to be nonzero even when the second term vanishes. Assuming that this is the case, we arrive at the expression

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},0) = 2P\sum_{\alpha\beta} f_\alpha\frac{\mathrm{Re}\left\{[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\right\}}{\epsilon_\alpha - \epsilon_\beta}\,. \tag{32}$$

Notice the formal similarity with Eq. (29). A fully gapped insulator with broken TR symmetry can support the accumulation of a TR-odd density, just as a TR-invariant insulator can support the *current* associated with a TR-odd density. We will see concrete examples of this in the following sections.

IV. GENERALIZED CONTINUITY EQUATION AND TORQUE

We now consider the generalized continuity equation (11) from the point of view of linear response theory. We start from the well-known identity[48]

$$-i\omega\chi_{A,B}(\omega) = -i\langle[\hat{A},\hat{B}]\rangle + \chi_{\dot{A},B}(\omega) \tag{33}$$

where the angular bracket denotes the average in the unperturbed ground state and $\dot{A} = \partial_t\hat{A} = i\hbar^{-1}[\hat{H}_0,\hat{A}]$. Applying this to our density response function, we get

$$-i\omega\chi_{n_O,\mathbf{r}}(\mathbf{r},\omega) = -i\langle[\hat{n}_O(\mathbf{r}),\hat{\mathbf{r}}]\rangle + \chi_{\dot{n}_O,\mathbf{r}}(\mathbf{r},\omega)\,. \tag{34}$$

After multiplication of this equation by $e\mathbf{E}$, on the left-hand side, we have the Fourier transform of the time derivative of the generalized density. Similarly, the first term on the right-hand side is easily recognized to be the density of "torque" exerted by the electric field, i.e., what we have called "extrinsic torque density" after Eq. (11). Because this torque is of the first order in $\mathbf{E}$, the average is taken over the equilibrium state (i.e., the state at zeroth order in $\mathbf{E}$).

The second term on the right-hand side can further be decomposed into two parts according to Eq. (6):

$$\chi_{\dot{n}_O,\mathbf{r}}(\mathbf{r},\omega) = -\nabla_{\mathbf{r}}\cdot\chi_{j_O,\mathbf{r}}(\mathbf{r},\omega) + \chi_{T^{int}_O,\mathbf{r}}(\mathbf{r},\omega), \tag{35}$$

where $\chi_{T^{int}_O,\mathbf{r}}(\mathbf{r},\omega)$ is the response function of the "intrinsic" torque density operator

$$\hat{T}^{int}_O(\mathbf{r}) = \sum_i[\partial_t\hat{O}_i]\star\delta(\mathbf{r} - \hat{\mathbf{r}}_i) \tag{36}$$

to the external field. Here $\partial_t\hat{O}_i = i\hbar^{-1}[\hat{H}_0,\hat{O}_i]$. We have thus recovered the structure of the generalized continuity equation announced in Sec. II, Eq. (11).

But now we can go further and obtain an expression for the torque by computing the linear response function $\chi_{T_O^{int},\mathbf{r}}$. The total torque term is given by

$$T_O(\mathbf{r},\omega) = e\left[-i\langle[\hat{n}_O(\mathbf{r}),\hat{\mathbf{r}}]\rangle + \chi_{T_O^{int},\mathbf{r}}(\mathbf{r},\omega)\right]\cdot\mathbf{E}. \qquad (37)$$

Let us focus on the *net torque* $\bar{T}_O \equiv \lim_{\omega\to0}\int T_O(\mathbf{r},\omega)d\mathbf{r}$. A non-zero value of this quantity implies that the total "charge" $\hat{O} \equiv \sum_i \hat{O}_i = \int \hat{n}_O(\mathbf{r})d\mathbf{r}$ is not conserved. The net torque associated to a TR-odd charge is even under time reversal: therefore utilizing once again Eq. (24) we find that its linear response is given by

$$\lim_{\omega\to0}\bar{\chi}_{T_O^{int},\mathbf{r}}(\omega) = P\sum_{\alpha\beta}\frac{f_\alpha - f_\beta}{\epsilon_\alpha-\epsilon_\beta}[\partial_t\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\,. \qquad (38)$$

Making use of the identity $[\partial_t\hat{O}]_{\alpha\beta} = i\hbar^{-1}(\epsilon_\alpha-\epsilon_\beta)[\hat{O}]_{\alpha\beta}$ and the definition of the principal part this can be rewritten as

$$\lim_{\omega\to0}\bar{\chi}_{T_O^{int},\mathbf{r}}(\omega) = -\frac{1}{i\hbar}\sum_{\alpha\beta}\frac{(\epsilon_\alpha-\epsilon_\beta)^2(f_\alpha-f_\beta)}{(\epsilon_\alpha-\epsilon_\beta)^2+\eta^2}[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\,. \qquad (39)$$

Eq. (39) is "almost" the negative of the first term in the square bracket of Eq. (37), which, after spatial integration, can be expressed as

$$\frac{1}{i\hbar}\langle[\hat{O},\hat{\mathbf{r}}]\rangle = \frac{1}{i\hbar}\sum_{\alpha\beta}(f_\alpha-f_\beta)[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\,. \qquad (40)$$

The difference between Eq. (39) and Eq. (40) arises from the seemingly innocent factor principal part factor, $\frac{(\epsilon_\alpha-\epsilon_\beta)^2}{(\epsilon_\alpha-\epsilon_\beta)^2+\eta^2}$, which excludes from the summation terms with $\epsilon_\alpha \simeq \epsilon_\beta$. Combining Eqs. (39) and (40) and using the identity $\frac{\eta}{(\epsilon_\alpha-\epsilon_\beta)^2+\eta^2} = \pi\delta(\epsilon_\alpha-\epsilon_\beta)$ (valid in the $\eta\longrightarrow0$ limit) we obtain

$$\bar{T}_O \;=\; \frac{e\pi\eta}{\hbar}\sum_{\alpha\beta}(f_\alpha-f_\beta)\delta(\epsilon_\alpha-\epsilon_\beta)\mathrm{Im}\left\{[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\right\}\cdot\mathbf{E} \qquad (41)$$

Observe that this expression coincides, apart from the factor η, with the expression for the spatially integrated density accumulation, which is given by Eq. (27). We can immediately conclude that the net torque is a Fermi surface property,just as the density accumulation. We process Eq. (41) following the same steps that led us to Eq. (28) for the density accumulation. In other words, replace $f_\alpha - f_\beta$ by $f'(\epsilon_\alpha)\delta(\epsilon_\alpha-\epsilon_\beta)$ and use $i(\epsilon_\beta-\epsilon_\alpha)[\hat{\mathbf{r}}]_{\beta\alpha} = [\hat{\mathbf{v}}]_{\beta\alpha}$. Then we replace the $\delta(\epsilon_\alpha-\epsilon_\beta)$ by $\frac{(\pi\tau)^{-1}}{(\epsilon_\alpha-\epsilon_\beta)^2+\tau^{-2}}$, where τ^{-1} is the spectral width of disorder-broadened Bloch states (this is the relaxation time approximation). The result is

$$\bar{T}_O = \eta\left(eN(\epsilon_F)\tau\left\langle\mathrm{Re}\{[\hat{O}]_{\mathbf{k},\mathbf{k}}[\hat{\mathbf{v}}]_{\mathbf{k},\mathbf{k}}\}\right\rangle_{FS}\cdot\mathbf{E}\right)\,, \qquad (42)$$

where we have assumed, for simplicity, that the Fermi surface is a single sheet with a single relaxation time. Thus, under the stated assumptions (i.e, in the relaxation time approximation) the net volume torque $\bar{T}_O$ in the zero-frequency limit is the integrated density accumulation Eq. (28) multiplied by η, or, equivalently, divided by the adiabatic switching-on time η^{-1}.

It is perhaps not surprising that Eq. (42) goes to zero in the limit $\eta\longrightarrow0$. The reason is that the integrated "torque" is the time derivative of the total "charge" and therefore is expected to vanish in the steady state, provided that (i) a steady state exists, and (ii) the total "charge" is described by a bounded operator. Both conditions are satisfied here, and Eq. (42) is consistent with the findings of Ref. [45].

However, the most interesting result here is not the steady-state torque (which vanishes), but the *external torque*, which, as discussed in Ref. ([32]) is the source of the density accumulation, acting for a time of order τ *before* the steady state is reached. The external torque, which we call $\bar{T}_O^{ext}$, can be extracted from Eq. (42) as

$$\bar{T}_O^{ext} = \frac{\bar{T}_O}{\eta\tau} = eN(\epsilon_F)\left\langle\mathrm{Re}\{[\hat{O}]_{\mathbf{k},\mathbf{k}}[\hat{\mathbf{v}}]_{\mathbf{k},\mathbf{k}}\}\right\rangle_{FS}\cdot\mathbf{E} \qquad (43)$$

i.e., it is the torque that we would compute from Eq. (37) if we chose the relaxation time τ to go to infinity while η goes to zero in such a way that $\eta\tau = 1$. In phenomenological theories of the density accumulation $\bar{T}_{O,ext}$ can be used as the torque that is generated by the external electric field acting on Bloch states before impurities have had a chance to act. Notice that Eq. (43) does not depend on the phenomenological parameter τ and can be evaluated from a purely microscopic theory. The additional torque exerted by impurities or other mechanisms can then be included phenomenologically, for example in the relaxation time approximation.

In Appendix IV we show that $\bar{T}_O$ can be formally expressed as a Fermi-surface average of the commutator between *projected* operators $\hat{O}_{FS}$ and $\hat{\mathbf{r}}_{FS}$, these being defined as the restrictions of $\hat{O}$ and $\hat{\mathbf{r}}$ respectively to the subspace of degenerate states at the Fermi level:

$$\bar{T}_O = \frac{ie}{\hbar}\langle\mathcal{D}\hat{O}\rangle_{FS}\cdot\mathbf{E}\,, \qquad (44)$$

where $\mathcal{D}\hat{O}$ is a short-hand for the commutator $i[\hat{O}_{FS},\hat{\mathbf{r}}_{FS}]$. The notation $\mathcal{D}$ is, of course, suggestive of the covariant derivative with respect to $\mathbf{k}$ to which the commutator between projected operators reduces when the eigenstates of the Hamiltonian are Bloch waves (see Appendix IV).

V. RECOVERING A CONVENTIONAL CONTINUITY EQUATION

We have so far considered only the net torque, i.e., the spatial average of the torque density. The difference between the actual torque density and its average will be referred to as "distributed torque". The distributed torque integrates to zero and, therefore, can be expressed as the spatial divergence of a vector field. In the limit of slow spatial variation, this vector field is the negative of the *torque dipole density*

$$D_{O,a}(\mathbf{r}) = \sum_i \langle \hat{r}_{i,a} \star \hat{T}_{O,i} \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \rangle \,, \tag{45}$$

where a is a cartesian index (for a proof of this see Appendix V). Thus we have

$$T_O(\mathbf{r}) - \frac{1}{V}\bar{T}_O = -\sum_a \nabla_a D_{O,a}(\mathbf{r}), \tag{46}$$

where V is the volume (or the area) of the system. As described in Sec. II B the distributed torque can be absorbed in a "proper" current density

$$\hat{\mathcal{J}}_{O,a}(\mathbf{r}) = \hat{j}_{O,a}(\mathbf{r}) + \hat{D}_{O,a}(\mathbf{r}) \,, \tag{47}$$

and the generalized continuity equation takes the form

$$\partial_t n_O(\mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \mathcal{J}_O(\mathbf{r}) = \frac{1}{V}\bar{T}_O \,. \tag{48}$$

Notice the appearance of the spatially integrated (net) torque on the right-hand side. In practical applications $\bar{T}_O$ would have to be approximated, for example by combining the microscopic expression (43) for the external contribution with a phenomenological expression for the internally generated relaxation.

Unlike the conventional current, the "proper current" is a Fermi surface property. To see this, consider its response to an external electric field. Just as we did previously in Eqs. (38) and (40) we can write the linear response of the proper current as the sum of two terms: an intrinsic one

$$\lim_{\omega\to 0} \bar{\chi}_{\mathcal{J}^{int}_{O,a},\mathbf{r}}(\omega) = P \sum_{\alpha\beta} \frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta} [\partial_t(\hat{r}_a \star \hat{O})]_{\alpha\beta} [\hat{\mathbf{r}}]_{\beta\alpha} \,, \tag{49}$$

where $\hat{r}_a \star \hat{O}$ is a short-hand for $\sum_i \hat{r}_{i,a} \star \hat{O}_i$, and an extrinsic one

$$\frac{1}{i\hbar}\langle [\hat{r}_a \star \hat{O}, \mathbf{r}] \rangle = \frac{1}{i\hbar} \sum_{\alpha\beta} (f_\alpha - f_\beta) [\hat{r}_a \star \hat{O}]_{\alpha\beta} [\hat{\mathbf{r}}]_{\beta\alpha} \,. \tag{50}$$

Crucially Eq. (49) involves the matrix elements of a time derivative (with dynamics controlled by H_0) so it can be treated exactly in the same way as we treated the net torque in the early part of this section and we arrive at the Fermi surface formula

$$\bar{\mathcal{J}}_{O,a} = \frac{ie}{\hbar} \left\langle \mathcal{D}\left(\hat{r}_a \star \hat{O}\right) \right\rangle_{FS} \cdot \mathbf{E} \,, \tag{51}$$

where $\mathcal{D}\left(\hat{r}_a \star \hat{O}\right)$ is a short-hand for the commutator $i[\left(\hat{r}_a \star \hat{O}\right)_{FS}, \hat{\mathbf{r}}_{FS}]$. The torque and the current calculated in this section are valid in the limit of slowly varying density. Formally exact, but necessarily more complicated formulas for the torque density are reported in the Appendix IV.

VI. CALCULATIONS OF OMM DENSITY RESPONSE

After presenting the general theory of density accumulations, we now turn to its practical application in the context of orbital magnetic responses [53–55]. Specifically, we focus on the accumulation of orbital magnetic moments in Dirac materials (graphene, hBN, TMDCs, etc.), both with and without TR symmetry [8, 56]. The OMM density operator is defined as

$$\hat{\mathbf{n}}_m(\mathbf{r}) = \sum_i \hat{\mathbf{m}}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \tag{52}$$

where $\hat{\mathbf{m}}_i = \frac{e}{4}(\hat{\mathbf{r}}_i \times \hat{\mathbf{v}}_i - \hat{\mathbf{v}}_i \times \hat{\mathbf{r}}_i)$ is the OMM operator.[2]

As a concrete example, we consider the OMM response of the Haldane model with a staggered sublattice potential Δ [1]. The Hamiltonian for this system is given by (see Appendix VI):

$$H = -t_0 \sum_{\langle ij \rangle} \hat{a}_i^\dagger \hat{b}_j + t_2 e^{i\phi} \sum_{\langle\langle ij \rangle\rangle} \hat{a}_i^\dagger \hat{a}_j + t_2 e^{-i\phi} \sum_{\langle\langle ij \rangle\rangle} \hat{b}_i^\dagger \hat{b}_j + h.c. \\ + \Delta \sum_i (\hat{a}_i^\dagger \hat{a}_i - \hat{b}_i^\dagger \hat{b}_i), \tag{53}$$

where $\hat{a}$ and $\hat{b}$ ($\hat{a}^\dagger$ and $\hat{b}^\dagger$) are the electron annihilation (creation) operators for the A and B sublattices of the honeycomb lattice. The parameter t_0 denotes the nearest-neighbor hopping amplitude between the A and B sublattices, while t_2 represents the next-nearest-neighbor hopping amplitude within the A (or B) sublattice. The phase parameter ϕ is a measure of the staggered magnetic flux, breaking the TR symmetry.

For simplicity, we set $\phi = \pi/2$. The model can be solved analytically, yielding two symmetrically placed bands with energies:

$$\epsilon_{1,2}(\mathbf{k}) = \pm\sqrt{t_0^2 |\gamma_{\mathbf{k}}|^2 + (\Delta - 2t_2\beta_{\mathbf{k}})^2}, \tag{54}$$

where $\gamma_{\mathbf{k}} = \sum_n e^{i\mathbf{k}\cdot\boldsymbol{\delta}_n}$, $\beta_{\mathbf{k}} = \sum_n \sin(\mathbf{k} \cdot \mathbf{l}_n)$, and $2t_2\beta_{\mathbf{k}}$ is known as the Haldane mass. Here, $\boldsymbol{\delta}_n$ represents the nearest neighbor bonds and $\mathbf{l}_n$ are the next nearest neighbor bonds [57]. The nonzero value of t_2 paired with a non-vanishing magnetic flux ϕ (modulo 2π) breaks the TR symmetry, while Δ breaks the inversion symmetry. Thus, the model has sufficient flexibility to illustrate both the TR-invariant case ($t_2 = 0$) and the case of broken TR symmetry ($t_2 \neq 0$ and $\phi \neq 2n\pi$).

In the next two sections, we will consider the Haldane model in a ribbon geometry with terminations (edges) parallel to the x-axis, and we will use the formalism developed in the previous sections to calculate the OMM density accumulations.

---

[2] It should be borne in mind that the OMM density is *not* equivalent to the thermodynamic magnetization density [44]. While thermodynamic magnetization includes terms arising from the change of the quasiparticle density of states in response to a magnetic field, the OMM density arises solely from the kinetic angular momentum of the electrons. In studies of the orbital Hall effect and related transport phenomena, the experimentally measured quantity is the induced OMM density, not the thermodynamic magnetization [22].

First, in Section VI A we will consider the TR-invariant case ($t_2 = 0$) where the Haldane model reduces to "gapped graphene" – the gap arising from the broken inversion symmetry. The band structure of the ribbon is shown in Fig. 1(b), where the energy levels are plotted vs k parallel to the edge. In this case an OMM density is generated only when the Fermi level crosses the conduction or valence bands and arises entirely from dissipative intraband processes, in agreement with the "non-dissipation no-accumulation theorem".

Next, in Section VI B we consider the Haldane model with $t_2 > 0$, $\phi = \pi/2$, which breaks TR symmetry. This model supports two distinct topological phases, determined by the competition between Δ and $2t_2\beta_{\mathbf{k}}$.

When $\Delta - 2t_2\beta_{\mathbf{k}} > 0$, the model is in the topologically trivial phase with a Chern number of zero for each band [1] and no edge states crossing the gap. The band structure of the ribbon in the topologically trivial phase is shown in Fig. 1(c). The essential difference between this and the TR invariant case is that an OMM density accumulation is generated even when the Fermi level lies in the gap: this nondissipative response arises from interband mixing and can be described as a magnetoelectric effect allowed by broken TR symmetry.

When $\Delta - 2t_2\beta_{\mathbf{k}} < 0$, the Haldane model enters a topologically nontrivial phase characterized by the quantum anomalous Hall effect [58]. The band structure of the ribbon is shown in Fig. 1(d). Because of the edge states crossing the gap, we now have both intraband and interband contribution to the OMM density accumulation. As a result, the system behaves as an orbital Chern insulator [59], where anomalous nondissipative transport leads to simultaneous accumulations of both charge and OMM densities.

A. The OMM density accumulation with TR symmetry

In this subsection, we perform calculations for finite ribbon structures based on Haldane's model. These ribbon structures maintain translational symmetry along the longitudinal direction (x-axis), while open boundary conditions are applied in the transverse direction (y-axis). The band structures of the ribbons are shown in Figs. 1(b)-(d) for various cases of interest.

The general form of the OMM density response along the transverse direction is given by (see Appendix VII for details):

$$\delta n_m^z(y) = e\tau E_x \int dx \sum_{n,k} \frac{\partial f_{nk}}{\partial k} n_{nn}^m(x,y,k) + eE_x \int dx \sum_{nn'k} \frac{f_{nk} - f_{n'k}}{\epsilon_{nk} - \epsilon_{n'k}} n_{nn'}^m(x,y,k) A_{n'n}(k). \tag{55}$$

The right-hand side consists of two terms: the first is an intraband contribution, which is associated with dissipative processes at the Fermi surface [60, 61] and involves the momentum relaxation time τ. The second term comes from interband contributions, which are intrinsic to the system and do not involve τ. Both contributions are strongly influenced by the presence or absence of TR and inversion symmetry. For the intraband term, the breaking of inversion symmetry is essential, while the intrinsic contribution requires the breaking of both inversion and TR symmetries. Notably, in TR symmetric systems only the intraband term survives, consistent with the no-dissipation, no-accumulation theorem.

Let us go back to Eq. (55), where the diagonal matrix elements of OMM density are

$$n_{nn}^m(x,y,k) = \sum_l \mathrm{Re}\Big\{ m_{nl}(k) \psi_{lk}^\dagger(x,y) \psi_{nk}(x,y) \Big\}. \tag{56}$$

and the matrix element of the OMM operator is given by (see Appendix VIII):

$$m_{nl}(k) = -\frac{e}{2\hbar}\partial_k[\epsilon_{nk} + \epsilon_{lk}] r_{nl}^y(k) + \frac{e}{4i\hbar} \sum_{n'} (\epsilon_{lk} + \epsilon_{nk} - 2\epsilon_{n'k}) \left[A_{nn'}(k) r_{n'l}^y(k) - r_{nn'}^y(k) A_{n'l}(k) \right], \tag{57}$$

where $r_{nl}^y(k)$ is the matrix element of the y-component of the position operator. The diagonal term $r_{nn}^y(k)$ indicates the average position of a state $|nk\rangle$ along the y axis and provides a useful illustration of the positions of the edge states. Eq. (57) exhibits a hybrid feature, combining the Berry connection in k-space with position matrix elements in real space. This distinguishes it from the conventional expression of the OMM in infinite systems [42, 62].

The OMMs, the average positions of the states along the y axis, and the accumulations of the OMM density, as computed for the ribbon band structure of Fig. 1(b) (gapped graphene ribbon) are presented in Fig. 2. In Fig. 2(a) we show the distribution of $m_{nn}(k)$ exhibiting an antisymmetric pattern in k-space, a direct consequence of TR symmetry. The calculations were done for a staggered onsite potential of $\Delta = 0.5t_0$ and $t_2 = 0.06t_0$. In Fig. 2(b), we plot the centers of the ribbon states along the y axis, $r_{nn}^y(k)$, which are used in calculation of the OMM density in Eq. (57). With these inputs, we calculate and plot the OMM density accumulation in Fig. 2(c), where we separate, for clarity, the intraband contribution (blue) and the interband contribution (orange). With a Fermi level crossing the bulk states at $E_f = 1.5t_0$ the interband contribution is zero, as predicted by the no-dissipation no-accumulation theorem. The accumulation arises entirely from the dissipative intraband term and its value depends on τ. Further analysis, presented in Fig. 2(d), shows that the induced OMM density can be decomposed into symmetric (orange squares) and antisymmetric (green triangles) components, with respect to the center of the ribbon. The symmetric component can be described as a dissipative magnetoelectric response induced by the difference between the two terminations (see Fig. Fig.1(a)) – while the antisymmetric component originates from the orbital Hall effect [47].

B. The OMM density accumulation with broken TR symmetry

When TR symmetry is broken, the OMM density response exhibits a new feature, which is especially evident when the Fermi level lies within the bulk band gap. We focus, therefore, on this case, and set $E_f = 0$. Because there is no proper Fermi surface (except for the edge states that may cross the Fermi

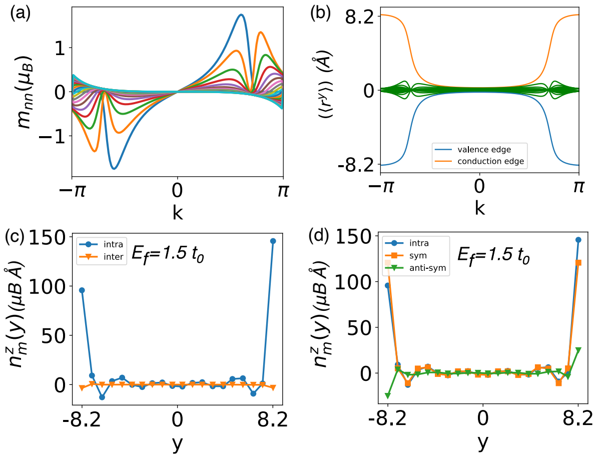


FIG. 2. Plots of calculated properties of TR-invariant "gapped graphene" nanoribbon with onsite staggered potential $\Delta = 0.5t_0$ and Fermi energy $E_f = 1.5t_0$ [see also Fig.1(b)]. (a) k dependence of the diagonal matrix element of the orbital magnetic moment $m_{nn}(k)$. Notice that $m_{nn}(k) = -m_{nn}(-k)$, as required by TR symmetry. (b) Plots of the centers of the ribbon states along the finite y direction ($r^y_{nn}(k)$). The ribbon, schematically shown in Fig.1(a)], has a width of 20 unit cells and the lattice constant is 1.42 Å. (c) The OMM density accumulation is resolved into an intraband contribution (blue dots) and an interband contribution (orange triangles). The interband contribution vanishes due to TR symmetry. (d) The intraband contribution to the OMM density accumulation is further resolved into components that are symmetric (orange squares) and antisymmetric (green triangles) with respect to the center of the ribbon ($y = 0$). The existence of the symmetric component is allowed by the fact that the edge terminations break mirror symmetry in a (z, x) plane passing through the center of the ribbon. (see Fig.1(a)).

level as will be discussed momentarily) the accumulation is dominated by the interband contribution:

$$\delta n^z_m(y) = eE_x \int dx \sum_{nn'k} \frac{f_{nk} - f_{n'k}}{\epsilon_{nk} - \epsilon_{n'k}} n^m_{nn'}(x, y, k) A_{n'n}(k). \tag{58}$$

Importantly, this contribution is nondissipative and intrinsic, in sharp contrast with the Fermi surface contributions that we found in gapped graphene.

We begin with the topologically trivial phase of the Haldane model, whose ribbon band structure is shown in Fig. 1(c). In this phase, there are no edge states connecting the conduction and the valence bands. Therefore, when the Fermi level is set at midgap (zero energy) there is strictly no Fermi surface, and the intraband contribution to the OMM density accumulation vanishes as shown by the blue dots in Fig. 3(b). However, the interband contribution is quite large and spatially asymmetric (due to the breaking of inversion symmetry), as shown by the orange triangles in Fig. 3(b). Crucially, the breaking of TR symmetry is essential for an orbital magnetization to appear. This can be understood by looking at the asymmetric distribution of the OMM in momentum space, which is illustrated in Fig. 3(a). The two valleys have OOMs of opposite signs but different magnitudes. In a TR invariant system, the magnitudes would be equal so that the OMM density generated in one valley (by virtual transitions between the conduction and valence bands) would be exactly canceled by an opposite contribution from the other valley. In the present system the cancellation fails, making non-dissipative orbital magnetization accumulation possible.

Further exploration of the OMM density accumulations in the topologically nontrivial phase of the Haldane model is illustrated in Fig. 3(c) and (d). The ribbon band structure is shown in Fig. 1(d). Two topologically protected edge states cross the Fermi level at midgap allowing a non-dissipative anomalous Hall current to flow in the bulk of the system. The precise mechanism is as follows: both electric charge and orbital magnetic moment (the latter shown by the blue dots in Fig. 3(d)) accumulate in the edge states. Notice that the accumulation is antisymmetric with respect to the center of the ribbon. The electric field generated by the charges accumulated on the edge drives the undergap Hall current. In addition, we observe an interband response of the OMM, shown by the orange triangles in Fig. 3(d), which is similar to the one observed in Fig. 3(b) but significantly larger. We attribute the larger size to the fact that the orbital moments, depicted in Fig. 3(c), have the same sign in the two valleys, leading to a much larger accumulation. This difference is evident when comparing Fig. 3(c) to Fig. 3(a) and is reflected in the larger magneto-electric response of the topologically nontrivial phase.

VII. CONCLUSION

We have systematically explored the nonconserved density accumulations arising from a steady electric field in several types of conducting and nonconducting materials. The formal work has been applied to a specific phenomenon–OMM density accumulations in the orbital Hall effect.

Firstly, we have found that in the presence of TR symmetry, the undergap contribution to density accumulation is absent. Consequently, density accumulations can only occur if there is a Fermi surface. This fact, first pointed out in Ref. [32] and aligning closely with recent experimental observations [22], can be seen more directly from the fact that the "proper current" defined in Section IV, absorbing the generalized torque dipole density, is a Fermi surface property. Examining gapped graphene ribbons with TR symmetry reveals the predominant influence of intraband contributions on the accumulation of OMM density. This contribution can be further dissected into a spatially uniform component (attributed to an orbital analogue of the Edelstein effect) and an antisymmetric component, which is linked to the orbital Hall effect.

When TR symmetry is disrupted, the no-dissipation no-accumulation is no longer valid. Nondissipative accumulations can arise, for example, from interband transitions in the Haldane model. We have illustrated the generation of a net OMM response by an electric field in a simplified Haldane model in which TR symmetry is broken. This phenomenon arises from states residing below the energy gap. It would be prohibited by the no-dissipation no-accumulation theorem in the TR symmetric scenario.

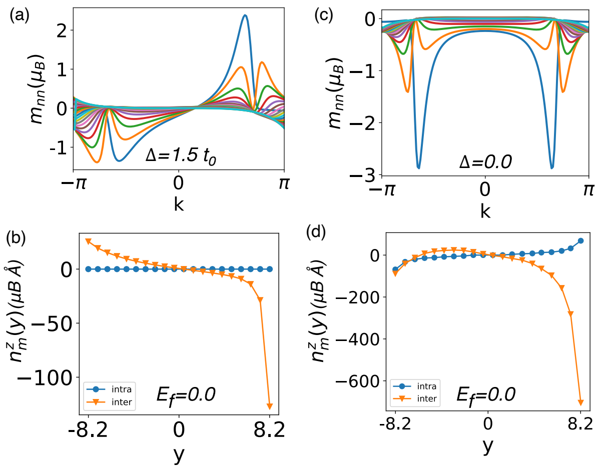


FIG. 3. OMM density response in Haldane model nanoribbons with flux $\phi = \frac{\pi}{2}$, and Fermi energy $E_f = 0$ at the center of the gap (a) Momentum space distribution of the orbital moment in the topologically trivial phase at on-site potential $\Delta = 0.5t_0$. Notice that $m_{nn}(k) \neq -m_{nn}(-k)$ due to TR symmetry breaking. (b) The OMM density response of the trivial phase. The intraband contribution (blue dots) is zero, while the interband contribution (orange triangles) is large and asymmetric. (c) Momentum space distribution of valence band orbital moment in the topologically nontrivial phase of the Haldane model at $\Delta = 0.0$. Notice that $m_{nn}(k) = m_{nn}(-k)$ due to inversion symmetry at $\Delta = 0$. (d) The OMM density response of the nontrivial phase. OMM density is contributed by both intraband (blue dots) and interband (orange triangles) terms. In this scenario, the intraband contribution is antisymmetric with respect to $y = 0$, which leads to opposite density accumulations at the two edges and zero net orbital magnetic response. However, the interband contribution gives rise to a large total OMM density response.

In summary, we have provided a fresh perspective on how TR-odd nonconserved densities respond to an electric field in various types of materials. Remarkably, this helps us to understand the fundamental difference that exists between the quantum spin Hall effect and the quantum Hall effect. In the quantum spin Hall effect TR symmetry requires the edge spin accumulations to go hand-in-hand with edge spin currents, the two being related by spin-momentum locking: undergap states play no role in the process. While the spin transport at the Fermi level is almost ballistic, being protected against elastic scattering by non-magnetic impurities – it is not truly non-dissipative, because it is not protected against other scattering mechanisms, e.g., inelastic scattering, operating at the Fermi level. In the quantum Hall effect, edge states still play a role as hosts for the electric charge accumulation, which produces an electric field perpendicular to the edges. However, the exactly quantized Hall current, which is perpendicular to the electric field and hence truly non-dissipative, is carried by undergap states in the incompressible bulk [63, 64].

ACKNOWLEDGMENTS

This research is supported by the Ministry of Education, Singapore, under its Research Centre of Excellence award to the Institute for Functional Intelligent Materials (I-FIM, project No. EDUNC-33-18-279-V12). A.P. and A.K. acknowledge support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS) and from the Leverhulme Trust under the grant agreement RPG-2023-253.

[1] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[2] B. A. Bernevig and S.-C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).
[3] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[4] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[5] J. Maciejko, T. L. Hughes, and S.-C. Zhang, Annu. Rev. Condens. Matter Phys. **2**, 31 (2011).
[6] I. Sodemann and L. Fu, Phys. Rev. Lett. **115**, 216806 (2015).
[7] K. Kang, T. Li, E. Sohn, J. Shan, and K. F. Mak, Nat. Mater. **18**, 324 (2019).
[8] Q. Ma, S.-Y. Xu, H. Shen, D. MacNeill, V. Fatemi, T.-R. Chang, A. M. Mier Valdivia, S. Wu, Z. Du, C.-H. Hsu, S. Fang, Q. D. Gibson, K. Watanabe, T. Taniguchi, R. J. Cava, E. Kaxiras, H.-Z. Lu, H. Lin, L. Fu, N. Gedik, and P. Jarillo-Herrero, Nature **565**, 337 (2019).
[9] K. von Klitzing, Rev. Mod. Phys. **58**, 519 (1986).
[10] N. Nagaosa, J. Sinova, S. Onoda, A. H. MacDonald, and N. P. Ong, Rev. Mod. Phys. **82**, 1539 (2010).
[11] A. Gianfrate, O. Bleu, L. Dominici, V. Ardizzone, M. De Giorgi, D. Ballarini, G. Lerario, K. W. West, L. N. Pfeiffer, D. D. Solnyshkov, D. Sanvitto, and G. Malpuech, Nature **578**, 381 (2020).
[12] Y. Zhang, Y. Sun, and B. Yan, Phys. Rev. B **97**, 041101 (2018).
[13] C.-X. Liu, S.-C. Zhang, and X.-L. Qi, Annu. Rev. Condens. Matter Phys. **7**, 301 (2016).
[14] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, Science **344**, 1489 (2014).
[15] J. E. Hirsch, Phys. Rev. Lett. **83**, 1834 (1999).
[16] J. Sinova, D. Culcer, Q. Niu, N. A. Sinitsyn, T. Jungwirth, and A. H. MacDonald, Phys. Rev. Lett. **92**, 126603 (2004).
[17] D. Go, D. Jo, C. Kim, and H.-W. Lee, Phys. Rev. Lett. **121**, 086602 (2018).
[18] Y. K. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Science **306**, 1910 (2004).
[19] S. Takahashi and S. Maekawa, Science and Technology of Advanced Materials **9**, 014105 (2008), pMID: 27877931.
[20] Y. Liu, J. Besbas, Y. Wang, P. He, M. Chen, D. Zhu, Y. Wu, J. M. Lee, L. Wang, J. Moon, N. Koirala, S. Oh, and H. Yang, Nat. Commun. **9**, 2492 (2018).
[21] S. Ruiz-Gómez, R. Guerrero, M. W. Khaliq, C. Fernández-González, J. Prat, A. Valera, S. Finizio, P. Perna, J. Camarero, L. Pérez, L. Aballe, and M. Foerster, Phys. Rev. X **12**, 031032 (2022).
[22] Y.-G. Choi, D. Jo, K.-H. Ko, D. Go, K.-H. Kim, H. G. Park, C. Kim, B.-C. Min, G.-M. Choi, and H.-W. Lee, Nature **619**, 52 (2023).
[23] E. Arrighi, V.-H. Nguyen, M. Di Luca, G. Maffione, Y. Hong, L. Farrar, K. Watanabe, T. Taniguchi, D. Mailly, J.-C. Charlier,

and R. Ribeiro-Palau, Nat. Commun. **14**, 8178 (2023).

[24] G. Sala, H. Wang, W. Legrand, and P. Gambardella, Phys. Rev. Lett. **131**, 156703 (2023).

[25] T. Tanaka, H. Kontani, M. Naito, T. Naito, D. S. Hirashima, K. Yamada, and J. Inoue, Phys. Rev. B **77**, 165117 (2008).

[26] T. Tanaka and H. Kontani, Phys. Rev. B **81**, 224401 (2010).

[27] G. Sundaram and Q. Niu, Phys. Rev. B **59**, 14915 (1999).

[28] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press, Cambridge, 2018).

[29] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010).

[30] N. A. Sinitsyn, J. Phys. Condens. Matter **20**, 023201 (2007).

[31] E. Drigo and R. Resta, Phys. Rev. B **101**, 165120 (2020).

[32] A. Kazantsev, A. Mills, E. O'Neill, H. Sun, G. Vignale, and A. Principi, Phys. Rev. Lett. **132**, 106301 (2024).

[33] S. Coh, D. Vanderbilt, A. Malashevich, and I. Souza, Phys. Rev. B **83**, 085108 (2011).

[34] A. M. Essin, A. M. Turner, J. E. Moore, and D. Vanderbilt, Phys. Rev. B **81**, 205104 (2010).

[35] A. Scaramucci, E. Bousquet, M. Fechner, M. Mostovoy, and N. A. Spaldin, Phys. Rev. Lett. **109**, 197203 (2012).

[36] A. A. Kovalev, Y. Tserkovnyak, K. Výborný, and J. Sinova, Phys. Rev. B **79**, 195129 (2009).

[37] S. A. Yang, H. Pan, Y. Yao, and Q. Niu, Phys. Rev. B **83**, 125122 (2011).

[38] M. Milletarì and A. Ferreira, Phys. Rev. B **94**, 134202 (2016).

[39] H. Ishizuka and N. Nagaosa, Sci. Adv. **4**, eaap9962 (2018).

[40] D. Go, T. S. Seifert, T. Kampfrath, K. Ando, H.-W. Lee, and Y. Mokrousov, Local and global reciprocity in orbital-charge-coupled transport (2024), arXiv:2407.00517.

[41] J. Shi, P. Zhang, D. Xiao, and Q. Niu, Phys. Rev. Lett. **96**, 076604 (2006).

[42] M.-C. Chang and Q. Niu, Phys. Rev. B **53**, 7010 (1996).

[43] D. Ceresoli, T. Thonhauser, D. Vanderbilt, and R. Resta, Phys. Rev. B **74**, 024408 (2006).

[44] J. Shi, G. Vignale, D. Xiao, and Q. Niu, Phys. Rev. Lett. **99**, 197202 (2007).

[45] C. Xiao and Q. Niu, Phys. Rev. B **104**, L241411 (2021).

[46] N. Sugimoto, S. Onoda, S. Murakami, and N. Nagaosa, Phys. Rev. B **73**, 113305 (2006).

[47] T. P. Cysne, F. S. M. Guimarães, L. M. Canonico, M. Costa, T. G. Rappoport, and R. B. Muniz, Phys. Rev. B **107**, 115402 (2023).

[48] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).

[49] I. Garate and M. Franz, Phys. Rev. Lett. **104**, 146802 (2010).

[50] C. Gorini, A. Maleki Sheikhabadi, K. Shen, I. V. Tokatly, G. Vignale, and R. Raimondi, Phys. Rev. B **95**, 205424 (2017).

[51] A. Maleki Sheikhabadi, I. Miatka, E. Y. Sherman, and R. Raimondi, Phys. Rev. B **97**, 235412 (2018).

[52] A. Johansson, B. Göbel, J. Henk, M. Bibes, and I. Mertig, Phys. Rev. Res. **3**, 013275 (2021).

[53] L. Salemi and P. M. Oppeneer, Phys. Rev. Mater. **6**, 095001 (2022).

[54] A. Pezo, D. García Ovalle, and A. Manchon, Phys. Rev. B **108**, 075427 (2023).

[55] S. Bhowal and G. Vignale, Phys. Rev. B **103**, 195309 (2021).

[56] Z. Z. Du, C. M. Wang, S. Li, H.-Z. Lu, and X. C. Xie, Nat. Commun. **10**, 3047 (2019).

[57] H. Sun, D. Bhowmick, B. Yang, and P. Sengupta, Phys. Rev. B **107**, 134426 (2023).

[58] C.-Z. Chang, J. Zhang, X. Feng, J. Shen, Z. Zhang, M. Guo, K. Li, Y. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S. Ji, X. Chen, J. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y. Wang, L. Lu, X.-C. Ma, and Q.-K. Xue, Science **340**, 167 (2013).

[59] H. Polshyn, J. Zhu, M. A. Kumar, Y. Zhang, F. Yang, C. L. Tschirhart, M. Serlin, K. Watanabe, T. Taniguchi, A. H. MacDonald, and A. F. Young, Nature **588**, 66 (2020).

[60] D. Go, F. Freimuth, J.-P. Hanke, F. Xue, O. Gomonay, K.-J. Lee, S. Blügel, P. M. Haney, H.-W. Lee, and Y. Mokrousov, Phys. Rev. Res. **2**, 033401 (2020).

[61] D. Go and H.-W. Lee, Phys. Rev. Res. **2**, 013177 (2020).

[62] D. Xiao, W. Yao, and Q. Niu, Phys. Rev. Lett. **99**, 236809 (2007).

[63] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[64] A. J. Bestwick, E. J. Fox, X. Kou, L. Pan, K. L. Wang, and D. Goldhaber-Gordon, Phys. Rev. Lett. **114**, 187201 (2015).

[65] J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (Cambridge University Press, 2020).

[66] V. Edelstein, Solid State Commun. **73**, 233 (1990).

Supplemental Materials

I. MICROSCOPIC RECIPROCITY RELATION IN TIME-REVERSAL INVARIANT SYSTEMS

In this section we provide a general poof of Eq. (22), which we refer to as *microscopic reciprocity relation* for any time-reversal invariant many-body system. Because we don't want to be limited to non-interacting systems we start from the general exact eigenstates (Lehmann) representation of the linear response function [48]

$$\chi_{AB}(\omega) = \sum_{nm} \frac{P_m - P_n}{\omega - \omega_{nm} + i\eta} A_{mn} B_{nm}. \tag{S1}$$

which describes the linear response of the Hermitian observable $\hat{A}$ to an external force that couples linearly to the Hermitian observable $\hat{B}$ (we have set $\hbar = 1$). $A_{mn} = \langle m|\hat{A}|n\rangle$ is the matrix element of $\hat{A}$ between exact many-body eigenstates $|n\rangle$ and $|m\rangle$ with energies E_n and E_m respectively, and $\omega_{nm} = E_n - E_m$. P_n and P_m are the occupation probabilities of states $|n\rangle$ and $|m\rangle$ in the canonical equilibrium ensemble at temperature T. We start from the basic identity (see [65], pp. 277)

$$\langle\alpha|\hat{A}|\beta\rangle = \langle\tilde{\beta}|\hat{T}\hat{A}^{\dagger}\hat{T}^{-1}|\tilde{\alpha}\rangle \tag{S2}$$

where $|\tilde{\alpha}\rangle = \hat{T}|\alpha\rangle$, $|\tilde{\beta}\rangle = \hat{T}|\beta\rangle$ and $\hat{T}$ is the anti-unitary time-reversal operator.

In a time-reversal invariant system for each $|n\rangle$ there is a time-reversed partner $|\tilde{n}\rangle$ with the same energy and the same occupation probability. The sum over all n is equivalent to the sum over all $\tilde{n}$. Therefore we can write

$$\chi_{AB}(\omega) = \sum_{nm} \frac{P_m - P_n}{\omega - \omega_{nm} + i\eta} \tilde{B}_{mn} \tilde{A}_{nm} = \chi_{\tilde{B}\tilde{A}}(\omega) \tag{S3}$$

where

$$\tilde{A} \equiv \hat{T}\hat{A}^{\dagger}\hat{T}^{-1}\,, \quad \tilde{B} \equiv \hat{T}\hat{B}^{\dagger}\hat{T}^{-1}\,. \tag{S4}$$

When dealing with operators of definite symmetry under time reversal we can further assume $\tilde{A} = \pm\hat{A} = \lambda_A \hat{A}$, where $\lambda_A = \pm 1$ is the signature of the operator under time reversal. Then we have

$$\chi_{AB}(\omega) = \lambda_A \lambda_B \chi_{B^{\dagger}A^{\dagger}}(\omega)\,. \tag{S5}$$

Finally, if the operators are Hermitian, we have

$$\chi_{AB}(\omega) = \lambda_A \lambda_B \chi_{BA}(\omega)\,. \tag{S6}$$

Now let us specialize in the case of a *non-interacting* system with $\hat{A}$ and $\hat{B}$ sums of one-particle-operators and $|\alpha\rangle$, $|\beta\rangle$ one-particle eigenstates. Making use of Eqs. (23), (17),(18) and (19) we obtain (with summation over repeated indices implied)

$$\begin{aligned}
\mathrm{Re}[\chi_{AB}(\omega)] =& \frac{1}{2}\,\mathrm{Re}\left[\chi_{AB}(\omega) + \lambda_A\lambda_B\chi_{BA}(\omega)\right] \\
=& \frac{1}{2}\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}] - \frac{1}{2}\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}] + \\
& \frac{1}{2}\lambda_A\lambda_B\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[B_{\alpha\beta}A_{\beta\alpha}] - \frac{1}{2}\lambda_A\lambda_B\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[B_{\alpha\beta}A_{\beta\alpha}] \\
=& \frac{1}{2}(1+\lambda_A\lambda_B)\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}] - \frac{1}{2}(1-\lambda_A\lambda_B)\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}].
\end{aligned} \tag{S7}$$

where $\mathcal{L}^{\eta}_{\alpha\beta}(\omega) = \frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta + \omega + i\eta}$ is the Lindhard factor and

$$\mathrm{Re}[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)] = (f_\alpha - f_\beta)\frac{\epsilon_\alpha - \epsilon_\beta + \omega}{(\epsilon_\alpha - \epsilon_\beta + \omega)^2 + \eta^2}, \quad \mathrm{Im}[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)] \quad = (f_\alpha - f_\beta)\frac{-\eta}{(\epsilon_\alpha - \epsilon_\beta + \omega)^2 + \eta^2}. \tag{S8}$$

For the imaginary part of the response function, similarly, we have:

$$\begin{aligned}
\mathrm{Im}[\chi_{AB}(\omega)] =& \frac{1}{2}\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}] + \frac{1}{2}\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}] + \\
& \frac{1}{2}\lambda_A\lambda_B\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[B_{\alpha\beta}A_{\beta\alpha}] + \frac{1}{2}\lambda_A\lambda_B\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[B_{\alpha\beta}A_{\beta\alpha}] \\
=& \frac{1}{2}(1+\lambda_A\lambda_B)\,\mathrm{Im}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}] + \frac{1}{2}(1-\lambda_A\lambda_B)\,\mathrm{Re}\left[\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\right]\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}].
\end{aligned} \tag{S9}$$

Eqs. (S7) and (S9) combined yield Eq. (24) of the main text:

$$\chi_{AB}(\omega) = \frac{1+\lambda_A\lambda_B}{2}\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\,\mathrm{Re}[A_{\alpha\beta}B_{\beta\alpha}] + i\frac{1-\lambda_A\lambda_B}{2}\mathcal{L}^{\eta}_{\alpha\beta}(\omega)\,\mathrm{Im}[A_{\alpha\beta}B_{\beta\alpha}]\,. \tag{S10}$$

II. GENERALIZED DENSITY ACCUMULATION IN A DIFFUSIVE METAL

In this Appendix we make use of Eq. (27) to calculate the static density accumulation. Making use of the identities $(f_\alpha - f_\beta)\delta(\epsilon_\alpha - \epsilon_\beta + \omega) \simeq (\epsilon_\alpha - \epsilon_\beta)f'(\epsilon_\alpha)\delta(\epsilon_\alpha - \epsilon_\beta + \omega)$ (valid in the limit $\omega \to 0$), where $f'(\epsilon_\alpha)$ is the derivative of the Fermi-Dirac distribution with respect to its own argument and $(\epsilon_\alpha - \epsilon_\beta)[\hat{\mathbf{r}}]_{\beta\alpha} = i\hbar[\hat{\mathbf{v}}]_{\beta\alpha}$ we can write

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},0) = \pi\hbar\sum_{\alpha\beta} f'(\epsilon_\alpha)\,\mathrm{Re}\{[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{v}}]_{\beta\alpha}\}\delta(\epsilon_\alpha - \epsilon_\beta), \tag{S11}$$

where the derivative of the Fermi-Dirac distribution, $f'(\epsilon_\alpha) \simeq -\delta(\epsilon_\alpha - \epsilon_F)$ (with ϵ_F the Fermi energy) forces ϵ_α (and therefore also ϵ_β) to be at the Fermi level. To make further progress, we need to make some assumptions about the nature of the single-particle eigenstates $|\alpha\rangle$ and $|\beta\rangle$. We work in the framework of the *relaxation time approximation*, namely, we assume that the eigenstates are scattering states characterized by a dominant Bloch wave vector $\mathbf{k}$ with a lifetime τ due to impurity scattering. Furthermore, we can assume that only states with the same $\mathbf{k}$ are connected by the operators $\hat{\mathbf{r}}$, $\hat{\mathbf{v}}$, etc... This implies that only energy eigenstates separated by energies of the order of $\hbar/\tau$ are connected. While this approach is often qualitatively correct, we caution the reader that it may occasionally be spectacularly wrong, as it neglects what in diagrammatic language is termed "vertex corrections". When vertex corrections are important, qualitative discussion is not sufficient, and a more careful treatment of the disorder is required.

Under the assumptions of the relaxation time approximation, the sum

$$\sum_{\beta}\mathrm{Re}\{[\hat{n}_O(\mathbf{r})]_{\alpha\beta}[\hat{\mathbf{v}}]_{\beta\alpha}\}\delta(\epsilon_F - \epsilon_\beta) \tag{S12}$$

can be approximated by a Lorentzian of width $1/\tau$ centered at ϵ_F:

$$\frac{1}{\pi}\frac{\tau^{-1}}{(\epsilon_\alpha - \epsilon_F)^2 + \tau^{-2}}\left\langle \mathrm{Re}\{\hat{n}_O(\mathbf{r})]_{\mathbf{k},\mathbf{k}}[\hat{\mathbf{v}}]_{\mathbf{k},\mathbf{k}}\}\right\rangle_{FS}\,, \tag{S13}$$

where the angular bracket denotes the average of the product of matrix elements between states at the Fermi level. Then Eq. (S11) works out to be

$$\chi_{n_O,\mathbf{r}}(\mathbf{r},0) = N(\epsilon_F)\tau\left\langle \mathrm{Re}\{[\hat{n}_O(\mathbf{r})]_{\mathbf{k},\mathbf{k}}[\hat{\mathbf{v}}]_{\mathbf{k},\mathbf{k}}\}\right\rangle_{FS}\,, \tag{S14}$$

where $N(\epsilon_F) = \sum_\alpha \delta(\epsilon_F - \epsilon_\alpha)$ is the density of states at the Fermi level. This result is for a single sheet of the Fermi surface. If there are multiple sheets with partial densities of states $N_i(\epsilon_F)$ and relaxation times τ_i we simply sum over them and recover Eq. (**??**) of the main text. This result diverges in the limit $\tau \to \infty$, consistent with the fact that in the absence of momentum relaxation, the electric field would drive an infinite parallel current. Indeed $N(\epsilon_F)\tau$ (times $e^2v_F^2$) is essentially the Drude conductivity of a disordered metal in the diffusive regime. It is a measure of the shift of the Fermi surface under the action of an electric field. The fact that the induced density is proportional to this shift clearly shows that the density response is inseparable from dissipation.

As a simple illustration of this formula, consider the two-dimensional Rashba electron gas with one-particle Hamiltonian

$$H = \frac{\hbar^2k^2}{2m} + \hbar\alpha(\hat{\mathbf{z}}\times\mathbf{k})\cdot\boldsymbol{\sigma} \tag{S15}$$

where α is the Rashba constant (a velocity), $\hat{\mathbf{z}}$ is a unit vector perpendicular to the plane, and $\boldsymbol{\sigma}$ is the spin. The Fermi "surface" (at sufficiently high electron density) consists of two concentric circles, one at $k = k_F + m\alpha/\hbar$ (denoted by +) and the other at $k = k_F - m\alpha/\hbar$, (denoted by $-$), where k_F is related to the two-dimensional electronic density n by $n = \frac{\hbar^2k_F^2 + m^2\alpha^2}{2\pi\hbar^2}$.

We want to calculate the spin density in the y direction induced by an electric field E in the x direction. Because the Hamiltonian (S15) is invariant under time-reversal we can use Eq. (27), which gives

$$S_y = \frac{\hbar}{2}\left\{N_+(\epsilon_F)\tau\langle[\hat{\sigma}_y]_{\mathbf{k},\mathbf{k}}[\hat{v}_x]_{\mathbf{k},\mathbf{k}}\rangle_+ + N_-(\epsilon_F)\tau\langle[\hat{\sigma}_y]_{\mathbf{k},\mathbf{k}}[\hat{v}_x]_{\mathbf{k},\mathbf{k}}\rangle_-\right\}eE\,, \tag{S16}$$

where $[\hat{\mathrm{v}}_x]_{\mathbf{k},\mathbf{k}} = \frac{1}{\hbar}\frac{\partial H}{\partial k_x} = \hbar k_x/m - \alpha[\hat{\sigma}_y]_{\mathbf{k},\mathbf{k}}$. We have assumed that the relaxation time has the same value on both sheets of the Fermi surface. On both Fermi circles the velocity points radially outward with a common value $v_{F+} = v_{F,-} = v_F = \hbar k_F/m$. The spin $[\hat{\boldsymbol{\sigma}}]_{\mathbf{k},\mathbf{k}}$ is tangential to the circles pointing clockwise for the + sheet and counterclockwise for the − sheet. Finally, the density of states on the two sheets is given by

$$N_{\pm}(\epsilon_F) = \frac{m}{2\pi\hbar^2}\left(1 \pm \frac{\alpha}{v_F}\right) . \tag{S17}$$

Making use of this information, it is a simple exercise to recover the well-known result for the current induced spin polarization (also known as Edelstein effect) [66]

$$S_y = \frac{\hbar}{2}\frac{m}{\pi\hbar^2}\alpha\tau e E . \tag{S18}$$

III. DERIVATION OF BERRY CURVATURE FORMULAS FOR GENERALIZED HALL CONDUCTIVITIES

Here we check that the well-known formulas connecting the Hall conductivity to the Berry curvature are recovered in our formalism. In Eq. (S1) we set $\hat{A} = \hat{\mathbf{J}}_O(\mathbf{r}) = \sum_i \hat{O}_i \star \hat{\mathbf{v}}_i \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i)$ (the generalized current density operator) and $\hat{B} = e\hat{\mathbf{r}}$ (the dipole operator). The response function in the DC limit ($\omega \to 0$) is given by

$$\chi_{j_O,\mathbf{r}}(\mathbf{r},0) = 2P\sum_{\alpha\beta} f_\alpha \frac{\mathrm{Re}\{[\hat{\mathbf{J}}_O]_{\alpha\beta}(\mathbf{r})[\hat{\mathbf{r}}]_{\beta\alpha}\}}{\epsilon_\alpha - \epsilon_\beta} , \tag{S19}$$

and the current response has the form

$$\delta J_O^a(\mathbf{r},0) = e\sum_c \chi^a_{J_O^b,r_c}(\mathbf{r},0)E_c. \tag{S20}$$

The conductivity is given by

$$\sigma^a_{bc}(\mathbf{r}) = e\chi^a_{J_O^b,r_c}(\mathbf{r},0). \tag{S21}$$

In a periodic system, the state $\alpha = (n,\mathbf{k})$ is labeled by the band index n and the reciprocal wave vector $\mathbf{k}$. Thus,

$$\begin{aligned}\sigma^a_{bc}(\mathbf{r}) =& 2P\sum_{n,n',\mathbf{k}} \frac{e f_{n\mathbf{k}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \mathrm{Re}\{[\hat{J}^b_{O^a}]_{n\mathbf{k},n'\mathbf{k}}(\mathbf{r})A^c_{n'n}\} \\ =& \frac{1}{2}\sum_{n\neq n',\mathbf{k}} \frac{e f_{n\mathbf{k}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \mathrm{Re}\left[\sum_l \left[\{\hat{O}^a,\hat{v}^b\}_{n,l}(\mathbf{k})\psi^\dagger_{l\mathbf{k}}\psi_{n'\mathbf{k}}(\mathbf{r}) + \{\hat{O}^a,\hat{v}^b\}_{l,n'}(\mathbf{k})\psi^\dagger_{n\mathbf{k}}\psi_{l\mathbf{k}}(\mathbf{r})\right] A^c_{n'n}\right].\end{aligned} \tag{S22}$$

Here, we have employed the expression $\mathbf{r}_{m\mathbf{k},n\mathbf{k}'} = \delta_{mn} i\frac{\partial}{\partial\mathbf{k}}\delta\left(\mathbf{k}-\mathbf{k}'\right) + \delta\left(\mathbf{k}-\mathbf{k}'\right)\mathbf{A}_{mn}(\mathbf{k})$ with $\mathbf{A}_{mn}(\mathbf{k}) = \langle u_{m\mathbf{k}}|i\partial_{\mathbf{k}}u_{n\mathbf{k}}\rangle$. Thus, the global conductivity $\sigma^a_{bc} = \int \sigma^a_{bc}(\mathbf{r})d\mathbf{r}$ is given by

$$\begin{aligned}\sigma^a_{bc} =& 2\sum_{n\neq n',\mathbf{k}} \frac{e f_{n\mathbf{k}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \mathrm{Re}\left\{[\hat{O}^a \star \hat{v}^b]_{n,n'}(\mathbf{k})A^c_{n'n}(\mathbf{k})\right\} \\ =& \frac{e}{\hbar}\sum_{n,\mathbf{k}} f_{n\mathbf{k}} \sum_{n'\neq n} \frac{-2\hbar^2\,\mathrm{Im}\left\{[\hat{O}^a \star \hat{v}^b]_{n,n'}(\mathbf{k})v^c_{n'n}(\mathbf{k})\right\}}{(\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}})^2} \\ =& \frac{e}{\hbar}\sum_{n,\mathbf{k}} f_{n\mathbf{k}}\Omega^a_{n,bc}(\mathbf{k}),\end{aligned} \tag{S23}$$

where $A^c_{n'n}(\mathbf{k}) = \frac{i\hbar v^c_{n'n}(\mathbf{k})}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}}$, and $\Omega^a_{n,bc}(\mathbf{k})$ is the generalized Berry curvature associated with the operator $\hat{O}$. Its explicit form is

$$\Omega^a_{n,bc}(\mathbf{k}) = \sum_{n'\neq n} \frac{-2\hbar^2\,\mathrm{Im}\left\{[\hat{O}^a \star \hat{v}^b]_{n,n'}(\mathbf{k})v^c_{n'n}(\mathbf{k})\right\}}{(\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}})^2}. \tag{S24}$$

IV. THE TORQUE DENSITY

In this section, we provide a brief derivation of Eq. (44) from the main text. We start from Eq. (37), which we restate here for convenience:

$$T_O(\mathbf{r},\omega) = e\left[\chi_{T_O,\mathbf{r}}(\omega,\mathbf{r}) + \frac{1}{i\hbar}\sum_i \langle[\hat{n}_O,\hat{\mathbf{r}}_i]\rangle_0\right]\cdot\mathbf{E}, \tag{S25}$$

The net torque we consider in the main text is given by

$$\bar{T}_O = \lim_{\omega\to 0}\int T_O(\mathbf{r},\omega)d\mathbf{r} = e\mathbf{E}\cdot\lim_{\omega\to 0}\bar{\chi}_{T_O,\mathbf{r}}(\omega) + \frac{e}{i\hbar}\int\sum_i\langle[\hat{n}_O,\hat{\mathbf{r}}_i]\rangle_0\cdot\mathbf{E}d\mathbf{r} \tag{S26}$$

where the first term on the right-hand side is given as

$$\lim_{\omega\to 0}\bar{\chi}_{T_O,\mathbf{r}}(\omega) = P\sum_{\alpha\beta}\frac{f_\alpha - f_\beta}{\epsilon_\alpha - \epsilon_\beta}[\hat{\mathbf{r}}]_{\beta\alpha}\int T^O_{\alpha\beta}(\mathbf{r})d\mathbf{r}\,. \tag{S27}$$

The matrix element of the torque density operator is given by

$$T^O_{\alpha\beta}(\mathbf{r}) = \langle\alpha|\hat{T}_O(\mathbf{r})|\beta\rangle = \frac{1}{i\hbar}\sum_i\langle\alpha|[\hat{O}_i,H_0]\star\delta(\mathbf{r}-\hat{\mathbf{r}}_i)|\beta\rangle\,. \tag{S28}$$

Making use of the identity

$$\int T^O_{\alpha\beta}(\mathbf{r})d\mathbf{r} = \frac{1}{i\hbar}\sum_i\langle\alpha|[\hat{O}_i,H_0]|\beta\rangle = \frac{1}{i\hbar}(\epsilon_\beta - \epsilon_\alpha)O_{\alpha\beta}\,, \tag{S29}$$

we obtain

$$\lim_{\omega\to 0}\bar{\chi}_{T_O,\mathbf{r}}(\omega) = -\frac{1}{i\hbar}\sum_{\alpha\beta}P\left(f_\alpha - f_\beta\right)[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}, \tag{S30}$$

where $[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}$ is shorthand for $\sum_i[\hat{O}_i]_{\alpha\beta}[\hat{\mathbf{r}}_i]_{\beta\alpha}$. This is Eq. (39) of the main text.

Now let us calculate the integrated commutator term in Eq. (S26), we have

$$\frac{1}{i\hbar}\int\sum_i\langle[\hat{n}_O,\hat{\mathbf{r}}_i]\rangle_0 d\mathbf{r} = \frac{1}{i\hbar}\sum_i\sum_\alpha[\hat{O}_i,\hat{\mathbf{r}}_i]_{\alpha\alpha}f_\alpha = \frac{1}{i\hbar}\sum_{\alpha\beta}\left(f_\alpha - f_\beta\right)[\hat{O}]_{\alpha\beta}[\hat{\mathbf{r}}]_{\beta\alpha}. \tag{S31}$$

We thus recover Eq. (40) in the main text.

The combination of Eqs. (S30) and (S31) leads to massive cancelation, after which only the terms only with $\epsilon_\alpha \simeq \epsilon_\beta$, as given by Eq. (V) in the main text. In order to express the result more compactly, we separate the position operator $\hat{\mathbf{r}}$ into two parts, intraband $\hat{\mathbf{R}}$ and interband $\hat{\mathbf{X}}$. After employing the commutators

$$\begin{aligned}
[\hat{O},\hat{\mathbf{R}}]_{\alpha,\beta} &= -i\frac{\partial O_{\alpha\beta}}{\partial\mathbf{k}} - O_{\alpha\beta}(\mathbf{A}_{\alpha\alpha} - \mathbf{A}_{\beta\beta}),\\
[\hat{O},\hat{\mathbf{X}}]_{\alpha,\beta} &= \sum_\gamma(1-\delta_{\gamma\beta})O_{\alpha\gamma}\mathbf{A}_{\gamma\beta} - \sum_\gamma(1-\delta_{\gamma\alpha})\mathbf{A}_{\alpha\gamma}O_{\gamma\beta},
\end{aligned} \tag{S32}$$

we finally obtain

$$\begin{aligned}
\bar{T}_O =& e\mathbf{E}\cdot\sum_\alpha\left\{-\frac{1}{\hbar}\frac{\partial O_{\alpha\alpha}}{\partial\mathbf{k}} - \frac{i}{\hbar}[\hat{O},\hat{\mathbf{A}}]_{\alpha\alpha} + \frac{i}{\hbar}\sum_{\beta\neq\alpha}\left[O_{\alpha\beta}A_{\beta\alpha} - O_{\beta\alpha}A_{\alpha\beta}\right]\right\}f_\alpha\\
=& -\frac{e}{\hbar}\sum_\alpha\mathbf{E}\cdot\frac{\partial O_{\alpha\alpha}}{\partial\mathbf{k}}f_\alpha\\
=& \frac{ie}{\hbar}\langle\mathcal{D}\hat{O}\rangle_{FS}\cdot\mathbf{E},
\end{aligned} \tag{S33}$$

where we have introduced the notation

$$\langle \mathcal{D}\hat{O}\rangle_{FS} = i\sum_{\alpha} \frac{\partial O_{\alpha\alpha}}{\partial \mathbf{k}} f_\alpha = -i\sum_{\alpha} O_{\alpha\alpha}\frac{\partial f_\alpha}{\partial \mathbf{k}}\,, \tag{S34}$$

which is the explicit form of the commutator $i[\hat{O}_{FS}, \hat{\mathbf{r}}_{FS}]$ in Bloch system. Using integration by parts, we can see this is the Fermi surface contribution of the net torque.

We now turn our attention to deriving the proper current formula, as presented in Eq. (51) of Sec. IV. Compared to Eq. (9), the proper current density operator shares a similar structure, expressed as

$$\hat{\mathcal{J}}_O = \sum_i \partial_t(\hat{O}_i \star \hat{\mathbf{r}}_i) \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i), \tag{S35}$$

Thus, deriving the proper current response becomes as straightforward as replacing $\hat{O}$ with $\hat{O} \star \hat{\mathbf{r}}$ in Eq. (S33). This yields

$$\bar{\mathcal{J}}_O = \frac{e}{\hbar}\sum_{\alpha} \mathbf{E}\cdot\frac{\partial f_\alpha}{\partial \mathbf{k}}[\hat{O}\star\hat{\mathbf{r}}]_{\alpha\alpha} = \frac{ie}{\hbar}\langle \mathcal{D}(\hat{O}\star\hat{\mathbf{r}})\rangle_{FS}\cdot\mathbf{E}, \tag{S36}$$

thus, we recover the Eq. (51), which clearly shows a Fermi surface contribution.

V. EXPRESSION OF TORQUE DENSITY IN THE LIMIT OF SLOW SPATIAL VARIATION

In this section, we show how the non-uniform component of the torque density can be expressed, in the limit of slow spatial variation, as the divergence of the torque dipole density. We know that $\hat{\mathcal{T}}_O(\mathbf{r}) = \sum_i (\partial_t\hat{O}_i) \star \delta(\mathbf{r} - \hat{\mathbf{r}}_i)$, which in $q-$space, with Fourier transformation $\hat{\mathcal{T}}_O(\mathbf{q}) = \int d\mathbf{r}e^{-i\mathbf{q}\cdot\mathbf{r}}\hat{\mathcal{T}}_O(\mathbf{r})$, is given by

$$\hat{\mathcal{T}}_O(\mathbf{q}) = \sum_i (\partial_t\hat{O}_i) \star e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}_i}. \tag{S37}$$

For $\mathbf{q} = 0$, $\hat{\mathcal{T}}_O(\mathbf{0})$ gives the uniform component of the torque, $\hat{\bar{\mathcal{T}}}_O$. For $\mathbf{q} \neq 0$ and for slow spatial variation ($qa \ll 1$) we expand

$$\hat{\mathcal{T}}_O(\mathbf{q}) \approx \mathbf{q}\cdot\sum_i \partial_{\mathbf{q}}\left[(\partial_t\hat{O}_i)\star e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}_i}\right] = -i\mathbf{q}\cdot\sum_i \hat{\mathbf{r}}_i \star (\partial_t\hat{O}_i)\star e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}_i} = -i\mathbf{q}\cdot\hat{\mathbf{D}}_O(\mathbf{q}). \tag{S38}$$

where $\hat{\mathbf{D}}_O(\mathbf{q}) = \sum_i \hat{\mathbf{r}}_i \star (\partial_t\hat{O}_i)\star e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}_i}$ is the torque dipole density operator. By transforming this expression to real space, we have

$$\hat{\mathbf{D}}_O(\mathbf{r}) = \int d\mathbf{q}\sum_i \hat{\mathbf{r}}_i \star (\partial_t\hat{O}_i)\star e^{i\mathbf{q}\cdot(\mathbf{r}-\hat{\mathbf{r}}_i)} = \sum_i \hat{\mathbf{r}}_i \star (\partial_t\hat{O}_i)\star \delta(\mathbf{r}-\hat{\mathbf{r}}_i). \tag{S39}$$

This is the right operator form of Eq. (45) in the main text. And the torque density operator now is rewritten as

$$\hat{\mathcal{T}}_O(\mathbf{r}) \approx \frac{1}{V}\hat{\bar{\mathcal{T}}}_O - \nabla_{\mathbf{r}}\cdot\hat{\mathbf{D}}_O(\mathbf{r}). \tag{S40}$$

where V is the volume of the system.

VI. SOLUTION OF THE HALDANE MODEL WITH STAGGERED ONSITE POTENTIALS

In this section, we discuss several solutions of the Haldane model. First we introduce the model in real space:

$$H = -t_0\sum_{\langle ij\rangle}\hat{a}_i^\dagger\hat{b}_j + t_2 e^{i\phi}\sum_{\langle\langle ij\rangle\rangle}\hat{a}_i^\dagger\hat{a}_j + t_2 e^{-i\phi}\sum_{\langle\langle ij\rangle\rangle}\hat{b}_i^\dagger\hat{b}_j + h.c. + \Delta\sum_i(\hat{a}_i^\dagger\hat{a}_j - \hat{b}_i^\dagger\hat{b}_j). \tag{S41}$$

Here, t_0 represents the nearest-neighbor hopping amplitude between the A and B sublattices, which are connected via three bond vectors, $\boldsymbol{\delta}_n$, defined as $\boldsymbol{\delta}_1 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$, $\boldsymbol{\delta}_2 = \left(\frac{-1}{2}, \frac{1}{2\sqrt{3}}\right)$, and $\boldsymbol{\delta}_3 = \left(0, \frac{-1}{\sqrt{3}}\right)$. The parameter t_2 corresponds to the next-nearest-neighbor hopping amplitude within the A (or B) sublattices, connected by the vectors $\mathbf{l}_n$, where $\mathbf{l}_1 = (1, 0)$, $\mathbf{l}_2 = \left(\frac{-1}{2}, \frac{\sqrt{3}}{2}\right)$, and

$\mathbf{l}_3 = \left(\frac{-1}{2}, \frac{-\sqrt{3}}{2}\right)$. The phase ϕ is the staggered magnetic flux, which breaks time-reversal symmetry, while Δ refers to the onsite potential that breaks inversion symmetry.

After Fourier transformation, we have the Hamiltonian in k-space,

$$H = \sum_{\mathbf{k}} (a^\dagger_{\mathbf{k}}, b^\dagger_{\mathbf{k}}) \begin{pmatrix} \Delta + 2t_2\beta^a_{\mathbf{k}} & -t_0\gamma_{\mathbf{k}} \\ -t_0\gamma^*_{\mathbf{k}} & -\Delta + 2t_2\beta^b_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ b_{\mathbf{k}} \end{pmatrix}, \tag{S42}$$

where $\gamma_{\mathbf{k}} = \sum_n e^{i\mathbf{k}\cdot\delta_n}$, $\beta^a_{\mathbf{k}} = \sum_n \mathrm{Re}\left[e^{i(\mathbf{k}\cdot\mathbf{l}_n+\phi)}\right] = \sum_n \cos(\mathbf{k}\cdot\mathbf{l}_n + \phi)$, and $\beta^b_{\mathbf{k}} = \sum_n \mathrm{Re}\left[e^{i(\mathbf{k}\cdot\mathbf{l}_n-\phi)}\right] = \sum_n \cos(\mathbf{k}\cdot\mathbf{l}_n - \phi)$. In the main text, we set $\phi = \frac{\pi}{2}$ for the convenience. The eigenvalues are given by the equation $\left(\Delta + 2t_2\beta^a_{\mathbf{k}} - E\right)\left(\Delta - 2t_2\beta^b_{\mathbf{k}} + E\right) + t_0^2|\gamma_{\mathbf{k}}|^2 = 0$, which has solutions

$$\epsilon_{1,2}(\mathbf{k}) = -t_2(\beta^a_{\mathbf{k}} + \beta^b_{\mathbf{k}}) \pm \sqrt{t_0^2|\gamma_{\mathbf{k}}|^2 + t_2^2(\beta^a_{\mathbf{k}} + \beta^b_{\mathbf{k}})^2 + \left(\Delta + 2t_2\beta^a_{\mathbf{k}}\right)\left(\Delta - 2t_2\beta^b_{\mathbf{k}}\right)}, \tag{S43}$$

The corresponding eigenfunctions are

$$|u_{1\mathbf{k}}\rangle = \begin{pmatrix} \frac{|\gamma_{\mathbf{k}}|u(\mathbf{k})}{\gamma^*_{\mathbf{k}}\sqrt{4t_0^2|\gamma_{\mathbf{k}}|^2+u^2(\mathbf{k})}} \\ \frac{2t_0|\gamma_{\mathbf{k}}|}{\sqrt{4t_0^2|\gamma_{\mathbf{k}}|^2+u^2(\mathbf{k})}} \end{pmatrix}, \quad |u_{2\mathbf{k}}\rangle = \begin{pmatrix} \frac{|\gamma_{\mathbf{k}}|v(\mathbf{k})}{\gamma^*_{\mathbf{k}}\sqrt{4t_0^2|\gamma_{\mathbf{k}}|^2+v^2(\mathbf{k})}} \\ \frac{2t_0|\gamma_{\mathbf{k}}|}{\sqrt{4t_0^2|\gamma_{\mathbf{k}}|^2+v^2(\mathbf{k})}} \end{pmatrix}. \tag{S44}$$

where $-2\Delta - 2t_2\beta^a_{\mathbf{k}} + 2t_2\beta^b_{\mathbf{k}} + \epsilon_{2\mathbf{k}} - \epsilon_{1\mathbf{k}} = u(\mathbf{k})$, and $-2\Delta - 2t_2\beta^a_{\mathbf{k}} + 2t_2\beta^b_{\mathbf{k}} - \epsilon_{2\mathbf{k}} + \epsilon_{1\mathbf{k}} = v(\mathbf{k})$.

When $t_2 = 0$ and $\Delta \neq 0$ the model reduces to TR symmetric "gapped graphene", whose band dispersion is

$$\epsilon_{1,2}(\mathbf{k}) = \pm\sqrt{t_0^2|\gamma_{\mathbf{k}}|^2 + \Delta^2}. \tag{S45}$$

The above formulas are for the bulk periodic system. For the nanoribbon geometry the bands are calculated numerically, for TR invariant and TR non-invariant systems, both in the trivial and nontrivial topological phases. The results sre shown in Fig. 1(b)-(d) of the main text.

VII. THE OMM DENSITY RESPONSE IN THE RIBBON GEOMETRY

This section focuses on the OMM response in the ribbon geometry. In the ribbon structure, we set x as the periodic (longitudinal) direction along which the electric field is applied and y as the transverse direction (with open boundary condition) along which the OMM density accumulation is computed. The OMM density induced by the electric field is given by

$$\delta n^z_m(y,\omega) = \frac{ieE_x}{i\eta + \omega}\int dx \sum_{n,k} \frac{\partial f_{nk}}{\partial k} n^m_{nn}(x,y,k) + eE_x \int dx \sum_{nn'k} \frac{f_{nk} - f_{n'k}}{\epsilon_{nk} - \epsilon_{n'k} + i\eta + \omega} n^m_{nn'}(x,y,k) A_{n'n}(k). \tag{S46}$$

The matrix elements of the OMM density operator are given by

$$n^m_{nn'}(x,y,k) = \frac{1}{2}\sum_l [m_{nl}(k)\psi^\dagger_{lk}(x,y)\psi_{n'k}(x,y) + m_{ln'}(k)\psi^\dagger_{nk}(x,y)\psi_{lk}(x,y)]. \tag{S47}$$

where $\psi_{nk}(x,y)$ (note that $\psi_{nk}(x,y)$ is not Bloch function) represents the wave function at real-space coordinates (x,y) for band n with wavevector k. In the tight-binding basis, $\psi_{nk}(x,y)$ is expressed as

$$\psi_{nk}(x,y) = \frac{1}{\sqrt{N}}\sum_R \sum_\mu C^{nk}_\mu e^{ik(R+r^x_\mu)}\phi_{R,\mu}(x,y), \tag{S48}$$

where R labels the periodic cells along the x-direction, μ denotes the sublattice within each periodic cell, and $\phi_{R,\mu}(x,y)$ is the tight-binding basis function. We assume that the latter is a δ-function centered at the lattice site:

$$\phi_{R,\mu}(x,y) = \delta(x - R - r^x_\mu)\delta(y - r^y_\mu), \tag{S49}$$

where (r^x_μ, r^y_μ) are the position coordinates of sublattice μ in the unit cell at $R = 0$. Thus, the x and the y-coordinates are replaced by discrete lattice-site coordinates. The coefficient C^{nk}_μ in Eq. (S48) can be obtained by diagonalizing the tight-binding Hamiltonian.

The integration of x in function $\psi^{\dagger}_{nk}(x,y)\psi_{lk}(x,y)$ gives

$$\begin{aligned}\int \psi^{\dagger}_{nk}(x,y)\psi_{lk}(x,y)dx =& \frac{1}{N}\sum_{RR'}\sum_{\mu\mu'} C^{nk*}_{\mu}C^{lk}_{\mu'}e^{ik(R+r^x_{\mu}-R'-r^x_{\mu'})}\delta(R+r^x_{\mu}-R'-r^x_{\mu'})\delta(y-r^y_{\mu'})\delta(y-r^y_{\mu})\\ =& \frac{1}{N}\sum_{RR'}\sum_{\mu\mu'} C^{nk*}_{\mu}C^{lk}_{\mu'}\delta_{R,R'}\delta_{r^x_{\mu},r^x_{\mu'}}\delta_{r^y_{\mu},r^y_{\mu'}}\delta(y-r^y_{\mu})\\ =& \sum_{\mu} C^{nk*}_{\mu}C^{lk}_{\mu}\delta(y-r^y_{\mu}). \end{aligned} \tag{S50}$$

After integrating over x, we obtain the matrix elements:

$$n^m_{nn'}(y,k) = \int n^m_{nn'}(x,y,k)dx = \frac{1}{2}\sum_{l\mu}[m_{nl}(k)C^{lk*}_{\mu}C^{n'k}_{\mu} + m_{ln'}(k)C^{nk*}_{\mu}C^{lk}_{\mu}]\delta(y-r^y_{\mu}). \tag{S51}$$

VIII. THE MATRIX ELEMENTS OF THE OMM OPERATOR

In this section, we calculate the matrix elements of the OMM operator in the ribbon geometry

$$\begin{aligned} m_{nl}(k) =& \frac{e}{4i\hbar}\langle nk|\{\hat{r}^x\hat{r}^y-\hat{r}^y\hat{r}^x, H\} - 2\hat{r}^xH\hat{r}^y + 2\hat{r}^yH\hat{r}^x|lk\rangle\\ =& \frac{e}{4i\hbar}\left(\epsilon_{lk}+\epsilon_{nk}\right)\sum_{n'k'} r^x_{nn'}(k,k')r^y_{n'l}(k',k) - r^y_{nn'}(k,k')r^x_{n'l}(k',k) - \frac{e}{4i\hbar}\langle nk|2\hat{r}^xH\hat{r}^y - 2\hat{r}^yH\hat{r}^x|lk\rangle\\ =& \frac{e}{4i\hbar}\sum_{n'k'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k'}\right)\left[r^x_{nn'}(k,k')r^y_{n'l}(k',k) - r^y_{nn'}(k,k')r^x_{n'l}(k',k)\right]. \end{aligned} \tag{S52}$$

Employing $\hat{\mathbf{v}} = \frac{1}{i\hbar}\left[\hat{\mathbf{r}}, H\right]$, and $r^x_{nn'}(k,k') = i\delta_{nn'}\partial_k\delta(k-k') + \delta(k-k')A_{nn'}(k)$ with $A_{nn'}(k) = i\langle u_{nk}|\partial_k u_{n'k}\rangle$, we obtain

$$\begin{aligned} \sum_{n'k'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k'}\right)r^x_{nn'}(k,k')r^y_{n'l}(k',k) =& \sum_{n'k'}\left[i\delta_{nn'}\partial_k\delta(k-k') + \delta(k-k')A_{nn'}(k)\right]\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k'}\right)r^y_{n'l}(k',k)\\ =& \, i\sum_{k'}\partial_k\delta(k-k')\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{nk'}\right)r^y_{nl}(k',k) + \sum_{n'}A_{nn'}(k)r^y_{n'l}(k)\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k}\right)\\ =& -2i\partial_k[\epsilon_{nk}]r^y_{nl}(k) + \sum_{n'}A_{nn'}(k)r^y_{n'l}(k)\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k}\right) \end{aligned} \tag{S53}$$

Similarly, we have the

$$\begin{aligned} \sum_{n'k'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k'}\right)r^y_{nn'}(k,k')r^x_{n'l}(k',k) =& \sum_{n'k'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k'}\right)r^y_{nn'}(k,k')\left[-i\delta_{n'l}\partial_k\delta(k-k') + \delta(k-k')A_{n'l}(k)\right]\\ =& \, i2\partial_k[\epsilon_{lk}]r^y_{nl}(k) + \sum_{n'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k}\right)r^y_{nn'}(k)A_{n'l}(k). \end{aligned} \tag{S54}$$

Finally, we have

$$m_{nl}(k) = \frac{-e}{2\hbar}\partial_k[\epsilon_{nk}+\epsilon_{lk}]r^y_{nl}(k) + \frac{e}{4i\hbar}\sum_{n'}\left(\epsilon_{lk}+\epsilon_{nk}-2\epsilon_{n'k}\right)\left[A_{nn'}(k)r^y_{n'l}(k) - r^y_{nn'}(k)A_{n'l}(k)\right]. \tag{S55}$$

arXiv:2401.02872v2 [cond-mat.str-el] 16 Oct 2024

Weak Coupling Theory of Magic-Angle Twisted Bilayer Graphene

Jihang Zhu,[1, 2] Iacopo Torre,[3] Marco Polini,[4, 5, 3] and Allan H. MacDonald[6]
[1]*Condensed Matter Theory Center and Joint Quantum Institute,*
Department of Physics, University of Maryland, College Park, Maryland 20742, USA
[2]*Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[3]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology,*
Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain
[4]*Dipartimento di Fisica dell'Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy*
[5]*Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy*
[6]*Physics Department, University of Texas at Austin, Austin TX 78712 USA*

Strong correlations occur in magic-angle twisted bilayer graphene (MATBG) when the octet of flat moiré minibands centered on charge neutrality (CN) is partially occupied. The octet consists of a single valence band and a single conduction band for each of four degenerate spin-valley flavors. Motivated by the importance of Hartree electrostatic interactions in determining the filling-factor dependent band structure, we use a time-dependent Hartree approximation to gain insight into electronic correlations. We find that the electronic compressibility is dominated by Hartree interactions, that paramagnetic states are stable over a range of density near CN, and that the dependence of energy on flavor polarization is strongly overestimated by mean-field theory.

Introduction— The energy bands of twisted bilayer graphene (TBG) have a four-fold spin-valley flavor degeneracy. As a magic twist angle near $\theta = 1^\circ$ is approached, the two sets of four-fold degenerate bands closest to the neutral system Fermi energy approach each other and narrow [1], converting graphene from a weakly-correlated Fermi liquid to a strongly correlated system [2–5] with a rich variety of competing states, including superconductors, insulating flavor ferromagnets, and metallic flavor ferromagnets. The ferromagnetism is reminiscent of but distinct from that exhibited by Bernal-stacked bilayer graphene in the quantum Hall regime [6–13] and is now clearly established [3, 5, 14–32] as a prominent part of the physics of MATBG. In contrast to the quantum Hall case, in which eight Landau bands are filled sequentially to minimize the exchange energy, MATBG ground states appear [33] not to have any broken symmetries for a range of filling factors near CN, and in broken symmetry states to keep the filling factors of partially occupied flavors ν_f inside an interval $(-\nu_h^*, \nu_e^*)$, where ν_h^* and ν_e^* are maximum hole and electron filling factors. ($\nu_f \equiv (N_f - M)/M$ where N_f is the number of flat band electrons with flavor f and M is the number of moiré cells in the system; $\nu = \sum_f \nu_f$.)

In this Letter, we address some unusual aspects of the correlation physics of MATBG from the weak-coupling point of view (one shot GW approximation). We find that the average compressibility is dominated by Hartree interactions, that unbroken symmetry states are stable over a range of density near CN, and that the dependence of energy on flavor polarization is strongly overestimated by mean-field theory. Below we first explain the technical details of our calculations and then discuss the relationship of our findings to those obtained using other approaches to MATBG interaction physics.

Moiré-Band Weak-coupling Theory— The one shot GW approximation, also known as the random phase approximation (RPA), is a perturbative method that accounts for dynamic screening of long-range Coulomb interactions. It is commonly used [34, 35] in *ab initio* electronic structure theory to understand collective electronic behaviors, especially as probed by optical or photoemission spectroscopy. Although rigorously justified [36] only in weakly interacting systems, it has recently attracted interest [37] as a universal and accurate method for total energy calculations in many real materials, including [38] strongly correlated Mott insulators.

In this Letter we employ RPA theory to approximate the dependence of energy on the total band filling factor and on the partitioning of electrons between the four spin-valley flavors of MATBG. Because the number of electrons for each flavor is a good quantum number, we can approximate the magnetic energy landscape by adding exchange-correlation (xc) corrections $E_{\rm xc}$ to the self-consistent Hartree (SCH) energies of flavor polarized states. The RPA theory is motivated by the unusual property of MATBG, illustrated in Fig. 2 by plotting SCH bands at a series of band filling factors, that the band filling dependence of its total energy is dominated [39–43] by a Hartree mean-field contribution. The SCH energy increases rapidly as the flat bands are filled as shown in Fig. 1(c), and dominates the experimentally measured compressibilty. The RPA accounts both for this energy, and for dynamic fluctuation corrections to it.

The xc correction to the SCH energy can be expressed [44], in terms of a coupling-constant integral of the pair correlation function. This quantity can in turn be related to the density response function by

$$E_{\rm xc} = \frac{1}{2} \sum_{\mathbf{q},\mathbf{g}}{}' V_{\mathbf{q}+\mathbf{g}} \Big[-\frac{1}{\pi} \int_0^1 d\lambda \int_0^\infty d\omega \, \chi^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega; \lambda) - 1 \Big], \tag{1}$$

where $V_{\mathbf{q}} = 2\pi e^2/q\epsilon_{\rm BN}$ is the two-dimensional (2D) Coulomb interaction accounting for hexagonal boron nitride (hBN) dielectric screening with the dielectric constant chosen to be $\epsilon_{\rm BN} = 5.1$ throughout the paper, $\mathbf{q}$ is a wavevector in the moiré Brillouin zone (MBZ), $\mathbf{g}$ is a moiré reciprocal lattice vector, and the prime on the sum excludes the $\mathbf{q} = \mathbf{g} = 0$ term which

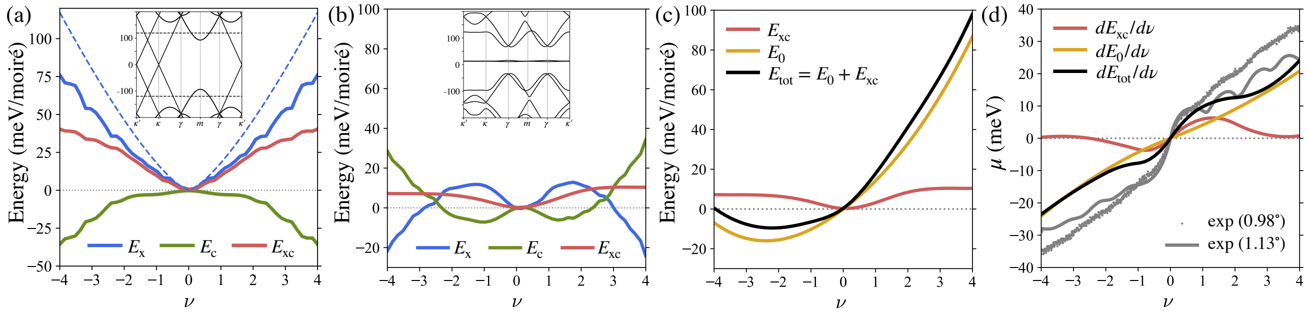


FIG. 1. Energies of paramagnetic states as a function of $\nu \in [-4, 4]$ for (a) a decoupled-bilayer and (b-d) 1.1°-TBG. (a-b) Exchange (E_x) and RPA correlation (E_c) energies as defined in Eqs. (4-5). The insets show the corresponding single-particle band structures. The black dashed lines in the inset of (a) mark the Fermi level for $\nu = \pm 4$. The blue dashed line in (a) is the exchange energy calculated using the approximate analytical expression Eq. (6). (c) The SCH energy E_0 [44] and the RPA total energy E_{tot}. (d) The calculated chemical potential $\mu = dE_{\mathrm{tot}}/d\nu$ with its zero shifted to the chemical potential at $\nu = 0$. The grey dots (0.98°) and the grey line (1.13°) plot measured chemical potentials from Ref.[19]. All energies are given relative to CN with the zero of energy at the neutral system Fermi level.

contributes only a gate-geometry-dependent constant [44]. In Eq. (1) $\chi^{\mathbf{gg}}$ is a diagonal matrix element of the density response function, which is a matrix in reciprocal lattice vectors because of system's discrete translational symmetry, and the frequency integration used to obtain equal time correlations has been rotated to the imaginary axis.

Equation (1) is formally exact. In RPA (time-dependent Hartree) we replace χ in Eq. (1) by

$$\begin{aligned}\chi(\lambda) &= \tilde{\chi}_{\mathrm{H}}(1-\lambda V\tilde{\chi}_{\mathrm{H}})^{-1}\\ &= \tilde{\chi}_{\mathrm{H}} + \lambda\tilde{\chi}_{\mathrm{H}} V\tilde{\chi}_{\mathrm{H}}(1-\lambda V\tilde{\chi}_{\mathrm{H}})^{-1},\end{aligned} \tag{2}$$

where $\tilde{\chi}_{\mathrm{H}}$ is the single-particle density response function calculated from the SCH bands [45], summing over independent contributions from all four flavors:

$$\tilde{\chi}_{\mathrm{H}} = \sum_{f=1}^{4} \tilde{\chi}_{\mathrm{H}}^{f}. \tag{3}$$

Possible improvements to this approximation are discussed later.

When inserted in Eq. (1), the second form for the right-hand-side of Eq. (2) separates the exchange energy E_{x}, the contribution that is first order in V, from the full fluctuation correction $E_{\mathrm{xc}} \equiv E_{\mathrm{x}} + E_{\mathrm{c}}$, allowing us to carefully account for its subtly convergent frequency integral. After integrating over λ, the exchange energy can be rewritten in the standard Slater determinant form [44]:

$$E_{\mathrm{x}} = -\frac{1}{2A}\sum_{\mathbf{q},\mathbf{g}}{}^{\prime} V_{\mathbf{q}+\mathbf{g}} \sum_{\substack{f,\mathbf{k},\alpha\\ \beta,\mathbf{g}_1,\mathbf{g}_2}} \left[\delta\bar{\rho}^{f}(\mathbf{k}) + 2\bar{\rho}^{0f}(\mathbf{k})\right]_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2} \delta\rho^{f}_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q}), \tag{4}$$

where $\delta\rho^f(\mathbf{k}) = \sum_n \left(\hat{z}_n(\mathbf{k})\hat{z}_n^{\dagger}(\mathbf{k})\Theta_{n\mathbf{k}} - \hat{z}_n^{0}(\mathbf{k})\hat{z}_n^{0\dagger}(\mathbf{k})\Theta^{0}_{n\mathbf{k}}\right)$ is the density matrix projected to flavor f relative to that of a charge neutral decoupled bilayer, $\delta\bar{\rho}$ is the complex conjugate of the corresponding matrix element of $\delta\rho$, $\hat{z}_n(\mathbf{k})$ and $\hat{z}_n^0(\mathbf{k})$ are plane-wave representation SCH and neutral-decoupled-bilayer quasiparticle eigenvectors, and $\Theta_{n\mathbf{k}}$ and $\Theta^0_{n\mathbf{k}}$ are the corresponding occupation numbers. In Eq. (4) $\mathbf{g}, \mathbf{g}_1, \mathbf{g}_2$ are moiré reciprocal lattice vectors, $\mathbf{k}$ and $\mathbf{q}$ are momenta in MBZ, α and β are layer and sublattice indices and A is the area of the 2D system. Because of their negative energy seas, continuum models of graphene multilayers are able to determine total energies only up to a reference energy (per area) that is a linear functions of electron density, $\varepsilon_{\mathrm{ref}} = \varepsilon_0 + \mu_0 n$; Eq. (4) chooses the zero of energy ε_0 to be the energy per area of neutral decoupled bilayers and the zero of chemical potential μ_0 to be the energy of states at the top of the decoupled bilayer valence band. The integration over the coupling-constant λ in Eq. (1) can be performed analytically to yield the correlation energy [44]

$$E_{\mathrm{c}} = \frac{1}{2\pi}\sum_{\mathbf{q}}{}^{\prime}\int_0^{\infty} d\omega \mathrm{Tr}\left[\sqrt{\mathbf{V}}\tilde{\chi}_{\mathrm{H}}\sqrt{\mathbf{V}} + \ln(1-\sqrt{\mathbf{V}}\tilde{\chi}_{\mathrm{H}}\sqrt{\mathbf{V}})\right], \tag{5}$$

where $\mathbf{V}$ and $\tilde{\chi}_{\mathrm{H}}$ are matrices in reciprocal lattice vector with implicit $\mathbf{q}$ and ω dependences. The correlation energy must be regularized by subtracting its value in unbroken symmetry states at CN; its contribution to the chemical potential at CN is close to zero because the models we study have approximate particle-hole symmetry.

Paramagnetic State Energy — We interpret our numerical results for the band filling ν dependence of the MATBG paramagnetic ground state energy (Fig. 1(b,c)) by comparing them

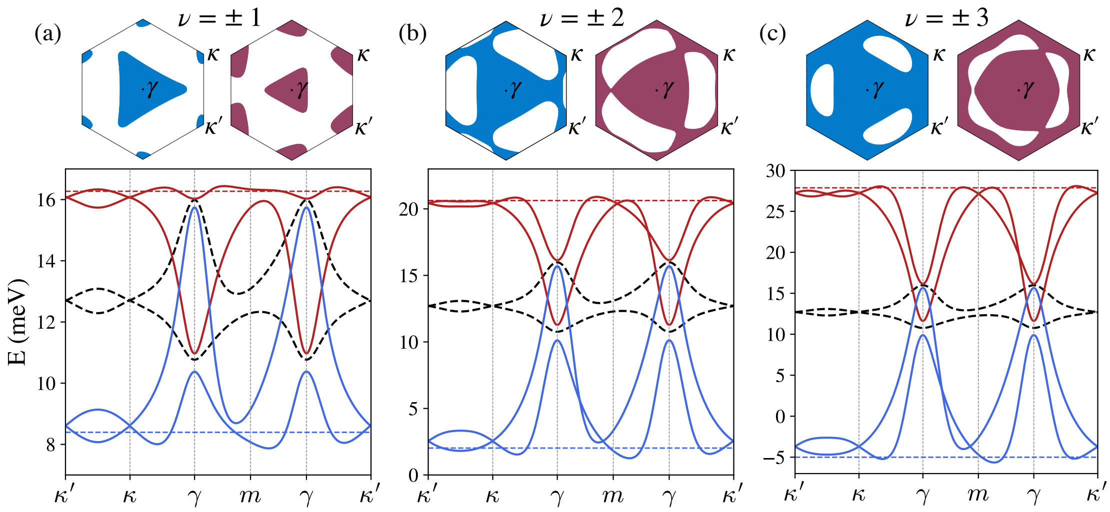


FIG. 2. The SCH paramagnetic state bands (colored lines) and the corresponding Fermi surfaces (shaded areas) at a series of ν values on hole-doped (blue) and electron-doped (red) sides. The black dashed line in each spectrum is the single-particle band structure and the colored dashed horizontal lines mark Fermi levels. At $\nu = -3$, the flat valence band is 1/4 full and the occupied states are those whose charge density is most peaked near minima of the external potential produced by remote band charges. At $\nu = -1$, the flat valence band is at 3/4 filling. Holes in the valence band remain near γ, which would be the valence band bottom if Hartree corrections were not included. Holes near γ are finally filled only around $\nu = -0.3$ (see Fig. 5 in SM III) as ν approaches zero and Hartree energies finally become small compared to band energies. The Fermi surfaces at filling factors $+\nu$ and $-\nu$ (for example $\nu = 2$ and $\nu = -2$) would be identical for any ν if the model had exact particle-hole symmetry. At filling factors away from $\nu = 0$, the SCH band width is dominated by the Hartree mean-field contribution.

with results for the decoupled bilayer [46] (Fig. 1(a)) calculated in exactly the same way. In both cases the exchange energy is positive at small $|\nu|$ because of [46] rapid changes in Bloch state spinors near the Dirac point. The blue dashed line in Fig. 1(a) is the exchange energy of an eight-Dirac-cone model [46]:

$$E_{\rm x}^{\rm D} = \frac{\alpha\hbar c}{24\pi}\frac{g}{\epsilon_{\rm BN}} k_{\rm F}^{3} \ln\left(\frac{k_{\rm c}}{k_{\rm F}}\right) + \text{regular terms}, \qquad (6)$$

where $g = 8$ and $k_{\rm F} = (4\pi n/g)^{1/2}$. The exchange energy of MATBG is smaller than that of decoupled bilayers because of the dominant role of the Hartree potential in shaping occupied band states wavefunctions. In contrast to the decoupled bilayer case, MATBG correlation energies are low near CN, because that is where the phase space for low-energy particle-hole excitations within the flat band octet is the largest. The correlation energy is highest near $|\nu| = 4$ because the gaps between flat and remote bands suppress fluctuations. In our calculations there is a small particle-hole asymmetry in all properties, including the exchange and correlation energies, because we include non-local interlayer tunneling corrections [47] to the Bistritzer-MacDonald (BM) MATBG model [1, 44].

Because of the partial cancellation between exchange and correlation effects, discussed again below in connection with flavor ferromagnetism, the difference between MATBG and decoupled bilayers is dominated by the SCH energy [44] plotted in Fig. 1(c). The SCH energy is calculated relative to its value at CN, and its slope at CN is finite because the bare flat bands are centered around $\varepsilon_{\rm fb} \approx 12$ meV (see Fig. 2) in the non-local BM model we employ. The chemical potential μ, the energy to add a single-electron increases steadily as the flat bands are filled mainly because of Hartree effects. We find that the chemical potential difference between full and empty flat bands is ~ 50 meV. When the bands are nearly empty, added electrons occupy regions in the moiré unit cell in which the mean-field potential from remote band electrons is most attractive. When the bands are nearly filled, on the other hand, it follows from approximate particle-hole symmetry that electrons occupy the same region but the Hartree mean-field potential is now repulsive.

In Fig. 1(d) we compare our results for the filling factor dependence of the chemical potential across the full range of flat band filling with experimental results published in Ref. [19]. The total shift in chemical potential is somewhat larger in experiment than in theory. Since the states near the full and empty flat band limit are not expected to be strongly correlated, we attribute this small discrepancy to weak mixing between flat and remote bands and small inaccuracies in the continuum model we employ. The most striking feature of these results is shared between theory and experiment, namely that the chemical potential increases approximately linearly with band filling factor [19, 20, 48, 49]. In MATBG experiments, structures do emerge at some filling factors that are thought to be due to first order flavor-symmetry breaking phase transitions at low temperatures, which we now address, and at higher temperatures to surviving local moment fluctuations [50].

Flat Band Flavor Ferromagnetism — The RPA energy calculation can be carried out for any set of flavor-dependent filling factors. Typical numerical results [51] are summarized in Fig. 3. The $\nu = 0$ polarized states in Fig. 3(c) have filling factor p for two flavors and filling factor $-p$ for the other two

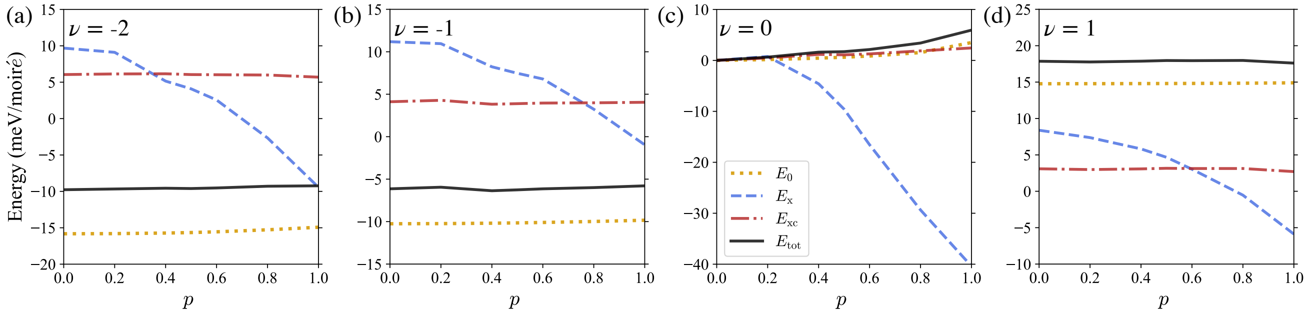


FIG. 3. SCH energy E_0 (yellow dotted lines), exchange energy $E_{\rm x}$ (blue dashed lines), exchange-correlation energy $E_{\rm xc}$ (red dash-dotted lines) and RPA total energy $E_{\rm tot}$ (black solid lines) as a function of polarization p at (a) $\nu = -2$, (b) $\nu = -1$, (c) $\nu = 0$ and (d) $\nu = 1$. p characterizes the degree of flavor polarization as explained in the main text. $p = 0$ corresponds to the paramagnetic state and $p = 1$ corresponds to full flavor polarization.

flavors. Increasing p shifts states from the valence bands of two flavors to the conduction bands of the other two flavors. Because of MATBG's approximate particle-hole symmetry, this polarization path does not strongly influence the charge density, which remains approximately uniform at this filling factor for all values of p, as illustrated in Figs. 13,14. The main point to notice is that fully polarized states are strongly favored by exchange energies, but this energy gain is almost perfectly cancelled by the correlation energy which strongly favors states in which each flavor is half filled. Similar results are obtained at other filling factors. The family of polarized states at $\nu = -2$ in Fig. 3(a) have filling factor $-(1+p)/2$ for two flavors and filling factor $-(1-p)/2$ for the other two flavors; increasing p shifts electrons between valence bands with different flavors and the charge density is non-uniform at all values of p. For $\nu = \pm 1$, the flavor polarization path illustrated in Fig. 3(b,d) is $\nu = \pm(1+3p)/4$ for one flavor and $\nu = \pm(1-p)/4$ for the remaining three flavors. The exchange energy gain upon polarization is again almost exactly cancelled by correlation, underscoring the dominance of the SCH energy. Once correlations are included the dependence of the SCH energy on p, which was judged to be insignificant in previous self-consistent Hartree-Fock [52] calculations, retains a role in the energy competition among different polarized states.

Within the RPA theory the cancellation between exchange and correlation for the polarization p dependence of the energy can be understood in terms of Eqs. (1-2). The p dependence of energy follows from that of $\tilde{\chi}_{\rm H}$, and this lies mainly in the range of low-frequency fluctuations within the flat band where the important matrix elements of $V\tilde{\chi}_{\rm H}$ are much larger than 1 so that $\chi(\lambda) \to V^{-1}$ (perfect screening), and the dependence of $E_{\rm xc}$ on polarization is lost. Physically, correlations are already strong even in the paramagnetic state and there is little left to gain by flavor ordering. Generally speaking, we find that the tendency toward flavor symmetry breaking is stronger at larger $|\nu|$ and stronger at positive ν than at negative ν, as summarized in Fig. 10, in agreement with most experiments [16, 19, 48, 53]. In addition we find that the difference in energy between polarized and paramagnetic states is drastically reduced by correlations from ~ 40 meV per moiré period to less than ~ 3 meV (Fig. 3(c)).

In MATBG broken C_2T symmetry opens up a gap between the conduction and valence bands. This type of broken symmetry within flavors is therefore common in mean-field calculations. In our RPA calculations we find, as summarized in Table II, that when C_2T symmetry is broken by adding a sublattice-dependent potential of the type produced by aligned hBN substrates, flavor ferromagnetism is favored at almost all filling factors including those proximate to CN. This finding aligns well with experimental evidence suggesting that hBN alignment tends to favor states with broken symmetries [14, 15, 54], including quantum anomalous Hall states at fractional flat band fillings [14].

The Magic-Angle Correlation Problem— In this Letter we have reported on the first RPA calculation for MATBG. The RPA weak-coupling approach has the advantage that it accounts for dynamic screening of long-range Coulomb interactions, but is less reliable than some other methods in accounting for short-distance correlations. Competing methods often require tight-binding models, which in the case of MATBG have the disadvantage that they require the introduction of additional bands [55] to compensate for fragile topology inherited from the isolated layer Dirac cones. Our theory establishes the crucial influence of correlations in compressible metallic states in expanding unbroken symmetry regions in the MATBG phase diagram. The RPA weak-coupling approach is also relevant for other moiré materials that exhibit strong correlations.

Our calculations include 146 remote valence and conduction bands per spin and flavor. Our calclations are consistent with experimental indications that flavor ferromagnetism is common in both insulating and metallic states when the MATBG flat bands are partially filled, less likely close to CN, and more likely at positive filling factors than at negative filling factors. The exchange energy gains that favor broken symmetry insulating ground states at integer ν, are comparable in size to correlation energy gains in closely competing metallic states with fewer or no broken symmetries. The resulting weak dependence of energy on magnetic state is consistent with small collective excitation energies of insulating states [56, 57] and with strong coupling approaches [22, 58] that can be applied close to integer band fillings. Our calculations demonstrate that [44] fluctuations in remote bands do not generally play a central role in MATBG properties except in the cases of nearly empty and nearly full bands. This finding justifies the flat-band projection that is required to make non-perturbative finite-size numerical calculations [59–62] feasible. Perturbative calculations are approximate, but have the advantage that finite-size effects can be eliminated by taking dense momentum space grids; our calculations employ 432 k-points in the MBZ.

Our calculation results can be compared directly to experimental results for the chemical potential μ, which increases by ~ 50 meV as the flat bands are filled. This compares to a dependence of energy on flavor polarization that is typically ~ 3 meV per moiré cell. The positive compressibility we find, in agreement with experiment, for MATBG electrons contrasts with the well-known negative compressibility of strongly interacting two-dimensional electron gas systems [63, 64], and is associated with unusual properties of the projected flat-band Hilbert space. In MATBG models with exact particle-hole symmetry, the flat conduction and valence bands at the Fermi energy spatial structure within the moiré unit cell that precisely complements the total density of remote occupied bands, so that the total density is uniform. The increase in chemical potential with filling factor is associated with the property that the non-uniform density of the remote bands is first eliminated and then restored with the opposite sign as the flat bands are filled. We emphasize that unlike most calculations in the literature, which overstate dielectric screening to suppress interaction scales, all our results are obtained using a physically realistic hBN dielectric constant $\epsilon_{\text{BN}} = 5.1$ [44].

The MATBG correlation problem is extraordinarily challenging and the RPA theory, like other approaches, has limitations. Even though the flat band eigenstates have weak dispersion, their wavefunctions vary in a complex way across the MBZ. For this reason there is no simple Hubbard-like lattice model representation of the correlation problem. Aside from the fascinating low-temperature superconducting instability, two key higher energy issues still do not have definitive answers. i) What is the ground state at CN? Is it the $p = 0$ state of Fig. 3, which has no broken symmetries and strong correlations, or the $p = 1$ state, which is a single Slater determinant with analytically calculable excitations when remote band fluctuations are neglected? ii) What is the Fermi surface in the range of filling factors surrounding $\nu = 0$? Is it the γ centered Fermi surface of the $p = 1$ state or the κ, κ' centered Fermi surface of the $p = 0$ state? In either case how does the Fermi surface, at least as indicated by weak-field Hall measurements [47], manage to avoid Liftshitz transitions over such a broad range of filling factors $-1.8 \lesssim \nu \lesssim 0.9$ surrounding $\nu = 0$? For the first question we do not consider the weak-coupling answer (that $p = 0$ is favored) to be definitive, but it certainly demonstrates that the two states are competitive. The second question is especially troublesome if one imagines that the ground state near $\nu = 0$ is a doped $p = 1$ state in which the band degeneracies have been reduced from four to two and Fermi surface areas must be correspondingly larger. The more likely option, in our view, is that the ground state near CN is an unpolarized state as predicted by RPA. Part of the motivation for this view is the absence of finite-temperature anomalies in experiment, which would signal a phase transitions to a paramagnetic state — expected to be at least weakly first order in MATBG as in other itinerant electron magnets [65]. If so, there is no hint experimentally of the emergence between $\nu = 0$ and $|\nu| = 1$ of the self-consistent Hartree multi-pocket Fermi surface topology illustrated in Fig. 2. Future work should explain why this pocket does not appear (or alternately why its appearance does not influence transport), perhaps due to a refinement of the single-particle model which changes flat band wavefunctions [66–69], exchange interactions within the doped flat bands that stabilize κ, κ' centered surfaces, broken C_2T symmetry related to chiral model physics [22, 56, 70, 71] and intervalley exchange interactions that we have neglected [71–73]. Systematic studies of the evolution of MATBG properties with gate induced interlayer displacement fields could play a role in sorting this confusing landscape.

Acknowledgements— Work in Austin was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award DE-SC0019481. M.P. is supported by the European Union's Horizon 2020 research and innovation program under the grant agreement No. 881603 - GrapheneCore3 and the Marie Sklodowska-Curie grant agreement No. 873028 and by the MUR - Italian Minister of University and Research under the "Research projects of relevant national interest - PRIN 2020" - Project No. 2020JLZ52N, title "Light-matter interactions and the collective behavior of quantum 2D materials (q-LIMA)". The authors acknowledge resources provided by the Texas Advanced Computing Center (TACC) at The University of Texas at Austin that have contributed to the research results reported in this paper.

[1] R. Bistritzer and A. H. MacDonald, *Moiré bands in twisted double-layer graphene,* Proceedings of the National Academy of Sciences **108**, 12233 (2011).

[2] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Unconventional superconductivity in magic-angle graphene superlattices,* Nature **556**, 43 (2018).

[3] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, *Correlated insulator behaviour at half-filling in magic-angle graphene superlattices,* Nature **556**, 80 (2018).

[4] E. Y. Andrei and A. H. MacDonald, *Graphene bilayers with a twist,* Nature Materials **19**, 1265 (2020).

[5] L. Balents, C. R. Dean, D. K. Efetov, and A. F. Young, *Superconductivity and strong correlations in moiré flat bands,* Nature Physics **16**, 725 (2020).

[6] E. McCann and V. I. Fal'ko, *Landau-level degeneracy and quantum Hall effect in a graphite bilayer,* Phys. Rev. Lett. **96**, 086805 (2006).

[7] Y. Barlas, R. Côté, K. Nomura, and A. H. MacDonald, *Intra-landau-level cyclotron resonance in bilayer graphene,* Phys.

Rev. Lett. **101**, 097601 (2008).

[8] R. Côté, J. Lambert, Y. Barlas, and A. H. MacDonald, *Orbital order in bilayer graphene at filling factor $\nu = -1$,* Phys. Rev. B **82**, 035445 (2010).

[9] Y. Barlas, K. Yang, and A. H. MacDonald, *Quantum Hall effects in graphene-based two-dimensional electron systems,* Nanotechnology **23**, 052001 (2012).

[10] B. E. Feldman, J. Martin, and A. Yacoby, *Broken-symmetry states and divergent resistance in suspended bilayer graphene,* Nature Physics **5**, 889 (2009).

[11] J. Martin, B. E. Feldman, R. T. Weitz, M. T. Allen, and A. Yacoby, *Local compressibility measurements of correlated states in suspended bilayer graphene,* Phys. Rev. Lett. **105**, 256806 (2010).

[12] R. T. Weitz, M. T. Allen, B. E. Feldman, J. Martin, and A. Yacoby, *Broken-symmetry states in doubly gated suspended bilayer graphene,* Science **330**, 812 (2010).

[13] J. Jung, F. Zhang, and A. H. MacDonald, *Lattice theory of pseudospin ferromagnetism in bilayer graphene: Competing interaction-induced quantum Hall states,* Phys. Rev. B **83**, 115408 (2011).

[14] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, *Intrinsic quantized anomalous Hall effect in a moiré heterostructure,* Science **367**, 900 (2020).

[15] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, *Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene,* Science **365**, 605 (2019).

[16] P. Stepanov, M. Xie, T. Taniguchi, K. Watanabe, X. Lu, A. H. MacDonald, B. A. Bernevig, and D. K. Efetov, *Competing zero-field Chern insulators in superconducting twisted bilayer graphene,* Phys. Rev. Lett. **127**, 197701 (2021).

[17] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, *Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene,* Nature **574**, 653 (2019).

[18] H. Polshyn, J. Zhu, M. A. Kumar, Y. Zhang, F. Yang, C. L. Tschirhart, M. Serlin, K. Watanabe, T. Taniguchi, A. H. MacDonald, and A. F. Young, *Electrical switching of magnetic order in an orbital Chern insulator,* Nature **588**, 66 (2020).

[19] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, *Cascade of phase transitions and dirac revivals in magic-angle graphene,* Nature **582**, 203 (2020).

[20] D. Wong, K. P. Nuckolls, M. Oh, B. Lian, Y. Xie, S. Jeon, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, *Cascade of electronic transitions in magic-angle twisted bilayer graphene,* Nature **582**, 198 (2020).

[21] I. Das, X. Lu, J. Herzog-Arbeitman, Z.-D. Song, K. Watanabe, T. Taniguchi, B. A. Bernevig, and D. K. Efetov, *Symmetry-broken Chern insulators and rashba-like landau-level crossings in magic-angle bilayer graphene,* Nature Physics **17**, 710 (2021).

[22] E. Khalaf, S. Chatterjee, N. Bultinck, M. P. Zaletel, and A. Vishwanath, *Charged skyrmions and topological origin of superconductivity in magic-angle graphene,* Science Advances **7**, eabf5299 (2021).

[23] Y. Saito, F. Yang, J. Ge, X. Liu, T. Taniguchi, K. Watanabe, J. I. A. Li, E. Berg, and A. F. Young, *Isospin pomeranchuk effect in twisted bilayer graphene,* Nature **592**, 220 (2021).

[24] Y. Xie, A. T. Pierce, J. M. Park, D. E. Parker, E. Khalaf, P. Ledwith, Y. Cao, S. H. Lee, S. Chen, P. R. Forrester, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, *Fractional Chern insulators in magic-angle twisted bilayer graphene,* Nature **600**, 439 (2021).

[25] S. Wu, Z. Zhang, K. Watanabe, T. Taniguchi, and E. Y. Andrei, *Chern insulators, van Hove singularities and topological flat bands in magic-angle twisted bilayer graphene,* Nature materials **20**, 488 (2021).

[26] K. P. Nuckolls, M. Oh, D. Wong, B. Lian, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, *Strongly correlated Chern insulators in magic-angle twisted bilayer graphene,* Nature **588**, 610 (2020).

[27] J. Kang and O. Vafek, *Strong coupling phases of partially filled twisted bilayer graphene narrow bands,* Phys. Rev. Lett. **122**, 246401 (2019).

[28] B. Lian, Z.-D. Song, N. Regnault, D. K. Efetov, A. Yazdani, and B. A. Bernevig, *Twisted bilayer graphene. IV. exact insulator ground states and phase diagram,* Phys. Rev. B **103**, 205414 (2021).

[29] X. Liu, G. Farahi, C.-L. Chiu, Z. Papic, K. Watanabe, T. Taniguchi, M. P. Zaletel, and A. Yazdani, *Visualizing broken symmetry and topological defects in a quantum Hall ferromagnet,* Science **375**, 321 (2022).

[30] L. Wang, E.-M. Shih, A. Ghiotto, L. Xian, D. A. Rhodes, C. Tan, M. Claassen, D. M. Kennes, Y. Bai, B. Kim, K. Watanabe, T. Taniguchi, X. Zhu, J. Hone, A. Rubio, A. N. Pasupathy, and C. R. Dean, *Correlated electronic phases in twisted bilayer transition metal dichalcogenides,* Nature materials **19**, 861 (2020).

[31] T. Wang, D. E. Parker, T. Soejima, J. Hauschild, S. Anand, N. Bultinck, and M. P. Zaletel, *Ground-state order in magic-angle graphene at filling $\nu = -3$: A full-scale density matrix renormalization group study,* Phys. Rev. B **108**, 235128 (2023).

[32] J. Zhu, J.-J. Su, and A. H. MacDonald, *Voltage-controlled magnetic reversal in orbital Chern insulators,* Phys. Rev. Lett. **125**, 227702 (2020).

[33] These statements are based on a minimal interpretation of measurements of electronic compressibility [19], tunneling spectroscopy [20] and weak-field Hall effects [47]. Other scenarios in which flavor symmetries are broken even at CN cannot be ruled out at present and are assumed in some theoretical work. Note that the cascades observed in the high-temperature compressibility experiment [23] (up to 100 K) are not indicative of ground state phase transitions. These high-temperature cascades may result from fluctuations of local moments, which can be explained by localized orbitals and the heavy fermion model [58, 74].

[34] D. Golze, M. Dvorak, and P. Rinke, *The GW compendium: A practical guide to theoretical photoemission spectroscopy,* Frontiers in Chemistry **7** (2019), 10.3389/fchem.2019.00377.

[35] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, *Berkeleygw: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures,* Computer Physics Communications **183**, 1269 (2012).

[36] M. Gell-Mann and K. A. Brueckner, *Correlation energy of an electron gas at high density,* Phys. Rev. **106**, 364 (1957).

[37] T. Olsen and K. S. Thygesen, *Random phase approximation applied to solids, molecules, and graphene-metal interfaces: From Van der Waals to covalent bonding,* Phys. Rev. B **87**, 075111 (2013).

[38] T. Olsen, *Assessing the performance of the random phase approximation for exchange and superexchange coupling constants in magnetic crystalline solids,* Phys. Rev. B **96**, 125143 (2017).

[39] M. Xie and A. H. MacDonald, *Nature of the correlated insulator*

states in twisted bilayer graphene, Phys. Rev. Lett. **124**, 097601 (2020).

[40] F. Guinea and N. R. Walet, *Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers,* Proceedings of the National Academy of Sciences **115**, 13174 (2018).

[41] T. Cea, N. R. Walet, and F. Guinea, *Electronic band structure and pinning of fermi energy to van Hove singularities in twisted bilayer graphene: A self-consistent approach,* Phys. Rev. B **100**, 205113 (2019).

[42] L. Rademaker, D. A. Abanin, and P. Mellado, *Charge smoothening and band flattening due to Hartree corrections in twisted bilayer graphene,* Phys. Rev. B **100**, 205114 (2019).

[43] T. Cea and F. Guinea, *Band structure and insulating states driven by coulomb interaction in twisted bilayer graphene,* Phys. Rev. B **102**, 045107 (2020).

[44] See Supplemental Material (SM) at XXX for the derivation of Eq. (1) in SM II; the derivations of Eqs. (4-5) in SM VII; the definition of SCH energy E_0 in SM III; the expression of $\rho^{0f}(\mathbf{k})$ in Eq. (4) in SM VIII; details of MATBG model parameters that we use in this article in SM I; a detailed discussion on fluctuations in remote bands in SM IX; Fig. S12 for the energy dependence on dielectric constant $\epsilon_{\rm BN}$. The SM also contains Refs. [1, 22, 47, 66–69, 75–85].

[45] P. Novelli, I. Torre, F. H. L. Koppens, F. Taddei, and M. Polini, *Optical and plasmonic properties of twisted bilayer graphene: Impact of interlayer tunneling asymmetry and ground-state charge inhomogeneity,* Phys. Rev. B **102**, 125403 (2020).

[46] Y. Barlas, T. Pereg-Barnea, M. Polini, R. Asgari, and A. H. MacDonald, *Chirality and correlations in graphene,* Phys. Rev. Lett. **98**, 236601 (2007).

[47] M. Xie and A. H. MacDonald, *Weak-field Hall resistivity and spin-valley flavor symmetry breaking in magic-angle twisted bilayer graphene,* Phys. Rev. Lett. **127**, 196401 (2021).

[48] S. L. Tomarken, Y. Cao, A. Demir, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and R. C. Ashoori, *Electronic compressibility of magic-angle graphene superlattices,* Phys. Rev. Lett. **123**, 046601 (2019).

[49] T. Cea, P. A. Pantaleón, N. R. Walet, and F. Guinea, *Electrostatic interactions in twisted bilayer graphene,* Nano Materials Science **4**, 27 (2022).

[50] One important limitation of the weak-coupling approach is that it does not capture the influence of unordered fluctuating local moments.

[51] A more complete set of numerical results for a wide variety of filling factors and flavor polarizations are gathered in SM X Table I.

[52] S. Liu, E. Khalaf, J. Y. Lee, and A. Vishwanath, *Nematic topological semimetal and insulator in magic-angle bilayer graphene at charge neutrality,* Phys. Rev. Res. **3**, 013033 (2021).

[53] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, *Tuning superconductivity in twisted bilayer graphene,* Science **363**, 1059 (2019).

[54] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, *Evidence of orbital ferromagnetism in twisted bilayer graphene aligned to hexagonal boron nitride,* Nano letters **21**, 4299 (2021).

[55] H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, *Faithful tight-binding models and fragile topology of magic-angle bilayer graphene,* Phys. Rev. B **99**, 195455 (2019).

[56] E. Khalaf, N. Bultinck, A. Vishwanath, and M. P. Zaletel, *Soft modes in magic angle twisted bilayer graphene,* arXiv:2009.14827 (2020).

[57] A. Kumar, M. Xie, and A. H. MacDonald, *Lattice collective modes from a continuum model of magic-angle twisted bilayer graphene,* Phys. Rev. B **104**, 035119 (2021).

[58] Z.-D. Song and B. A. Bernevig, *Magic-angle twisted bilayer graphene as a topological heavy fermion problem,* Phys. Rev. Lett. **129**, 047601 (2022).

[59] P. Potasz, M. Xie, and A. H. MacDonald, *Exact diagonalization for magic-angle twisted bilayer graphene,* Phys. Rev. Lett. **127**, 147203 (2021).

[60] C. Repellin, Z. Dong, Y.-H. Zhang, and T. Senthil, *Ferromagnetism in narrow bands of moiré superlattices,* Phys. Rev. Lett. **124**, 187601 (2020).

[61] F. Xie, A. Cowsik, Z.-D. Song, B. Lian, B. A. Bernevig, and N. Regnault, *Twisted bilayer graphene. VI. an exact diagonalization study at nonzero integer filling,* Phys. Rev. B **103**, 205416 (2021).

[62] T. Soejima, D. E. Parker, N. Bultinck, J. Hauschild, and M. P. Zaletel, *Efficient simulation of moiré materials using the density matrix renormalization group,* Phys. Rev. B **102**, 205111 (2020).

[63] J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, *Negative compressibility of interacting two-dimensional electron and quasiparticle gases,* Phys. Rev. Lett. **68**, 674 (1992).

[64] G. L. Yu, R. Jalil, B. Belle, A. S. Mayorov, P. Blake, F. Schedin, S. V. Morozov, L. A. Ponomarenko, F. Chiappini, S. Wiedmann, U. Zeitler, M. I. Katsnelson, A. K. Geim, K. S. Novoselov, and D. C. Elias, *Interaction phenomena in graphene seen through quantum capacitance,* Proceedings of the National Academy of Sciences **110**, 3282 (2013).

[65] M. Brando, D. Belitz, F. M. Grosche, and T. R. Kirkpatrick, *Metallic quantum ferromagnets,* Rev. Mod. Phys. **88**, 025006 (2016).

[66] S. Fang, S. Carr, Z. Zhu, D. Massatt, and E. Kaxiras, *Angle-dependent ab initio low-energy Hamiltonians for a relaxed twisted bilayer graphene heterostructure,* arXiv:1908.00058 (2019).

[67] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, *Exact continuum model for low-energy electronic states of twisted bilayer graphene,* Phys. Rev. Res. **1**, 013001 (2019).

[68] O. Vafek and J. Kang, *Continuum effective Hamiltonian for graphene bilayers for an arbitrary smooth lattice deformation from microscopic theories,* Phys. Rev. B **107**, 075123 (2023).

[69] J. Kang and O. Vafek, *Pseudomagnetic fields, particle-hole asymmetry, and microscopic effective continuum Hamiltonians of twisted bilayer graphene,* Phys. Rev. B **107**, 075408 (2023).

[70] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, *Origin of magic angles in twisted bilayer graphene,* Phys. Rev. Lett. **122**, 106405 (2019).

[71] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, *Ground state and hidden symmetry of magic-angle graphene at even integer filling,* Phys. Rev. X **10**, 031034 (2020).

[72] E. Khalaf, P. Ledwith, and A. Vishwanath, *Symmetry constraints on superconductivity in twisted bilayer graphene: Fractional vortices, 4e condensates, or nonunitary pairing,* Phys. Rev. B **105**, 224508 (2022).

[73] M. Kashiwagi, T. Taen, K. Uchida, K. Watanabe, T. Taniguchi, and T. Osada, *Weak localization on moiré superlattice in twisted double bilayer graphene,* Jpn. J. Appl. Phys. **61**, 100907 (2022).

[74] J. Kang, B. A. Bernevig, and O. Vafek, *Cascades between light and heavy fermions in the normal state of magic-angle twisted bilayer graphene,* Phys. Rev. Lett. **127**, 266402 (2021).

[75] R. Geick, C. H. Perry, and G. Rupprecht, *Normal modes in hexagonal boron nitride,* Phys. Rev. **146**, 543 (1966).

[76] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (2005).

[77] A. Görling, *Hierarchies of methods towards the exact kohn-sham correlation energy based on the adiabatic-connection fluctuation-dissipation theorem,* Phys. Rev. B **99**, 235120 (2019).

[78] Y. H. Kwan, G. Wagner, T. Soejima, M. P. Zaletel, S. H. Simon, S. A. Parameswaran, and N. Bultinck, *Kekulé spiral order at all nonzero integer fillings in twisted bilayer graphene,* Phys. Rev. X **11**, 041063 (2021).

[79] P. J. Ledwith, E. Khalaf, and A. Vishwanath, *Strong coupling theory of magic-angle graphene: A pedagogical introduction,* Annals of Physics **435**, 168646 (2021).

[80] G. Shavit, E. Berg, A. Stern, and Y. Oreg, *Theory of correlated insulators and superconductivity in twisted bilayer graphene,* Phys. Rev. Lett. **127**, 247703 (2021).

[81] G. Shavit, K. c. v. Kolář, C. Mora, F. von Oppen, and Y. Oreg, *Strain disorder and gapless intervalley coherent phase in twisted bilayer graphene,* Phys. Rev. B **107**, L081403 (2023).

[82] M. Oh, K. P. Nuckolls, D. Wong, R. L. Lee, X. Liu, K. Watanabe, T. Taniguchi, and A. Yazdani, *Evidence for unconventional superconductivity in twisted bilayer graphene,* Nature **600**, 240 (2021).

[83] J. P. Hong, T. Soejima, and M. P. Zaletel, *Detecting symmetry breaking in magic angle graphene using scanning tunneling microscopy,* Phys. Rev. Lett. **129**, 147001 (2022).

[84] J. Yu, B. A. Foutty, Y. H. Kwan, M. E. Barber, K. Watanabe, T. Taniguchi, Z.-X. Shen, S. A. Parameswaran, and B. E. Feldman, *Spin skyrmion gaps as signatures of intervalley-coherent insulators in magic-angle twisted bilayer graphene,* Nature Communications **14**, 6679 (2023).

[85] K. P. Nuckolls, R. L. Lee, M. Oh, D. Wong, T. Soejima, J. P. Hong, D. Călugăru, J. Herzog-Arbeitman, B. A. Bernevig, K. Watanabe, T. Taniguchi, N. Regnault, M. P. Zaletel, and A. Yazdani, *Quantum textures of the many-body wavefunctions in magic-angle graphene,* Nature **620**, 525 (2023).

[86] See e.g. Ref.[77] and references therein to earlier work.

Supplemental Material:
Weak Coupling Theory of Competing Phases in Magic-Angle Twisted Bilayer Graphene

I. BISTRITZER-MACDONALD MODEL WITH NON-LOCAL INTERLAYER TUNNELING

In the original Bistritzer-MacDonald (BM) model, Eq. (5) in Ref. [1], the interlayer tunneling matrix element is

$$\begin{aligned} T^{\alpha\beta}_{\mathbf{k}\mathbf{k}'} &= \frac{1}{A_{uc}}\sum_{j=1}^{3} t_{\mathbf{k}+\mathbf{G}_j} e^{i[\mathbf{G}_j\cdot(\tau_\alpha-\tau_\beta+\tau_0)-\mathbf{G}'_j\cdot\mathbf{d}]}\delta_{\mathbf{k}+\mathbf{G}_j,\mathbf{k}'+\mathbf{G}'_j} \\ &= \frac{1}{A_{uc}}\sum_{j=1}^{3} t_{\mathbf{k}+\mathbf{G}_j} T^{\alpha\beta}_j \delta_{\mathbf{k}+\mathbf{G}_j,\mathbf{k}'+\mathbf{G}'_j}, \end{aligned} \tag{7}$$

where $\mathbf{G}$ and $\mathbf{G}'$ are reciprocal lattice vectors of the top (layer 1) and bottom (layer 2) graphene layers respectively, momenta $\mathbf{k}$ and $\mathbf{k}'$ are measured relative to the Γ point of the original monolayer graphene and are both near the Dirac point $\mathbf{K}$. Greek indices label sublattices. The interlayer tunneling t and matrices T_j are defined later. The top layer (layer 1) is anticlockwise rotated by $\theta/2$ and characterized by vectors without $'$ in the notation, bottom layer (layer 2) is clockwise rotated by $\theta/2$ and characterized by vector with $'$. The reciprocal lattice vectors are related by the rotation operator $\mathbf{G}'_i = \mathcal{R}_{-\theta}\mathbf{G}_i$, with

$$\mathcal{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{8}$$

Three $\mathbf{G}$'s that are most relevant to the interlayer hopping in the two-center approximation are, as shown in Fig. 4(a),

$$\mathbf{G}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{G}_2 = \mathcal{R}_{\theta/2}\cdot\frac{4\pi}{\sqrt{3}a}\begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \mathbf{G}_3 = \mathcal{R}_{\theta/2}\cdot\frac{4\pi}{\sqrt{3}a}\begin{pmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}. \tag{9}$$

$a = 0.246$ nm is graphene's lattice constant. Starting from AB-stacked bilayer graphene, we choose

$$\tau_A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \tau_B = \tau_0 = \mathcal{R}_{\theta/2}\cdot\frac{a}{\sqrt{3}}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{10}$$

Three largest interlayer hopping terms, corresponding to three $\mathbf{G}$'s in Eq. (9), are

$$T_1 = \begin{pmatrix} u & 1 \\ 1 & u \end{pmatrix}, \; T_2 = e^{-i\mathbf{G}'_2\cdot\mathbf{d}}\begin{pmatrix} ue^{i\phi} & 1 \\ e^{-i\phi} & ue^{i\phi} \end{pmatrix}, \; T_3 = e^{-i\mathbf{G}'_3\cdot\mathbf{d}}\begin{pmatrix} ue^{-i\phi} & 1 \\ e^{i\phi} & ue^{-i\phi} \end{pmatrix}, \tag{11}$$

where $\phi = 2\pi/3$, $u = w_{AA}/w_{AB}$ is the ratio of interlayer tunneling between the same sublattice and between different sublattices. The phase factor dependent on $\mathbf{d}$ can be eliminated by a unitary transformation of the plane-wave expanded basis.

The Fourier transform of interlayer tunneling strength $t(\mathbf{r})$ in Eq. (7) is

$$t_{\mathbf{q}} = \int d^2\mathbf{r}e^{i\mathbf{q}\cdot\mathbf{r}}t(\mathbf{r}). \tag{12}$$

Because $t(\mathbf{r})$ is a smooth and slowly varying function of the projected 2D coordinate $\mathbf{r}$, $t_{\mathbf{q}}$ rapidly decays with respect to q and we only keep three terms in the interlayer tunneling Eq. (7). Because of the symmetry of carbon p_z orbitals, $t(\mathbf{r})$ and therefore $t_{\mathbf{q}}$ are orientation-independent, i.e. $t_{\mathbf{q}} = t_q$.

As shown in Fig. 4(b), the three allowed momentum boosts, defined by the difference between the momenta in the top and bottom layers $\tilde{\mathbf{g}}_j = \mathbf{k} - \mathbf{k}' = \mathbf{G}'_j - \mathbf{G}_j$, are

$$\tilde{\mathbf{g}}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{g}}_2 = \frac{4\pi}{\sqrt{3}a_M}\begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad \tilde{\mathbf{g}}_3 = \frac{4\pi}{\sqrt{3}a_M}\begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \tag{13}$$

where $a_M = a/(2\sin(\theta/2) \approx a/\theta$ is the moiré lattice constant.

Momenta $\mathbf{k} + \mathbf{G}_j$ in Eq. (7) are near Brillouin zone corners. To zeroth order,

$$t_{\mathbf{k}+\mathbf{G}_j} \approx t_{\mathbf{k}_{D,j}} = t_{k_D} = w_0 A_{uc}, \tag{14}$$

where $\mathbf{k}_{D,j}$ are Dirac points of the unrotated graphene

$$\mathbf{k}_{D,j} = e^{i2\pi(j-1)/3}(1,0)\frac{4\pi}{3a}. \tag{15}$$

The interlayer tunneling is local and therefore momentum independent

$$T_{\mathbf{k}\mathbf{k}'} = w_0 \sum_{j=1}^{3} T_j \delta_{\mathbf{k}-\mathbf{k}',\tilde{\mathbf{g}}_j}. \tag{16}$$

The approximate particle-hole symmetry is broken by taking into account non-local interlayer tunneling effects[47]. Keeping the expansion in $t_{\mathbf{k}+\mathbf{G}_j}$ till the first order,

$$t_{\mathbf{k}} \approx t_{k_D} + t'(k - k_D), \tag{17}$$

where

$$t' = \left.\frac{dt}{dk}\right|_{k=k_D} < 0 \tag{18}$$

is the tunable non-local tunneling parameter. The momentum-dependent interlayer tunneling becomes

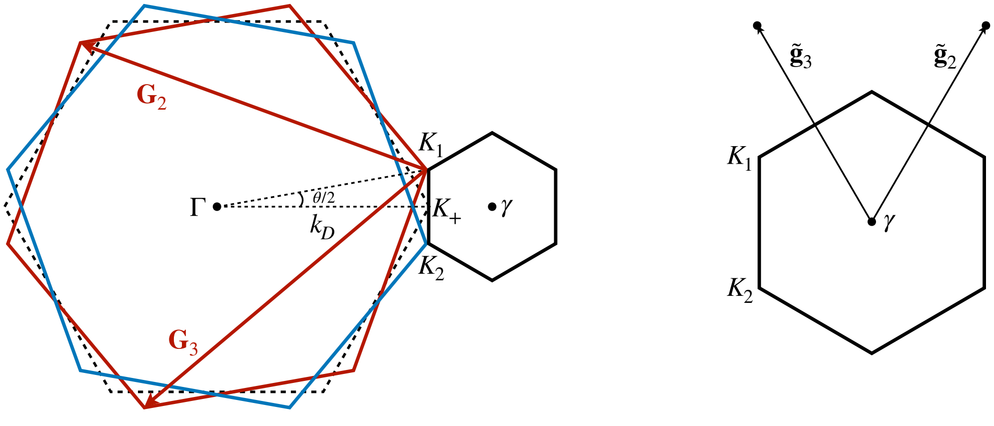
$$\begin{aligned} T_{\mathbf{k}\mathbf{k}'} &= \frac{1}{A_{uc}} \sum_{j=1}^{3} t_{\mathbf{k}+\mathbf{G}_j} T_j \delta_{\mathbf{k}-\mathbf{k}',\tilde{\mathbf{g}}_j} \\ &= \sum_{j=1}^{3} \left[w_0 + \frac{w_{nl}}{g_M}(|\mathbf{k}+\mathbf{G}_j| - k_D) \right] T_j \delta_{\mathbf{k}-\mathbf{k}',\tilde{\mathbf{g}}_j}. \end{aligned} \tag{19}$$

The non-local tunneling parameter w_{nl} is defined as

$$w_{nl} = \frac{g_M}{A_{uc}} t', \tag{20}$$

where $g_M = 4\pi/\sqrt{3}a_M$ is the length of moiré primitive reciprocal lattice vector.

In the numerical calculations in this Letter, we take $v_F = 0.866 \times 10^6$ m/s, $u = w_{AA}/w_{AB} = 0.6$, $w_0 = 110$ meV and $w_{nl} = -20$ meV for MATBG of twist angle $1.1°$. For other efforts to improve the accuracy of the single-particle Hamiltonian of twisted bilayer graphene see Refs. [66–69].



(a) (b)

FIG. 4. (a) Rotated Brillouin zones of top (red) and bottom (blue) graphene layers. The top layer is rotated anticlockwise with respect to the bottom layer by θ. The black dashed hexagonal is the Brillouin zone of unrotated graphene. K_+, K_1 and K_2 are Dirac points of unrotated graphene, the top graphene layer and the bottom graphene layer respectively. k_D is the length of $\Gamma - K_+$. The smaller black hexagon on the right is the moiré Brillouin zone. (b) The moiré Brillouin zone. Three momentum boosts $\tilde{\mathbf{g}}_j$ in Eq. (13) are shown.

II. TBG HAMILTONIAN

The Hamiltonian of a periodic crystal is

$$\hat{H} = \hat{T}_e + \hat{H}_{e-e} + \hat{H}_{e-b} + \hat{H}_{b-b}, \tag{21}$$

where $\hat{T}_e$ is the kinetic energy operator. Instead of the uniform background charge in the Jellium model, a periodic background of the positive charge

$$n_b(\mathbf{r}) = \frac{1}{A}\sum_{\mathbf{G}} n_{b,\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \tag{22}$$

is assumed. At CN, the background charge cancels with excess electron charge

$$n_{b,\mathbf{0}} = \int n_b(\mathbf{r}) d\mathbf{r} = N. \tag{23}$$

Written in electron number density operators $\hat{n}(\mathbf{q})$ and the total number operator $\hat{N}$

$$\begin{aligned} \hat{n}(\mathbf{q}) &= \sum_{\mathbf{k},\mathbf{G}} c^\dagger_{\mathbf{k}+\mathbf{G}-\mathbf{q}} c_{\mathbf{k}+\mathbf{G}}, \\ \hat{N} = \hat{n}(\mathbf{q}=0) &= \sum_{\mathbf{k},\mathbf{G}} c^\dagger_{\mathbf{k}+\mathbf{G}} c_{\mathbf{k}+\mathbf{G}}, \end{aligned} \tag{24}$$

the interacting Hamiltonians are

$$\begin{aligned} \hat{H}_{e-b} &= -e^2 \int d\mathbf{r}d\mathbf{r}' \frac{\hat{n}(\mathbf{r}) n_b(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = -\frac{1}{A}\sum_{\mathbf{G}} V(\mathbf{G}) n_{b,\mathbf{G}} \hat{n}_{-\mathbf{G}}, \\ \hat{H}_{b-b} &= \frac{e^2}{2}\int d\mathbf{r}d\mathbf{r}' \frac{n_b(\mathbf{r}) n_b(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{2A}\sum_{\mathbf{G}} V(\mathbf{G}) n_{b,\mathbf{G}} n_{b,-\mathbf{G}}, \\ \hat{H}_{e-e} &= \frac{e^2}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = \frac{1}{2A}\sum_{\mathbf{q},\mathbf{G}} V(\mathbf{q}+\mathbf{G})\big[\hat{n}(\mathbf{q}+\mathbf{G})\hat{n}(-\mathbf{q}-\mathbf{G}) - \hat{N}\big]. \end{aligned} \tag{25}$$

Similar to the Jellium model, the $\mathbf{G}=0$ terms of $\hat{H}_{e-b}$ and $\hat{H}_{b-b}$, and the $\mathbf{q}=\mathbf{G}=0$ term of $\hat{H}_{e-e}$ cancel in the thermodynamic limit $A, N \to \infty$. The full Hamiltonian in Eq. (21) can be written as

$$\begin{aligned} \hat{H} &= \hat{T}_e + \hat{H}_{e-e} + \hat{V}_b \\ &= \hat{T}_e + \frac{1}{2A}\sideset{}{'}\sum_{\mathbf{q},\mathbf{G}} V(\mathbf{q}+\mathbf{G})\big[\hat{n}(\mathbf{q}+\mathbf{G})\hat{n}(-\mathbf{q}-\mathbf{G}) - \hat{N}\big] + \frac{1}{2A}\sum_{\mathbf{G}\neq 0} V(\mathbf{G})\Big[-2n_{b,\mathbf{G}}\hat{n}_{-\mathbf{G}} + n_{b,\mathbf{G}} n_{b,-\mathbf{G}}\Big], \end{aligned} \tag{26}$$

where $'$ on the summation symbol means $\mathbf{q}=\mathbf{G}=0$ term is excluded in the momentum summation.

It is straightforward to generalize the electron-electron interacting Hamiltonian in Eq. (26) to the TBG case, where two layers and two sublattices degrees of freedom are explicitly included,

$$\hat{H}_{e-e} = \frac{1}{2A}\sideset{}{'}\sum_{\substack{\mathbf{q},\mathbf{g},\alpha,\beta \\ \mathbf{k}_1,\mathbf{k}_2,\mathbf{g}_1,\mathbf{g}_2}} V_{\alpha\beta}(\mathbf{q}+\mathbf{g}) c^\dagger_{\alpha,\mathbf{k}_1+\mathbf{g}_1-\mathbf{q}-\mathbf{g}} c^\dagger_{\beta,\mathbf{k}_2+\mathbf{g}_2+\mathbf{q}+\mathbf{g}} c_{\beta,\mathbf{k}_2+\mathbf{g}_2} c_{\alpha,\mathbf{k}_1+\mathbf{g}_1}, \tag{27}$$

where $\mathbf{g}, \mathbf{g}_1, \mathbf{g}_2$ are moiré reciprocal lattice vectors and $\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}$ are momenta in the first moiré Brillouin zone. α, β label layers and sublattices. The $'$ on top of the summation means the $\mathbf{q}=\mathbf{g}=0$ term is excluded, which is cancelled by the periodic background of positive charge in the thermodynamic limit.

Alternatively, $\hat{H}_{e-e}$ can be expressed in electron number density operators

$$\hat{H}_{e-e} = -\frac{1}{2A}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V_S(\mathbf{q}+\mathbf{g})\hat{N} + \frac{1}{2A}\sideset{}{'}\sum_{\substack{\mathbf{q},\mathbf{g}\\ \alpha,\beta}} V_{\alpha\beta}(\mathbf{q}+\mathbf{g}) \sum_{\mathbf{g}_1,\mathbf{g}_2} \hat{n}^{\mathbf{g}_1}_{\alpha}(\mathbf{q}+\mathbf{g})\hat{n}^{\mathbf{g}_2}_{\beta}(-\mathbf{q}-\mathbf{g}), \tag{28}$$

where

$$\begin{aligned}\hat{N} &= \sum_{\mathbf{k},\mathbf{g},\alpha} c^{\dagger}_{\alpha,\mathbf{k}+\mathbf{g}} c_{\alpha,\mathbf{k}+\mathbf{g}},\\ \hat{n}(\mathbf{q}) &= \sum_{\alpha,\mathbf{g}} \hat{n}^{\mathbf{g}}_{\alpha}(\mathbf{q}) = \sum_{\alpha,\mathbf{g}} \sum_{\mathbf{k}} c^{\dagger}_{\alpha,\mathbf{k}+\mathbf{g}-\mathbf{q}} c_{\alpha,\mathbf{k}+\mathbf{g}}.\end{aligned} \tag{29}$$

For isotropic dielectric media, the Coulomb interaction within the same 2D layer (V_S) and between different 2D layers (V_D) are respectively

$$\begin{aligned}V_S(\mathbf{q}) &= \frac{2\pi e^2}{\epsilon q},\\ V_D(\mathbf{q}) &= \frac{2\pi e^2}{\epsilon q} e^{-qd},\end{aligned} \tag{30}$$

where ϵ is the relative dielectric constant of the surrounding environment and d is the distance between adjacent layers.

For anisotropic dielectric media, for example hBN, $\epsilon_{\text{BN}} = \sqrt{\epsilon_{zz}\epsilon_{\perp}} \approx 4.5$ with $\epsilon_{\perp} = 6.9$[75], $\epsilon_{zz} = 3$.

$$\begin{aligned}V_S(\mathbf{q}) &= \frac{2\pi e^2}{q\sqrt{\epsilon_{\perp}\epsilon_{zz}}},\\ V_D(\mathbf{q}) &= \frac{2\pi e^2}{q\sqrt{\epsilon_{\perp}\epsilon_{zz}}} e^{-qd\sqrt{\epsilon_{\perp}/\epsilon_{zz}}}.\end{aligned} \tag{31}$$

With dual metallic gates that is equidistant from the target 2D system, with distance d_m, the screened Coulomb potential energy is

$$V_{sc}(\mathbf{q}) = \frac{2\pi e^2}{q\sqrt{\epsilon_{\perp}\epsilon_{zz}}} \tanh\left(q d_m \sqrt{\frac{\epsilon_{\perp}}{\epsilon_{zz}}}\right). \tag{32}$$

In calculations throughout this paper, the d-dependence of the Coulomb potential in Eq. (31) is ignored, *i.e.*,

$$V(\mathbf{q}) = V_S(\mathbf{q}) = V_D(\mathbf{q}) = \frac{2\pi e^2}{q\epsilon_{\text{BN}}}, \tag{33}$$

and the hBN dielectric constant is chosen to be $\epsilon_{\text{BN}} = 5.1$. Then Eq. (28) simplifies to

$$\hat{H}_{e-e} = \frac{1}{2A} \sum_{\mathbf{q},\mathbf{g}}{}' V(\mathbf{q}+\mathbf{g}) \left[\hat{n}(\mathbf{q}+\mathbf{g})\hat{n}(-\mathbf{q}-\mathbf{g}) - \hat{N}\right]. \tag{34}$$

III. SELF-CONSISTENT HARTREE APPROXIMATION IN TBG

With approximate SU(4) symmetry, the exchange interaction is only between electrons within the same flavor. The inter-flavor interaction is only through the Hartree potential. The matrix element of the flavor- and momentum-independent Hartree self-energy is

$$\Sigma^{H}_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2} = \frac{\delta_{\alpha\beta}}{A} \sum_{\alpha',\mathbf{g}'_1,\mathbf{g}'_2} V_{\alpha\alpha'}(\mathbf{g}_2 - \mathbf{g}_1) \delta\rho_{\alpha',\mathbf{g}'_1;\alpha',\mathbf{g}'_2} \delta_{\mathbf{g}'_2-\mathbf{g}'_1,\mathbf{g}_2-\mathbf{g}_1}, \tag{35}$$

where $\delta\boldsymbol{\rho}$ is the relative density matrix defined by subtracting the density matrix of the decoupled graphene bilayer at CN

$$\delta\boldsymbol{\rho} = \boldsymbol{\rho} - \boldsymbol{\rho}_0. \tag{36}$$

The flavor-polarization-dependence of the self-energy is implicitly incorporated in the density matrix through the Fermi-Dirac distribution $\Theta(\mu_f - \varepsilon^{f}_{n\mathbf{k}})$:

$$\rho_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2} = \sum_{f,n,\mathbf{k}} z^{nf}_{\alpha,\mathbf{g}_1}(\mathbf{k}) \bar{z}^{nf}_{\beta,\mathbf{g}_2}(\mathbf{k}) \Theta(\mu_f - \varepsilon^{f}_{n\mathbf{k}}). \tag{37}$$

$f = 1, 2, 3, 4$ represent four flavors. If the inversion symmetry is retained, the diagonal elements of Σ^H are constant and can be set to zero.

For any specific flavor polarization $(\nu_1, \nu_2, \nu_3, \nu_4)$, the Hamiltonian is solved self-consistently for each flavor

$$H^f(\mathbf{k}, \nu_1, \nu_2, \nu_3, \nu_4) = H_0^f(\mathbf{k}) + \Sigma^H(\nu_1, \nu_2, \nu_3, \nu_4), \tag{38}$$

where $H_0^f(\mathbf{k})$ is the single-particle Hamiltonian of flavor f.

Self-consistent Hartree band structures and corresponding Fermi surfaces of flavor-symmetry unbroken states at filling factors $\nu \in (-1, 1)$ are shown in Fig. 5.

The electrostatic Hartree energy is

$$\begin{aligned} E_H &= \frac{1}{2} \sum_{f,n,\mathbf{k}} \langle \Psi_{fn\mathbf{k}} | \Sigma^H | \Psi_{fn\mathbf{k}} \rangle \Theta(\mu_f - \varepsilon^f_{n\mathbf{k}}) \\ &= \frac{1}{2A} \sum_{\substack{\mathbf{g}_1, \mathbf{g}_2 \\ \mathbf{g}'_1, \mathbf{g}'_2}} \sum_{\alpha, \alpha'} V_{\alpha\alpha'}(\mathbf{g}_2 - \mathbf{g}_1) \bar{\rho}_{\alpha, \mathbf{g}_1; \alpha, \mathbf{g}_2} \delta\rho_{\alpha', \mathbf{g}'_1; \alpha', \mathbf{g}'_2} \delta_{\mathbf{g}'_2 - \mathbf{g}'_1, \mathbf{g}_2 - \mathbf{g}_1}, \end{aligned} \tag{39}$$

and is further regularized by subtracting the negative Fermi sea contribution

$$E_H = \frac{1}{2A} \sum_{\substack{\mathbf{g}_1, \mathbf{g}_2 \\ \mathbf{g}'_1, \mathbf{g}'_2}} \sum_{\alpha, \alpha'} V_{\alpha\alpha'}(\mathbf{g}_2 - \mathbf{g}_1) \delta\bar{\rho}_{\alpha, \mathbf{g}_1; \alpha, \mathbf{g}_2} \delta\rho_{\alpha', \mathbf{g}'_1; \alpha', \mathbf{g}'_2} \delta_{\mathbf{g}'_2 - \mathbf{g}'_1, \mathbf{g}_2 - \mathbf{g}_1}. \tag{40}$$

The SCH energy E_0 in the main text is defined to include both the SCH band dispersion and the electrostatic Hartree energy:

$$\begin{aligned} E_0 &= E_{\text{band}} - E_H \\ &= \sum_{f,n,\mathbf{k}} \varepsilon^f_{n\mathbf{k}} \Theta(\mu_f - \varepsilon^f_{n\mathbf{k}}) - E_H, \end{aligned} \tag{41}$$

where $\varepsilon^f_{n\mathbf{k}}$ are eigenvalues of SCH Hamiltonian Eq. (38) and E_H is subtracted to avoid double-counting of the Coulomb energy.

IV. SELF-CONSISTENT HARTREE-FOCK FERMI SURFACES

To compare with the self-consistent Hartree approximation, we present in Fig. 6 the Fermi surfaces of self-consistent Hartree-Fock calculations with momentum- and flavor-dependent Fock self-energy

$$\Sigma^{F,f}_{\alpha, \mathbf{g}_1; \beta, \mathbf{g}_2}(\mathbf{k}) = -\frac{1}{A} \sum_{\mathbf{k}', \mathbf{g}'_1, \mathbf{g}'_2} V_{\alpha\beta}(\mathbf{k}' - \mathbf{k} + \mathbf{g}'_1 - \mathbf{g}_1) \delta\rho^f_{\alpha, \mathbf{g}'_1; \beta, \mathbf{g}'_2}(\mathbf{k}') \delta_{\mathbf{g}'_2 - \mathbf{g}'_1, \mathbf{g}_2 - \mathbf{g}_1}. \tag{42}$$

V. THE DENSITY RESPONSE FUNCTION OF TBG

At zero temperature, the matrix elements of the flavor-specified density response function $\tilde{\chi}^f$ of MATBG can be derived following the Lindhard formula

$$[\tilde{\chi}^f]^{\mathbf{g}\mathbf{g}'}(\mathbf{q}, \omega) = \frac{1}{A} \sum_{n,m,\mathbf{k}} \frac{\Theta^f_{n\mathbf{k}} - \Theta^f_{m\mathbf{k}+\mathbf{q}}}{\omega + \varepsilon^f_{n\mathbf{k}} - \varepsilon^f_{m\mathbf{k}+\mathbf{q}} + i\eta} \Big[\sum_{\alpha, \mathbf{g}_1} z^{nf}_{\alpha, \mathbf{g}_1}(\mathbf{k}) \bar{z}^{mf}_{\alpha, \mathbf{g}_1 + \mathbf{g}}(\mathbf{k} + \mathbf{q}) \Big]^* \sum_{\beta, \mathbf{g}_2} z^{nf}_{\beta, \mathbf{g}_2}(\mathbf{k}) \bar{z}^{mf}_{\beta, \mathbf{g}_2 + \mathbf{g}'}(\mathbf{k} + \mathbf{q}). \tag{43}$$

where ε and z are quasiparticle eigen-energy and eigenvector of the self-consistent Hartree approximation. It can be easily proved that Eq. (43) satisfies the symmetry

$$[\tilde{\chi}^f]^{\mathbf{g}\mathbf{g}'}(\mathbf{q}, \omega) = \big[[\tilde{\chi}^f]^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q}, -\omega) \big]^*. \tag{44}$$

It obeys the general property of the response function

$$\tilde{\chi}_{AB}(\omega) = [\tilde{\chi}_{A^\dagger B^\dagger}(-\omega)]^*. \tag{45}$$

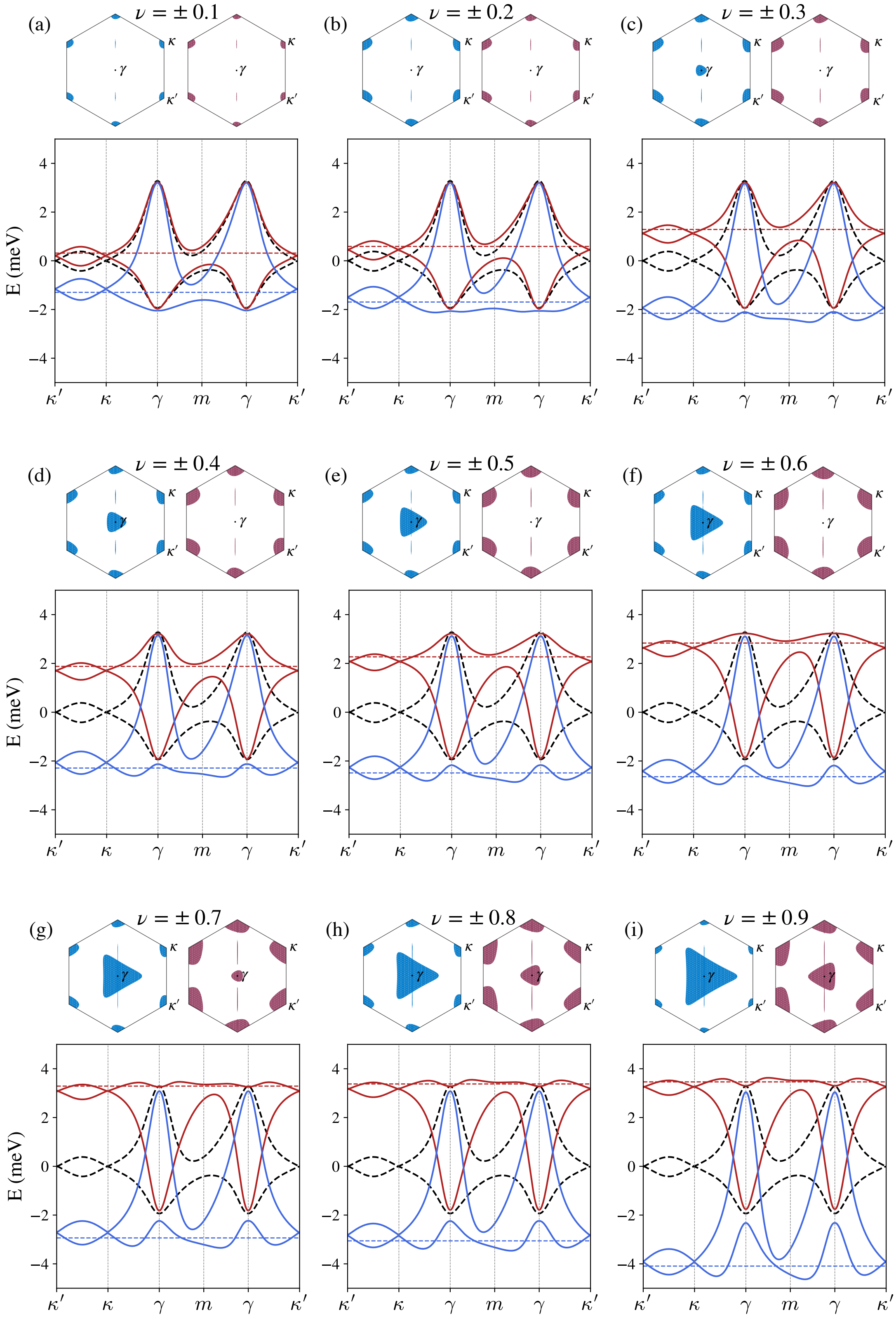


FIG. 5. Self-consistent Hartree band structures (colored lines) and corresponding Fermi surfaces (colored shaded areas) of flavor paramagnetic states at filling factors $\nu \in (-1, 1)$. Red (blue) represents electron (hole) doping. The black dashed line in each spectrum is the single-particle band structure. Note that the energy bands and the Fermi levels are shifted such that the non-interacting band energy at κ is zero.

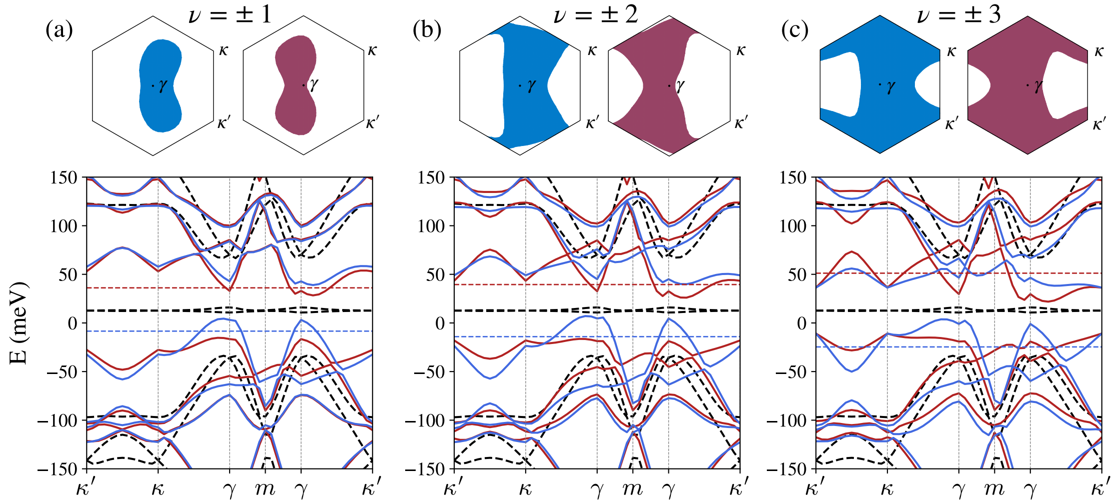


FIG. 6. Self-consistent Hartree-Fock band structures (colored lines) and corresponding Fermi surfaces (colored shaded areas) of flavor paramagnetic states at filling factors $\nu = \pm 1, \pm 2, \pm 3$. Red (blue) represents electron (hole) doping. The black dashed line in each spectrum is the single-particle band structure.

The density response functions of opposite valleys with the same electrostatic doping level are related by the spinless time-reversal symmetry

$$[\tilde{\chi}^-]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) = [\tilde{\chi}^+]^{-\mathbf{g}'-\mathbf{g}}(-\mathbf{q},\omega), \tag{46}$$

following reciprocity relations

$$\tilde{\chi}^-_{AB}(\omega) = \tilde{\chi}^+_{B^T A^T}(\omega), \tag{47}$$

where ± represent two opposite valleys in TBG. If the TBG system is flavor unpolarized, the total density response function $\tilde{\chi}$ summed over four flavors is time-reversal invariant and satisfies

$$\begin{aligned}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) &= 2[\tilde{\chi}^+]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) + 2[\tilde{\chi}^-]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega)\\ &= 2[\tilde{\chi}^+]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) + 2[\tilde{\chi}^+]^{-\mathbf{g}'-\mathbf{g}}(-\mathbf{q},\omega)\\ &= \tilde{\chi}^{-\mathbf{g}'-\mathbf{g}}(-\mathbf{q},\omega).\end{aligned} \tag{48}$$

At finite temperatures, the density response function is

$$[\tilde{\chi}^f]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) = \frac{1}{A}\sum_{n,m,\mathbf{k}} \frac{\Theta^f_{n\mathbf{k}} - \Theta^f_{m\mathbf{k}+\mathbf{q}}}{i\omega + \varepsilon^f_{n\mathbf{k}} - \varepsilon^f_{m\mathbf{k}+\mathbf{q}}} \Big[\sum_{\alpha,\mathbf{g}_1} z^{nf}_{\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^{mf}_{\alpha,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}+\mathbf{q})\Big]^* \sum_{\beta,\mathbf{g}_2} z^{nf}_{\beta,\mathbf{g}_2}(\mathbf{k})\bar{z}^{mf}_{\beta,\mathbf{g}_2+\mathbf{g}'}(\mathbf{k}+\mathbf{q}). \tag{49}$$

Matrix $\tilde{\chi}^f(\mathbf{q},i\omega)$ is Hermitian along the imaginary frequency axis, *i.e.*,

$$\begin{aligned}\tilde{\chi}^f(\mathbf{q},i\omega) &= [\tilde{\chi}^f]^\dagger(\mathbf{q},-i\omega),\\ [\tilde{\chi}^f]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) &= \big[[\tilde{\chi}^f]^{\mathbf{g}'\mathbf{g}}(\mathbf{q},-i\omega)\big]^*.\end{aligned} \tag{50}$$

$\tilde{\chi}^f(\mathbf{q},i\omega)$ also satisfies

$$[\tilde{\chi}^f]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) = [\tilde{\chi}^f]^{-\mathbf{g}'-\mathbf{g}}(-\mathbf{q},-i\omega) = \big[[\tilde{\chi}^f]^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},i\omega)\big]^*. \tag{51}$$

Opposite valleys with the same electrostatic doping level are related by the time-reversal symmetry

$$[\tilde{\chi}^-]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) = [\tilde{\chi}^+]^{-\mathbf{g}'-\mathbf{g}}(-\mathbf{q},i\omega) = \big[[\tilde{\chi}^+]^{\mathbf{g}'\mathbf{g}}(\mathbf{q},i\omega)\big]^* = \big[[\tilde{\chi}^-]^{\mathbf{g}'\mathbf{g}}(\mathbf{q},-i\omega)\big]^*. \tag{52}$$

For valley unpolarized state, the total proper density response function in Eq. (3) satisfies

$$\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) = 2[\tilde{\chi}^+]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) + 2[\tilde{\chi}^-]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) = 2[\tilde{\chi}^+]^{\mathbf{g}\mathbf{g}'}(\mathbf{q},i\omega) + 2\big[[\tilde{\chi}^+]^{\mathbf{g}'\mathbf{g}}(\mathbf{q},i\omega)\big]^*, \tag{53}$$

and therefore the density response function of flavor unpolarized state is Hermitian, *i.e.*,

$$\tilde{\chi}(\mathbf{q},i\omega) = \tilde{\chi}^\dagger(\mathbf{q},i\omega). \tag{54}$$

We see that symmetries of $\tilde{\chi}(\mathbf{q},i\omega)$ are different from those of $\tilde{\chi}(\mathbf{q},\omega)$ at zero temperature.

VI. THE COUPLING-CONSTANT INTEGRATION AS AN EVALUATION OF THE GROUND STATE ENERGY

The ground-state energy of an electron system can be easily connected to the density-density linear-response function—for example at the level of the popular RPA — by the integration over the coupling constant theorem (see, for example, Sect. I.8.3 of Ref. [76]). However, the ground-state electron density of a moiré crystal is inhomogeneous on the moiré superlattice length scale. The application of the integration over the coupling constant theorem to an inhomogeneous many-body system is not straightforward and presents some subtleties. On the contrary, the Hohenberg-Kohn and Kohn-Sham theorems (see, for example, Chapter 7 of Ref. [76]) of density functional theory are the natural theoretical framework to deal with inhomogeneous many-body electron systems. This is why we formulate the problem of the calculation of the ground-state energy of twisted bilayer graphene within the framework of Adiabatic Connection Fluctuation and Dissipation Theorem (ACFDT) [86] and then recover the RPA by taking a suitable limit. This theory applies the integration over the coupling constant theorem between the Kohn-Sham (KS) ground state and the real ground state of the electronic Hamiltonian as detailed in the following.

A. The Adiabatic Connection Fluctuation and Dissipation Theorem

The electronic Hamiltonian can be written as

$$\hat{\mathcal{H}} = \hat{T}_{\rm e} + \hat{\mathcal{H}}_{\rm e-e} + \hat{V}_{\rm ext} \ , \tag{55}$$

where $\hat{T}_{\rm e}$ is the kinetic operator, $\hat{\mathcal{H}}_{\rm e-e}$ is the electron-electron interaction, i.e.

$$\hat{H}_{\rm e-e} = \frac{1}{2A} \sum_{\mathbf{q},\mathbf{G}}{}' V(\mathbf{q}+\mathbf{G}) \left[\hat{n}(\mathbf{q}+\mathbf{G}) \hat{n}(-\mathbf{q}-\mathbf{G}) - \hat{N} \right], \tag{56}$$

and $\hat{V}_{\rm ext}$ is the crystal potential, i.e.

$$\hat{V}_{\rm ext} = \int d^2\mathbf{r} V_{\rm ext}(\mathbf{r}) \hat{n}(\mathbf{r}) \ . \tag{57}$$

In Eq. (56), $V(\mathbf{q}+\mathbf{G})$ is the Fourier transform of the electron-electron interaction potential, evaluated at the wave vector $\mathbf{q}+\mathbf{G}$, where $\mathbf{q}$ is in the first BZ and $\mathbf{G}$ is an arbitrary reciprocal lattice vector. Similarly, $\hat{n}(\mathbf{q}+\mathbf{G})$ is the Fourier transform of the ground-state density operator $\hat{n}(\mathbf{r})$.

Let $|\Phi\rangle$ be the ground state of the full Hamiltonian and $n(\mathbf{r})$ the associated ground-state density, i.e.

$$n(\mathbf{r}) = \langle \Phi | \hat{n}(\mathbf{r}) | \Phi \rangle \ . \tag{58}$$

We now introduce a key auxiliary system, which is described by the so-called *Kohn-Sham Hamiltonian* $\hat{\mathcal{H}}_{\rm KS}$:

$$\hat{\mathcal{H}}_{\rm KS} = \hat{T}_{\rm e} + \hat{V}_{\rm ext} + \hat{V}_{\rm H} + \hat{V}_{\rm xc} \equiv \hat{T}_{\rm e} + \hat{V}_{\rm KS} \ . \tag{59}$$

This Hamiltonian is a (self-consistent) one-particle Hamiltonian whose ground state $|\psi_{\rm KS}\rangle$ is a Slater determinant of Kohn-Sham orbitals. The fundamental property of this Hamiltonian is that *it yields the exact same density of the full Hamiltonian* (55).

We now define a family $\hat{\mathcal{H}}_\lambda$ of Hamiltonians depending on a real dimensionless parameter $\lambda \in [0, 1]$:

$$\hat{\mathcal{H}}_\lambda = \hat{T}_{\rm e} + \lambda \hat{\mathcal{H}}_{\rm e-e} + \hat{V}_\lambda \ . \tag{60}$$

Let $|\psi_\lambda\rangle$ be the normalized ground state of $\hat{\mathcal{H}}_\lambda$, i.e. $\langle \psi_\lambda | \psi_\lambda \rangle = 1 \ \forall \lambda$. In Eq. (60), $\hat{T}_{\rm e}$ and $\hat{\mathcal{H}}_{\rm e-e}$ have the exact same meaning as above. The key new quantity is $\hat{V}_\lambda$, which is a local potential that interpolates between the Kohn-Sham potential in the limit $\lambda = 0$, i.e. $\hat{V}_{\lambda=0} = \hat{V}_{\rm KS}$, and the exact, physical crystal potential in the limit $\lambda = 1$, i.e. $\hat{V}_{\lambda=1} = \hat{V}_{\rm ext}$. Crucially, $\hat{V}_\lambda$ varies with λ in such a way that the correct electronic density is reproduced at every value of λ, i.e.

$$n_\lambda(\mathbf{r}) \equiv \langle \psi_\lambda | \hat{n}(\mathbf{r}) | \psi_\lambda \rangle = n(\mathbf{r}) \ \forall \lambda \ . \tag{61}$$

The uniqueness (up to a constant) of $\hat{V}_\lambda$ can be proved by applying the Hohenberg-Kohn theorem to the electronic Hamiltonian with reduced coupling constant. In the following, we will fix the arbitrary constant by setting the average value of $\hat{V}_\lambda$ to zero.

The ground state energy at coupling constant λ is

$$E(\lambda) = \langle \psi_\lambda | \hat{\mathcal{H}}_\lambda | \psi_\lambda \rangle \ . \tag{62}$$

By applying the Hellman-Feynman theorem we obtain

$$\frac{dE(\lambda)}{d\lambda} = \langle\psi_\lambda|\hat{\mathcal{H}}_{\rm e-e}|\psi_\lambda\rangle + \int d^2\mathbf{r} n(\mathbf{r})\partial_\lambda V_\lambda(\mathbf{r}) \ . \tag{63}$$

Integrating the previous differential equation between $\lambda = 0$ and $\lambda = 1$ we find:

$$E(1) - E(0) = \int_0^1 d\lambda\langle\psi_\lambda|\hat{\mathcal{H}}_{\rm e-e}|\psi_\lambda\rangle + \int d^2\mathbf{r} n(\mathbf{r})[V_{\rm ext}(\mathbf{r}) - V_{\rm KS}(\mathbf{r})] \ . \tag{64}$$

In writing the previous two equations we made use of the crucial fact that $n_\lambda(\mathbf{r}) = n(\mathbf{r})$ for every λ in the integration interval.

We now recall that, in DFT, the exact ground-state energy E of the system, which in the notation of Eq. (64) coincides with $E(1)$, is given by [76]

$$E = E(1) = T_{\rm s} + \int d^2\mathbf{r} n(\mathbf{r})V_{\rm ext}(\mathbf{r}) + E_{\rm H} + E_{\rm xc} \ , \tag{65}$$

where $T_{\rm s}$ is the non-interacting kinetic energy functional, i.e. the kinetic energy of a non-interacting system whose ground-state density is $n(\mathbf{r})$, $E_{\rm H}$ is the Hartree energy, i.e.

$$E_{\rm H} = \frac{1}{2}\int d^2\mathbf{r}\int d^2\mathbf{r}' V(|\mathbf{r}-\mathbf{r}'|)n(\mathbf{r})n(\mathbf{r}') \ , \tag{66}$$

and $E_{\rm xc}$ is the exchange-correlation energy functional.

On the other hand, the quantity $E(0)$ is the average of the Kohn-Sham Hamiltonian over the Kohn-Sham ground-state

$$E(0) = \sum_{\alpha\in{\rm occ.}} \epsilon_\alpha^{\rm KS} = T_{\rm s} + \int d^2\mathbf{r} n(\mathbf{r})V_{\rm KS}(\mathbf{r}) \ , \tag{67}$$

where $\epsilon_\alpha^{\rm KS}$ are the eigenvalues of the Kohn-Sham equations and the sum runs over the occupied states.

Taking the difference between Eq. (65) and Eq. (67) and comparing the result with Eq. (64) we find the following important result:

$$E_{\rm H} + E_{\rm xc} = \int_0^1 d\lambda\langle\psi_\lambda|\hat{\mathcal{H}}_{\rm e-e}|\psi_\lambda\rangle \ . \tag{68}$$

We can conveniently express the matrix element $\langle\psi_\lambda|\hat{\mathcal{H}}_{\rm e-e}|\psi_\lambda\rangle$ of the interaction Hamiltonian in terms of the density-density response function by using the fluctuation-dissipation theorem [76]. At zero temperature and assuming a non-degenerate ground state, we obtain

$$\begin{aligned}\langle\psi_\lambda|\hat{n}_{\mathbf{q}+\mathbf{G}}\hat{n}_{-\mathbf{q}-\mathbf{G}}|\psi_\lambda\rangle &= -\frac{\hbar}{\pi}\int_0^\infty {\rm Im}\left[\chi_{\hat{n}_{\mathbf{q}+\mathbf{G}}\hat{n}_{-\mathbf{q}-\mathbf{G}}}(\omega,\lambda)\right]d\omega + \langle\Phi_\lambda|\hat{n}_{\mathbf{q}+\mathbf{G}}|\Phi_\lambda\rangle\langle\Phi_\lambda|\hat{n}_{-\mathbf{q}-\mathbf{G}}|\Phi_\lambda\rangle \\ &= -\frac{\hbar A}{\pi}\int_0^\infty {\rm Im}\left[\chi_{nn}^{\mathbf{G}\mathbf{G}}(\mathbf{q},\omega,\lambda)\right]d\omega + n_{\mathbf{q}+\mathbf{G}}n_{-\mathbf{q}-\mathbf{G}} \ .\end{aligned} \tag{69}$$

Making use of (69) and (56) we get

$$\langle\psi_\lambda|\hat{\mathcal{H}}_{\rm e-e}|\psi_\lambda\rangle = \frac{N}{2}\sum_{\mathbf{G}}\int_{\rm BZ}\frac{d^2\mathbf{q}}{(2\pi)^2}V_{\mathbf{q}+\mathbf{G}}\left\{\frac{n_{\mathbf{q}+\mathbf{G}}n_{-\mathbf{q}-\mathbf{G}}}{N} - \frac{\hbar}{n\pi}\int_0^\infty d\omega{\rm Im}[\chi_{nn}^{\mathbf{G}\mathbf{G}}(\mathbf{q},\omega,\lambda)] - 1\right\}. \tag{70}$$

The first term is independent of λ and coincides with the Hartree energy,

$$E_{\rm H} = \frac{1}{2A}\sum_{\mathbf{G}\neq\mathbf{0}} V_{\mathbf{G}}n_{\mathbf{G}}n_{-\mathbf{G}} \ . \tag{71}$$

We are therefore left with

$$E_{\rm xc} = \int_0^1 d\lambda\frac{N}{2}\sum_{\mathbf{G}}\int\frac{d^2\mathbf{q}}{(2\pi)^2}V_{\mathbf{q}+\mathbf{G}}\left\{-\frac{\hbar}{n\pi}\int_0^\infty d\omega{\rm Im}[\chi_{nn}^{\mathbf{G}\mathbf{G}}(\mathbf{q},\omega,\lambda)] - 1\right\} \ . \tag{72}$$

It can be further shown that the exchange energy can be written as

$$E_{\rm x} = \frac{N}{2} \sum_{\mathbf{G}} \int \frac{d^2\mathbf{q}}{(2\pi)^2} V_{\mathbf{q}+\mathbf{G}} \left\{ -\frac{\hbar}{n\pi} \int_0^\infty d\omega {\rm Im}[\chi_{\rm KS}^{\mathbf{G}\mathbf{G}}(\mathbf{q}, \omega)] - 1 \right\} , \tag{73}$$

where $\chi_{\rm KS}^{\mathbf{G}\mathbf{G}}(\mathbf{q}, \omega)$ is the Kohn-Sham response function. Note that this is the exchange energy calculated on the KS orbitals, which is *different* from the Hartree-Fock exchange.

The difference between $E_{\rm xc}$ and $E_{\rm x}$ is the correlation energy:

$$E_{\rm c} = \frac{N}{2} \int_0^1 d\lambda \sum_{\mathbf{G}} \int_{\rm BZ} \frac{d^2\mathbf{q}}{(2\pi)^2} V_{\mathbf{q}+\mathbf{G}} \left\{ -\frac{\hbar}{n\pi} \int_0^\infty d\omega {\rm Im}[\chi_{nn}^{\mathbf{G}\mathbf{G}}(\mathbf{q}, \omega, \lambda) - \chi_{\rm KS}^{\mathbf{G}\mathbf{G}}(\mathbf{q}, \omega)] \right\} . \tag{74}$$

In a crystal, the linear response relation—see Eq. (7.182) in Ref. [76]—relating the full density response at coupling constant λ to the Kohn-Sham density response reads as following:

$$\delta n_{\mathbf{q}+\mathbf{G}}(\omega) = \sum_{\mathbf{G}'} \chi_{\rm KS}^{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) \left\{ V_{\mathbf{q}+\mathbf{G}'}^{\rm ext}(\omega) + \sum_{\mathbf{G}''} [\lambda V_{\mathbf{q}+\mathbf{G}'} \delta_{\mathbf{G}'\mathbf{G}''} + f_{\rm xc,L}^{\mathbf{G}'\mathbf{G}''}(\mathbf{q}, \omega, \lambda)] \delta n_{\mathbf{q}+\mathbf{G}''}(\omega) \right\} , \tag{75}$$

where $f_{\rm xc,L}^{\mathbf{G}'\mathbf{G}''}(\mathbf{q}, \omega, \lambda)$ is the wave vector and frequency-dependent exchange-correlation kernel [76] evaluate at coupling constant λ. Treating functions in Eq. (75) as matrices with respect to reciprocal lattice vectors indices, we can finally rewrite Eq. (75) as

$$\delta n = \chi_{\rm KS} \cdot \left\{ V^{\rm ext} + [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \delta n \right\} . \tag{76}$$

Using the same notation, the definition of the full response function at coupling constant λ reads as following:

$$\delta n = \chi(\lambda) \cdot V^{\rm ext} . \tag{77}$$

Substituting this definition into Eq. (76) yields

$$\chi(\lambda) = \chi_{\rm KS} \cdot \left\{ 1 + [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \chi(\lambda) \right\} . \tag{78}$$

Carrying out some simple algebraic manipulation we finally find

$$\chi(\lambda) = \chi_{\rm KS} \cdot \{ 1 - [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \chi_{\rm KS} \}^{-1} , \tag{79}$$

which can be further rearranged into

$$\chi(\lambda) - \chi_{\rm KS} = \chi_{\rm KS} \cdot [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \chi_{\rm KS} \cdot \{ 1 - [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \chi_{\rm KS} \}^{-1} . \tag{80}$$

Substituting into the formula for the correlation energy we get the exact expression

$$E_{\rm c} = -\frac{\hbar N}{2\pi n} \int_{\rm BZ} \frac{d^2\mathbf{q}}{(2\pi)^2} \int_0^\infty d\omega \int_0^1 d\lambda {\rm Im} \left\{ \left[V \cdot \chi_{\rm KS} \cdot [\lambda V + f_{\rm xc,L}(\lambda)] \cdot \chi_{\rm KS} \cdot [1 - (\lambda V + f_{\rm xc,L}(\lambda)) \cdot \chi_{\rm KS}]^{-1} \right] (\mathbf{q}, \omega) \right\} . \tag{81}$$

The total energy is finally given by

$$E = \sum_{\alpha \in {\rm occ.}} \epsilon_\alpha^{\rm KS} - \int d^2\mathbf{r} n(\mathbf{r}) V_{\rm xc}(\mathbf{r}) - E_{\rm H} + E_{\rm x} + E_{\rm c} . \tag{82}$$

B. Taking the RPA limit

The RPA is obtained by setting to zero $\hat{V}_{\rm xc}$ in the KS equations and $f_{\rm xc,L}$ in (81). In this limit, the KS equations become the Hartree equations, the KS orbitals become the Hartree orbitals, and the KS response function becomes the Hartree response function.

The coupling-constant integral can then be done analytically, yielding

$$\begin{aligned} E_{\rm c,RPA} &= -\frac{\hbar N}{2\pi n} \int \frac{d^2\mathbf{q}}{(2\pi)^2} \int_0^\infty d\omega \int_0^1 d\lambda \lambda {\rm Im} \Big\{ {\rm Tr} \big[[{\rm V} \cdot \chi_{\rm H}]^2 \cdot [1 - \lambda {\rm V} \cdot \chi_{\rm H}]^{-1} \big] (\mathbf{q}, \omega) \Big\} \\ &= -\frac{\hbar N}{2\pi n} \int \frac{d^2\mathbf{q}}{(2\pi)^2} \int_0^\infty d\omega \int_0^1 d\lambda \lambda {\rm Im} \Big\{ {\rm Tr} \big[(\sqrt{\rm V} \cdot \chi_{\rm H} \cdot \sqrt{\rm V})^2 \cdot (1 - \lambda \sqrt{\rm V} \cdot \chi_{\rm H} \cdot \sqrt{\rm V})^{-1} \big] (\mathbf{q}, \omega) \Big\} \\ &= \frac{N}{2n} \int \frac{d^2\mathbf{q}}{(2\pi)^2} \int_0^\infty \frac{\hbar d\omega}{\pi} {\rm Im} \Big\{ {\rm Tr} [\sqrt{V} \cdot \chi_{\rm H} \cdot \sqrt{V} + \ln(1 - \sqrt{V} \cdot \chi_{\rm H} \cdot \sqrt{V})] (\mathbf{q}, \omega) \Big\} . \end{aligned} \tag{83}$$

Setting $\hat{V}_{\mathrm{xc}} = 0$ in (82) we obtain the final expression for the RPA (or time-dependent Hartree) ground-state energy of an inhomogeneous system,

$$E = \sum_{\alpha}^{\mathrm{occ.}} \epsilon_{\alpha}^{\mathrm{H}} - E_{\mathrm{H}} + E_{\mathrm{x}} + E_{\mathrm{c,RPA}} \ . \tag{84}$$

Note that the Hartree and exchange energies are now calculated on the Hartree orbitals. The first two terms in the previous equation coincide with the Hartree expression of the ground-state energy, avoiding double counting of the Coulomb interaction energy.

VII. THE EXCHANGE-CORRELATION ENERGY OF TBG

Following the ACFDT in SM VI, the exchange-correlation energy of TBG is,

$$\begin{aligned} E_{xc} &= \frac{n}{2}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V(\mathbf{q}+\mathbf{g})\Big[-\frac{1}{\pi n}\int_0^1 d\lambda \int_0^\infty \mathrm{Im}\chi^{\mathbf{g}\mathbf{g}}(\mathbf{q},\omega;\lambda)d\omega - 1\Big] \\ &= \frac{n}{2}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V(\mathbf{q}+\mathbf{g})\Big[-\frac{1}{\pi n}\int_0^1 d\lambda \int_0^\infty \mathrm{Re}\chi^{\mathbf{g}\mathbf{g}}(\mathbf{q},i\omega;\lambda)d\omega - 1\Big] \\ &= \frac{n}{2}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V(\mathbf{q}+\mathbf{g})\Big[-\frac{1}{\pi n}\int_0^1 d\lambda \int_0^\infty \chi^{\mathbf{g}\mathbf{g}}(\mathbf{q},i\omega;\lambda)d\omega - 1\Big]. \end{aligned} \tag{85}$$

In the last two expressions above, the integral along the real axis is rotated to the imaginary axis using the contour deformation, which is justified below.

The response function can be expressed in the entire complex plane using the spectral representation,

$$\chi(z) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{Im}\chi(\omega)}{z-\omega} d\omega. \tag{86}$$

In TBG, the density response function satisfies (as in Eq. (44))

$$\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) = \left[\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},-\omega)\right]^*, \tag{87}$$

therefore

$$\begin{aligned} \tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},z) &= -\frac{1}{\pi}\int_0^\infty \left[\frac{\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},-\omega)}{z+\omega} + \frac{\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega)}{z-\omega}\right] d\omega \\ &= -\frac{1}{\pi}\int_0^\infty \left[\frac{-\mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},\omega)}{z+\omega} + \frac{\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega)}{z-\omega}\right] d\omega, \\ \tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},z) &= -\frac{1}{\pi}\int_0^\infty \left[\frac{\mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},-\omega)}{z+\omega} + \frac{\mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},\omega)}{z-\omega}\right] d\omega \\ &= -\frac{1}{\pi}\int_0^\infty \left[\frac{-\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega)}{z+\omega} + \frac{\mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},\omega)}{z-\omega}\right] d\omega. \end{aligned} \tag{88}$$

Combine $\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},z)$ and $\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},z)$,

$$\begin{aligned} \tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},z) + \tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},z) &= -\frac{1}{\pi}\int_0^\infty \Big(\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) + \mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},\omega)\Big)\Big(\frac{1}{z-\omega} - \frac{1}{z+\omega}\Big) d\omega \\ &= -\frac{1}{\pi}\int_0^\infty \Big(\mathrm{Im}\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q},\omega) + \mathrm{Im}\tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q},\omega)\Big) \frac{2\omega(z_1^2 - z_2^2 - \omega^2) - i4z_1z_2\omega}{(z_1^2+z_2^2-\omega^2)^2 + 4z_2^2\omega^2} d\omega, \end{aligned} \tag{89}$$

where $z = z_1 + iz_2$. Along the imaginary axis, the imaginary part of Eq. (89) vanishes. Therefore the integral along the real axis can be rotated to the imaginary axis:

$$\begin{aligned}\int_0^\infty \Big(\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q}, i\omega) + \tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q}, i\omega)\Big)d\omega &= \int_0^\infty \mathrm{Re}\Big(\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q}, i\omega) + \tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q}, i\omega)\Big)d\omega \\ &= \int_0^\infty \mathrm{Im}\Big(\tilde{\chi}^{\mathbf{g}\mathbf{g}'}(\mathbf{q}, \omega) + \tilde{\chi}^{-\mathbf{g}-\mathbf{g}'}(-\mathbf{q}, \omega)\Big)d\omega\end{aligned} \tag{90}$$

Return to the xc energy, since only the diagonal in $\mathbf{g}$ elements of matrix $\boldsymbol{\chi}(\mathbf{q}, i\omega; \lambda)$ is relevant in Eq. (85), and the Coulomb matrix $\mathbf{V}(\mathbf{q})$ is diagonal in $\mathbf{g}$, the exchange-correlation energy can also be written in the matrix product form,

$$E_{xc} = \frac{n}{2}\sum_{\mathbf{q}}{}' \Big[-\frac{1}{\pi n}\int_0^1 d\lambda \int_0^\infty d\omega \mathrm{Tr}\big(\mathbf{V}(\mathbf{q})\boldsymbol{\chi}(\mathbf{q}, i\omega; \lambda)\big) - \mathrm{Tr}\big(\mathbf{V}(\mathbf{q})\big)\Big]. \tag{91}$$

The coupling-constant-dependent density response function in Eq. (85) and Eq. (91) is approximated, within the RPA, with

$$\begin{aligned}\chi(\lambda) &= \tilde{\chi}_H(1 - \lambda V \tilde{\chi}_H)^{-1} \\ &= \tilde{\chi}_H + \lambda\tilde{\chi}_H V \tilde{\chi}_H (1 - \lambda V \tilde{\chi}_H)^{-1},\end{aligned} \tag{92}$$

where $\tilde{\chi}_H$ is the proper density response function of the self-consistent Hartree approximation. The expression of $\tilde{\chi}_H$ of a specific flavor is shown in Eq. (49).

The exchange energy E_x is the first order contribution in Eq. (85) and Eq. (91), *i.e.* arising from the first term in Eq. (92)

$$\begin{aligned}E_x &= \frac{n}{2}\sum_{\mathbf{q},\mathbf{g}}{}' V(\mathbf{q}+\mathbf{g})\Big[-\frac{1}{\pi n}\int_0^\infty \tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega)d\omega - 1\Big] \\ &= \frac{n}{2}\sum_{\mathbf{q}}{}' \Big[-\frac{1}{\pi n}\mathrm{Tr}\big(\mathbf{V}(\mathbf{q})\tilde{\boldsymbol{\chi}}_H(\mathbf{q})\big) - \mathrm{Tr}\big(\mathbf{V}(\mathbf{q})\big)\Big].\end{aligned} \tag{93}$$

Using the Lindhard formula, the diagonal elements of $\tilde{\chi}_H$ is

$$\tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega) = \frac{1}{A}\sum_{n,m,\mathbf{k}}\left(\frac{\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}}}{\omega^2 + (\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})^2} - i\frac{\omega}{\omega^2 + (\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})^2}\right)(f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}})\Big|\sum_{\alpha,\mathbf{g}_1} z^n_{\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^m_{\alpha,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}+\mathbf{q})\Big|^2 \tag{94}$$

and its real and imaginary parts are respectively

$$\begin{aligned}\mathrm{Re}\tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega) &= \frac{1}{A}\sum_{n,m,\mathbf{k}}\frac{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})(f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}})}{\omega^2 + (\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})^2}\Big|\sum_{\alpha,\mathbf{g}_1} z^n_{\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^m_{\alpha,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}+\mathbf{q})\Big|^2, \\ \mathrm{Im}\tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega) &= -\frac{1}{A}\sum_{n,m,\mathbf{k}}\frac{\omega(f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}})}{\omega^2 + (\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})^2}\Big|\sum_{\alpha,\mathbf{g}_1} z^n_{\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^m_{\alpha,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}+\mathbf{q})\Big|^2.\end{aligned} \tag{95}$$

$\varepsilon_{n\mathbf{k}}$, $z^n(\mathbf{k})$ are eigen-energies and eigenvectors of the self-consistent Hartree approximation. Using the integration equality

$$\int_0^\infty d\omega\frac{1}{\omega^2 + a^2} = \frac{1}{a}\arctan\left(\frac{\omega}{a}\right)\Big|_0^\infty = \frac{\pi}{2|a|}, \tag{96}$$

the frequency integration of the real part of $\tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega)$ can be done analytically:

$$\int_0^\infty \mathrm{Re}\tilde{\chi}_H^{\mathbf{g}\mathbf{g}}(\mathbf{q}, i\omega)d\omega = \frac{\pi}{2A}\sum_{n,m,\mathbf{k}}\frac{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}})(f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}})}{|\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}}|}\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle. \tag{97}$$

Rewrite the Fermi-Dirac occupation difference

$$f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}} = f_{n\mathbf{k}}(1 - f_{m\mathbf{k}+\mathbf{q}}) - (1 - f_{n\mathbf{k}})f_{m\mathbf{k}+\mathbf{q}} \tag{98}$$

and it is clear that

$$\frac{\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}}}{|\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}}|} = \begin{cases} -1, \text{ if } f_{n\mathbf{k}} = 1 \text{ and } f_{m\mathbf{k}+\mathbf{q}} = 0, \\ 1, \text{ if } f_{n\mathbf{k}} = 0 \text{ and } f_{m\mathbf{k}+\mathbf{q}} = 1. \end{cases} \tag{99}$$

Then Eq. (97) becomes

$$\begin{aligned}\int_0^\infty \mathrm{Re}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q},i\omega)d\omega &= -\frac{\pi}{2A}\sum_{n,m,\mathbf{k}}\left[f_{n\mathbf{k}}(1-f_{m\mathbf{k}+\mathbf{q}})+(1-f_{n\mathbf{k}})f_{m\mathbf{k}+\mathbf{q}}\right]\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle \\ &= -\frac{\pi}{2A}\sum_{n,m,\mathbf{k}}(f_{n\mathbf{k}}+f_{m\mathbf{k}+\mathbf{q}}-2f_{n\mathbf{k}}f_{m\mathbf{k}+\mathbf{q}})\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle.\end{aligned} \tag{100}$$

The first two terms are simply total occupation number

$$\begin{aligned}\sum_{n,m,\mathbf{k}} f_{n\mathbf{k}}\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle &= \sum_{n,\mathbf{k}} f_{n\mathbf{k}} = N, \\ \sum_{n,m,\mathbf{k}} f_{m\mathbf{k}+\mathbf{q}}\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle &= \sum_{m,\mathbf{k}} f_{m\mathbf{k}+\mathbf{q}} = N.\end{aligned} \tag{101}$$

Equation (100) becomes

$$\int_0^\infty \mathrm{Re}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q},i\omega)d\omega = -\pi n + \frac{\pi}{A}\sum_{n,m,\mathbf{k}} f_{n\mathbf{k}}f_{m\mathbf{k}+\mathbf{q}}\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|m\mathbf{k}+\mathbf{q}\rangle\langle m\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}}|n\mathbf{k}\rangle. \tag{102}$$

Substitute into the exchange energy Eq. (93), the $-\pi n$ term above cancels exactly with the self-interacting term, and the exchange energy is simply

$$\begin{aligned}E_x &= -\frac{1}{2A}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V(\mathbf{q}+\mathbf{g})\sum_{n,m,\mathbf{k}} f_{n\mathbf{k}}f_{m\mathbf{k}+\mathbf{q}}\left|\sum_{\alpha,\mathbf{g}_1} z^n_{\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^m_{\alpha,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}+\mathbf{q})\right|^2 \\ &= -\frac{1}{2A}\sideset{}{'}\sum_{\mathbf{q},\mathbf{g}} V(\mathbf{q}+\mathbf{g})\sum_{n,m,\mathbf{k}} f_{n\mathbf{k}}f_{m\mathbf{k}+\mathbf{q}}\sum_{\alpha,\beta,\mathbf{g}_1,\mathbf{g}_2}\bar{\rho}^n_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2}(\mathbf{k})\rho^m_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q}).\end{aligned} \tag{103}$$

This exchange energy looks like the HF exchange but it is not since it is calculated over the SCH wavefunctions instead of the self-consistent HF wavefunctions.

On the other hand, the correlation energy is given by higher order contributions in RPA. Use

$$\mathrm{Tr}(\mathbf{V}\chi) = \mathrm{Tr}(\sqrt{\mathbf{V}}\chi\sqrt{\mathbf{V}}), \tag{104}$$

the correlation energy is

$$\begin{aligned}E_c &= -\frac{1}{2\pi}\sideset{}{'}\sum_{\mathbf{q}}\int_0^1 d\lambda\int_0^\infty d\omega\mathrm{Tr}\left[\lambda(\sqrt{\mathbf{V}}\tilde{\chi}_H\sqrt{\mathbf{V}})^2(1-\lambda\sqrt{\mathbf{V}}\tilde{\chi}_H\sqrt{\mathbf{V}})^{-1}\right] \\ &= \frac{1}{2\pi}\sideset{}{'}\sum_{\mathbf{q}}\int_0^\infty d\omega\mathrm{Tr}\left[\sqrt{\mathbf{V}}\tilde{\chi}_H\sqrt{\mathbf{V}}+\ln(1-\sqrt{\mathbf{V}}\tilde{\chi}_H\sqrt{\mathbf{V}})\right].\end{aligned} \tag{105}$$

For any diagonalizable and non-singular (invertible) matrix $\mathbf{A}$, its logarithm is

$$\ln\mathbf{A} = \mathbf{v}(\ln\boldsymbol{\lambda})\mathbf{v}^{-1}, \tag{106}$$

where $\boldsymbol{\lambda}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$,

$$\boldsymbol{\lambda} = \mathbf{v}^{-1}\mathbf{A}\mathbf{v}, \tag{107}$$

and $\mathbf{v}$ is the matrix with eigenvectors in each column. The trace of $\ln\mathbf{A}$

$$\begin{aligned}\mathrm{Tr}\ln\mathbf{A} &= \mathrm{Tr}\left[\mathbf{v}(\ln\boldsymbol{\lambda})\mathbf{v}^{-1}\right] \\ &= \mathrm{Tr}\left[(\ln\boldsymbol{\lambda})\mathbf{v}^{-1}\mathbf{v}\right] \\ &= \mathrm{Tr}\ln\boldsymbol{\lambda} \\ &= \ln\det\mathbf{A}.\end{aligned} \tag{108}$$

The correlation energy

$$
\begin{aligned}
E_c &= \frac{1}{2\pi} \sum_{\mathbf{q}}{}' \int_0^\infty d\omega \mathrm{Tr}\Big[\mathbf{V}(\mathbf{q})\tilde{\chi}_H(\mathbf{q}, i\omega) + \ln\left(1 - \sqrt{\mathbf{V}(\mathbf{q})}\tilde{\chi}_H(\mathbf{q}, i\omega)\sqrt{\mathbf{V}(\mathbf{q})}\right)\Big] \\
&= \frac{1}{2\pi} \sum_{\mathbf{q}}{}' \int_0^\infty d\omega \Big[\mathrm{Tr}\big(\mathbf{V}(\mathbf{q})\tilde{\chi}_H(\mathbf{q}, i\omega)\big) + \sum_i \ln \lambda_i(\mathbf{q}, i\omega)\Big],
\end{aligned}
\tag{109}
$$

where $\lambda_i(\mathbf{q}, i\omega)$ is the i-th eigenvalue of $(1 - \sqrt{\mathbf{V}}\tilde{\chi}_H\sqrt{\mathbf{V}})$.

VIII. EXCHANGE ENERGY REGULARIZATION

The exchange energy in Eq. (103) must be regularized to deal with the negative energy sea of the Dirac model. In the main text we denote E_x as the regularized exchange energy

$$
\begin{aligned}
E_x &= -\frac{1}{2A} \sum_{\mathbf{q},\mathbf{g}}{}' V(\mathbf{q}+\mathbf{g}) \sum_{\substack{\mathbf{k},\mathbf{g}_1,\mathbf{g}_2 \\ \alpha,\beta}} \left[\bar{\rho}_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2}(\mathbf{k})\rho_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q}) - \bar{\rho}^0_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2}(\mathbf{k})\rho^0_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q})\right] \\
&= -\frac{1}{2A} \sum_{\mathbf{q},\mathbf{g}}{}' V(\mathbf{q}+\mathbf{g}) \sum_{\substack{\mathbf{k},\mathbf{g}_1,\mathbf{g}_2 \\ \alpha,\beta}} \left[\delta\bar{\rho}(\mathbf{k}) + 2\bar{\rho}^0(\mathbf{k})\right]_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2} \delta\rho_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q}).
\end{aligned}
\tag{110}
$$

ρ^0 is the density matrix of the charge neutral decoupled bilayer and therefore it's diagonal in **g**'s:

$$
\rho^0_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2}(\mathbf{k}) = \delta_{\mathbf{g}_1\mathbf{g}_2} \sum_{n\in v} z^0_{n,\alpha,\mathbf{g}_1}(\mathbf{k})\bar{z}^0_{n,\beta,\mathbf{g}_2}(\mathbf{k}).
\tag{111}
$$

Summing over two valence bands from top and bottom layer Dirac cones, the 4 × 4 density matrix with an explicit **g** is

$$
\rho^0_{\mathbf{g};\mathbf{g}}(\mathbf{k}) \equiv \rho^0(\mathbf{k}+\mathbf{g}) = \frac{1}{2}\begin{pmatrix}
1 & -e^{-i(\theta_{\mathbf{k}+\mathbf{g}-\mathbf{K}_1}-\theta/2)} & 0 & 0 \\
-e^{i(\theta_{\mathbf{k}+\mathbf{g}-\mathbf{K}_1}-\theta/2)} & 1 & 0 & 0 \\
0 & 0 & 1 & -e^{-i(\theta_{\mathbf{k}+\mathbf{g}-\mathbf{K}_2}+\theta/2)} \\
0 & 0 & -e^{i(\theta_{\mathbf{k}+\mathbf{g}-\mathbf{K}_2}+\theta/2)} & 1
\end{pmatrix}.
\tag{112}
$$

Therefore, the second part of the second line of Eq. (110) is equivalent to

$$
E_x^{(2)} = -\frac{1}{A} \sum_{\substack{\mathbf{k}'\in\mathrm{MBZ} \\ \mathbf{k}\in k_c\sim 1/a}} V(\mathbf{k}'-\mathbf{k}+\mathbf{g}) \sum_{\mathbf{g}_1,\alpha,\beta} \bar{\rho}^0_{\alpha\beta}(\mathbf{k}+\mathbf{g}_1)\delta\rho_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_1+\mathbf{g}}(\mathbf{k}')
\tag{113}
$$

Because the diagonal terms of ρ^0 contribute a constant energy shift which is proportional to the occupation number at $\mathbf{k}'$, only off-diagonal terms of ρ^0 matter and it's

$$
\begin{aligned}
E_x^{(2)} = \frac{2\pi e^2}{\epsilon}\frac{1}{8\pi} \sum_{\mathbf{k}'\in\mathrm{MBZ},\mathbf{g}} \Big[& e^{i\xi(\theta_{\mathbf{k}'+\mathbf{g}-\mathbf{K}_1}-\theta/2)}[\delta\rho_{\mathbf{g};\mathbf{g}}]_{12}(\mathbf{k}')|\mathbf{k}'+\mathbf{g}-\mathbf{K}_1| \ln\frac{k_c}{|\mathbf{k}'+\mathbf{g}-\mathbf{K}_1|} \\
&+ e^{i\xi(\theta_{\mathbf{k}'+\mathbf{g}-\mathbf{K}_2}+\theta/2)}\delta[\rho_{\mathbf{g};\mathbf{g}}]_{34}(\mathbf{k}')|\mathbf{k}'+\mathbf{g}-\mathbf{K}_2| \ln\frac{k_c}{|\mathbf{k}'+\mathbf{g}-\mathbf{K}_2|} + h.c\Big]
\end{aligned}
\tag{114}
$$

Therefore the regularized exchange energy is

$$
E_x = -\frac{1}{2A} \sum_{\mathbf{q},\mathbf{g}}{}' V(\mathbf{q}+\mathbf{g}) \sum_{\substack{\mathbf{k},\mathbf{g}_1,\mathbf{g}_2 \\ \alpha,\beta}} \delta\bar{\rho}_{\alpha,\mathbf{g}_1;\beta,\mathbf{g}_2}(\mathbf{k})\delta\rho_{\alpha,\mathbf{g}_1+\mathbf{g};\beta,\mathbf{g}_2+\mathbf{g}}(\mathbf{k}+\mathbf{q}) + E_x^{(2)}.
\tag{115}
$$

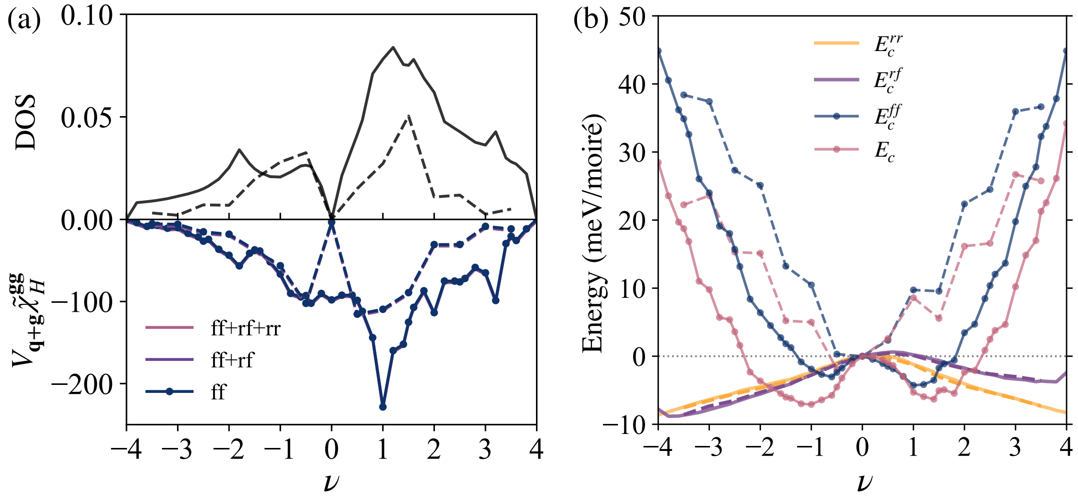


FIG. 7. (a) Density of state in unit of meV^{-1}·nm^{-2} and $V_{\mathbf{q}+\mathbf{g}}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q}, i\omega = 0)$ as a function of ν, with $\mathbf{g} = 0$ and $\mathbf{q} = (-\sqrt{3}/2, -1/2)g_{\mathrm{M}}/\sqrt{3}$, where g_{M} is the length of moiré primitive reciprocal lattice vector. Different colors in $V_{\mathbf{q}+\mathbf{g}}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q}, i\omega = 0)$ plot label distinct particle-hole excitations as indicated by the legend. (a) shows that excitations between flat bands (ff) dominate when the flat bands are not empty or fully occupied. (b) The correlation energy E_c as a function of ν. E_c^{ff}, E_c^{rf} and E_c^{rr} are contributions to the correlation energy from excitations between flat bands, between remote bands and flat bands and between remote bands respectively. E_c^{rr} and E_c^{rf} are independent of flavor polarizations. In both figures, solid lines represent flavor paramagnetic states and dashed lines represent flavor fully polarized states, for example at $\nu = 2$, $\nu_f = (1, 1, 0, 0)$.

IX. SCREENING EFFECTS FROM REMOTE BANDS

In Fig. 7(a) we show the density of state (DOS) and $V_{\mathbf{q}+\mathbf{g}}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q}, i\omega = 0)$ as a function of ν, for the specific $\mathbf{g}$ and $\mathbf{q}$. For both flavor paramagnetic state (solid lines) and flavor polarized state (dashed lines), electron-hole excitations between flat bands (ff) dominate over excitations between remote and flat bands (rf) and between remote bands (rr), as long as flat bands are not entirely empty or fully occupied. In Fig. 7(b), contributions to the correlation energy (E_c) from electron-hole excitations between flat bands (E_c^{ff}), between remote and flat bands (E_c^{rf}) and between remote bands (E_c^{rr}) are separately shown. E_c^{ff} dominates and is responsible for the tendency of E_c with respect to ν. By comparing correlation energies of flavor paramagnetic state (solid lines) with flavor polarized state (dashed lines), E_c^{rr} and E_c^{rf} are almost independent of flavor polarizations.

Figure 8 shows $V_{\mathbf{q}+\mathbf{g}}\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q}, i\omega)$ at $\nu = -1$ as a function of frequency, for the specific $\mathbf{g}$ and $\mathbf{q}$. It illustrates that the xc energy are dominated by diagonal elements of $\mathbf{V}\tilde{\boldsymbol{\chi}}_H$ with the smallest $|\mathbf{g}|$ and by excitations between flat bands (ff).

Figure 9 shows the sum of eigenvalues of $\ln(1 - \mathbf{V}\tilde{\boldsymbol{\chi}})$, *i.e.* the second term of Eq. (5), as a function of q and unitless frequency $\tilde{\omega}$. Again the correlation effect is dominated by ff excitations.

X. ENERGIES OF COMPETING BROKEN FLAVOR-SYMMETRY STATES

Figure 10 schematically summarizes our findings in Table I by plotting the total energy (Fig. 10(a)) and xc energy (Fig. 10(b)) relative to the flavor paramagnetic state for various flavor polarizations (y-axis) as a function of filling factor ν (x-axis). $E_{tot}^S - E_{tot} > 0$ indicates the ground state favors flavor polarized states. $E_{xc}^S - E_{xc} > 0$ indicates the xc effect, when the single-quasiparticle energy is ignored, favors flavor polarized states. Figure 10(b) clearly shows that the xc effect predicts broken flavor symmetry for $|\nu| \geq 1$ on both electron- and hole-doped sides and is stronger on the electron-doped side. After including the single-quasiparticle energy which in general prefers unpolarized state, however, the ground state is predicted to be flavor paramagnetic on hole-doped side. This is a result of the fact that on the hole-doped side the Hartree energy is twice stronger and the xc effect is weaker than that on the electron-doped side.

We further explore the effects of C_2T symmetry breaking on energies, as summarized in Table II. We find that flavor ferromagnetic states are favored at all filling factors including those near CN, consistent with experiments that hBN alignment favors broken symmetry states.

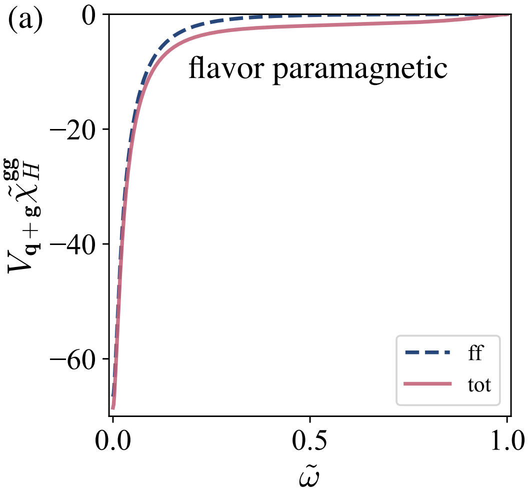


FIG. 8. $V(\mathbf{q}+\mathbf{g})\tilde{\chi}_H^{\mathbf{gg}}(\mathbf{q},i\omega)$ at $\nu=-1$ as a function of the unitless frequency $\tilde{\omega}$, with $\mathbf{g}=0$ and $\mathbf{q}=(-\sqrt{3}/2,-1/2)g_{\mathrm{M}}/\sqrt{3}$. $\tilde{\omega}=\omega/(\omega+\omega_0)$, where $\omega_0=30$ meV is chosen to be around the size of flat-band bandwidth. Electron-hole excitations between flat bands (ff) play a dominant role in the correlation effect.

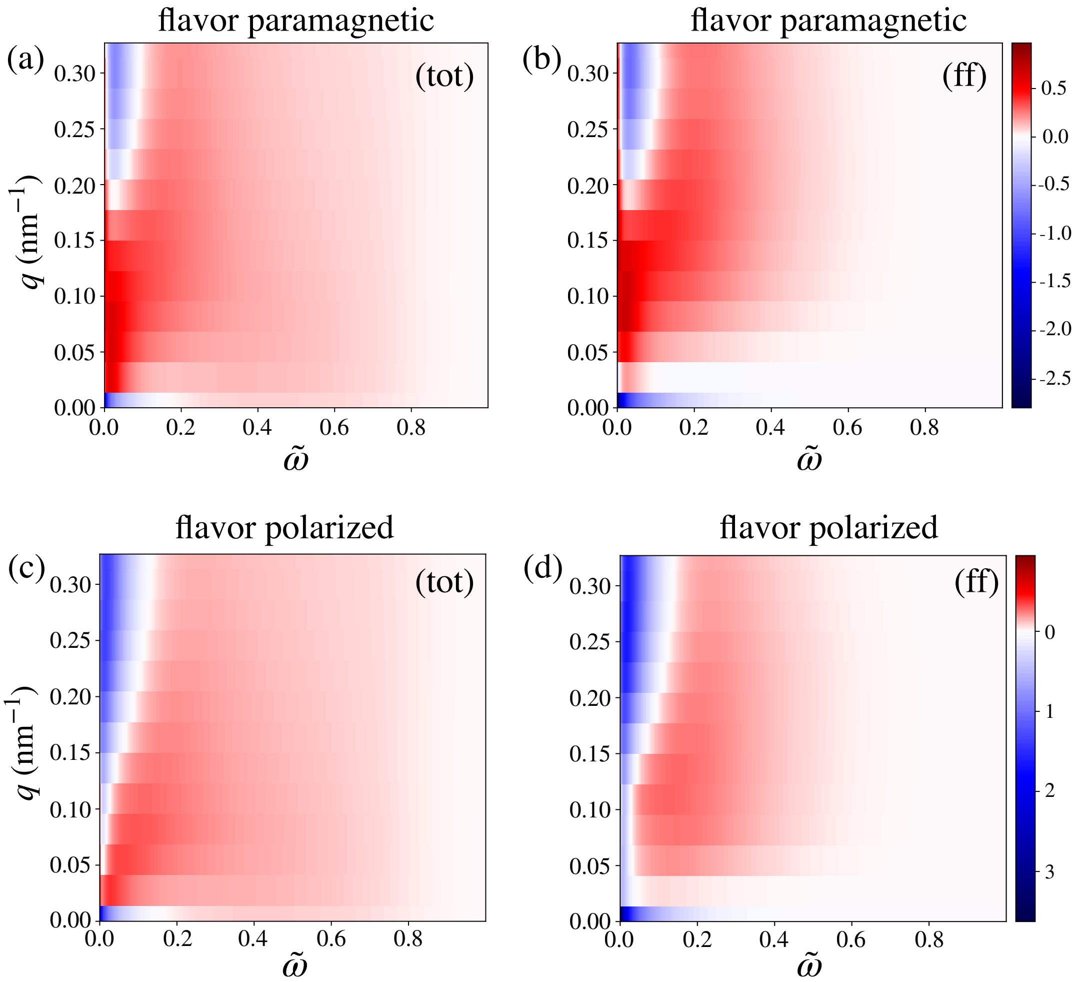


FIG. 9. $\sum_j \ln(1-\lambda_j)$ with respect to q and unitless frequency $\tilde{\omega}$, where λ_j is the j-th eigenvalue of $\mathbf{V}\tilde{\chi}_H$.

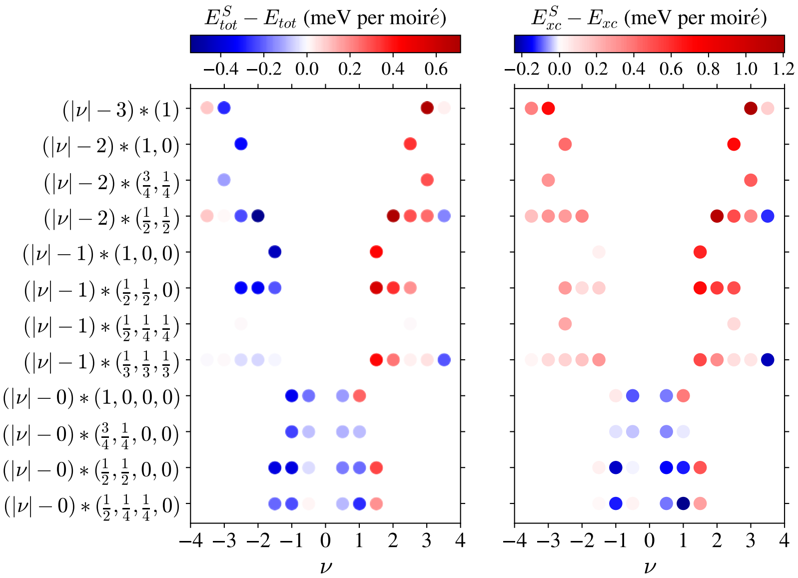


FIG. 10. (a) Total energy E_{tot} and (b) xc energy (E_{xc}) relative to those of the flavor paramagnetic state (E^S_{tot}, E^S_{xc}), for various flavor-polarized states (y-axis) as a function of ν (x-axis). The y-axis is labeled by $(|\nu| - n) * (f_1, ...)$, where the integer n indicates the number of flavors that are fully occupied or empty and fractions inside the right bracket, multiplied by $(|\nu| - n)$, denote filling factors of the remaining partially occupied flavors. For example, the data point at $\nu = -1$ and $(|\nu| - 0) * (1/2, 1/4, 1/4, 0)$ represents the filling factors of four flavors are $-(1/2, 1/4, 1/4, 0)$ respectively. $E^S_{tot} - E_{tot} > 0$ indicates the ground state favors flavor polarized states and $E^S_{xc} - E_{xc} > 0$ indicates the xc energy favors flavor polarized states.

XI. MAGNETIC ANISOTROPY IN SU(4) FERROMAGNETS

MATBG has $SU(4)$ ferromagnetism because of its four degenerate spin-valley flavors. In the continuum model we employ the $SU(4)$ symmetry is reduced to $SU(2) \times SU(2) \times U(1)$ by the difference between two valley projected band Hamiltonians, which contribute explicitly to the model's $SU(4)$ magnetic anisotropy. Valley-exchange and spin-orbit interactions, which we neglect, also contribute to magnetic anisotropy. In our RPA approach, we neglect anisotropy by focusing only on flavor-dependent filling factors, the eigenvalues of the spin-valley density matrix. Indeed our explicit calculations assume that the ferromagnet's density matrix remains diagonal in the spin-valley representation because total valley occupation number remains a good quantum number, simplifying the use of a coupling-constant integral representation of the energy. There is in fact theoretical[22, 78–81] and experimental [82–85] work that intervalley coherence, which breaks valley number symmetry, is present in many MATBG ferromagnets. Our calculations make no effort to distinguish between ferromagnets that differ only in the orientation of the spin-valley magnetization and not in its magnitude as characterized by the differences between spin-valley density-matrix eigenvalues. There is evidence in recent experiments [85], in the form of the presence of sample-specific domain walls and vortices in various partial order parameters, that as in most conventional ferromagnets the energy scale associated with anisotropy is smaller than the energy scale associated with ordering.

ν	$(\nu_1, \nu_2, \nu_3, \nu_4)$	E_0	E_x	E_c	E_{xc}	E_{tot}	ν	$(\nu_1, \nu_2, \nu_3, \nu_4)$	E_0	E_x	E_c	E_{xc}	E_{tot}
-0.5	(-1/8, -1/8, -1/8, -1/8)	-5.477	6.097	-4.520	1.577	-3.900	0.5	(1/8, 1/8, 1/8, 1/8)	6.852	1.570	-0.930	**0.640**	**7.492**
	(-1/4, -1/8, -1/8, 0)	-5.458	5.534	-3.992	**1.542**	**-3.916**		(1/4, 1/8, 1/8, 0)	6.864	1.163	-0.443	0.720	7.584
	(-1/4, -1/4, 0, 0)	-5.437	5.417	-3.834	1.583	-3.854		(1/4, 1/4, 0, 0)	6.876	0.756	0.033	0.789	7.665
	(-3/8, -1/8, 0, 0)	-5.426	4.650	-3.039	1.611	-3.815		(3/8, 1/8, 0, 0)	6.882	0.006	0.702	0.708	7.590
	(-1/2, 0, 0, 0)	-5.386	3.586	-1.913	1.673	-3.713		(1/2, 0, 0, 0)	6.905	-1.863	2.579	0.716	7.621
-1.0	(-1/4, -1/4, -1/4, -1/4)	-10.236	11.187	-7.073	4.114	**-6.122**	1.0	(1/4, 1/4, 1/4, 1/4)	14.774	8.384	-5.293	3.091	17.865
	(-1/2, -1/4, -1/4, 0)	-10.138	9.902	-5.659	4.243	-5.895		(1/2, 1/4, 1/4, 0)	14.806	6.313	-2.983	3.330	18.136
	(-1/2, -1/2, 0, 0)	-10.044	8.465	-4.162	4.303	-5.741		(1/2, 1/2, 0, 0)	14.833	4.721	-1.500	3.221	18.054
	(-3/4, -1/4, 0, 0)	-10.006	6.076	-1.946	4.130	-5.876		(3/4, 1/4, 0, 0)	14.848	1.755	1.349	3.104	17.952
	(-1, 0, 0, 0)	-9.816	-0.957	5.008	**4.051**	-5.765		(1, 0, 0, 0)	14.907	-5.881	8.586	**2.705**	**17.612**
-1.5	(-3/8, -3/8, -3/8, -3/8)	-13.764	11.482	-6.009	5.473	**-8.291**	1.5	(3/8, 3/8, 3/8, 3/8)	23.162	11.808	-5.054	6.754	29.916
	(-3/4, -3/8, -3/8, 0)	-13.559	7.287	-1.825	5.462	-8.097		(3/4, 3/8, 3/8, 0)	23.239	7.646	-1.159	6.487	29.726
	(-3/4, -3/4, 0, 0)	-13.335	3.008	2.418	5.426	-7.909		(3/4, 3/4, 0, 0)	23.328	3.450	2.815	6.265	29.593
	(-1, -1/6, -1/6, -1/6)	-13.450	1.424	3.749	**5.173**	-8.277		(1, 1/6, 1/6, 1/6)	23.261	2.345	3.882	6.227	29.488
	(-1, -1/4, -1/4, 0)	-13.410	1.267	4.071	5.338	-8.072		(1, 1/4, 1/4, 0)	23.297	1.984	4.065	**6.049**	**29.346**
	(-1, -1/2, 0, 0)	-13.266	0.215	5.212	5.427	-7.839		(1, 1/2, 0, 0)	23.355	0.570	5.556	6.126	29.481
-2.0	(-1/2, -1/2, -1/2, -1/2)	-15.756	9.684	-3.635	6.049	**-9.707**	2.0	(1/2, 1/2, 1/2, 1/2)	32.588	11.054	-2.341	8.713	41.301
	(-1, -1/3, -1/3, -1/3)	-15.530	1.340	4.537	5.877	-9.653		(1, 1/3, 1/3, 1/3)	32.681	4.395	3.984	8.379	41.060
	(-1, -1/2, -1/2, 0)	-15.314	0.227	5.721	5.948	-9.366		(1, 1/2, 1/2, 0)	32.793	1.151	6.991	8.142	40.935
	(-1, -1, 0, 0)	-14.862	-9.434	15.127	**5.693**	-9.169		(1, 1, 0, 0)	32.985	-8.577	16.158	**7.581**	**40.566**
-2.5	(-5/8, -5/8, -5/8, -5/8)	-16.095	2.991	3.595	6.586	-9.509	2.5	(5/8, 5/8, 5/8, 5/8)	43.627	7.017	2.494	9.511	53.138
	(-1, -1/2, -1/2, -1/2)	-15.919	0.494	5.962	6.456	-9.463		(1, 1/2, 1/2, 1/2)	43.685	3.354	6.068	9.422	53.107
	(-1, -3/4, -3/8, -3/8)	-15.844	-2.101	8.436	6.335	**-9.509**		(1, 3/4, 3/8, 3/8)	43.734	0.136	9.266	9.402	53.136
	(-1, -3/4, -2/4, -1/4)	-15.812	-2.242	8.691	6.449	-9.363		(1, 3/4, 2/4, 1/4)	43.746	-0.106	9.424	9.318	53.064
	(-1, -3/4, -3/4, 0,)	-15.483	-6.518	12.820	6.302	-9.181		(1, 3/4, 3/4, 0)	43.925	-4.951	13.961	9.010	52.935
	(-1, -1, -1/4, -1/4)	-15.582	-7.584	13.887	6.303	-9.279		(1, 1, 1/4, 1/4)	43.845	-6.466	15.462	8.996	52.841
	(-1, -1, -1/2, 0)	-15.361	-9.125	15.295	**6.170**	-9.191		(1, 1, 1/2, 0)	43.973	-7.790	16.593	**8.803**	**52.776**
-3.0	(-3/4, -3/4, -3/4, -3/4)	-14.731	-2.788	9.788	7.000	-7.731	3.0	(3/4, 3/4, 3/4, 3/4)	56.420	0.058	10.219	10.277	66.697
	(-1, -2/3, -2/3, -2/3)	-14.627	-5.009	11.905	6.896	-7.731		(1, 2/3, 2/3, 2/3)	56.443	-1.905	12.109	10.204	66.647
	(-1, -1, -1/2, -1/2)	-14.425	-7.827	14.517	6.690	**-7.735**		(1, 1, 1/2, 1/2)	56.523	-6.657	16.577	9.920	66.443
	(-1, -1, -3/4, -1/4)	-14.301	-10.579	17.277	6.698	-7.603		(1, 1, 3/4, 1/4)	56.585	-9.340	19.149	9.809	66.394
	(-1, -1, -1, 0)	-13.748	-17.295	23.588	**6.293**	-7.455		(1, 1, 1, 0)	56.897	-17.628	26.698	**9.070**	**65.967**
-3.5	(-7/8, -7/8, -7/8, -7/8)	-11.595	-11.455	18.756	7.301	-4.294	3.5	(7/8, 7/8, 7/8, 7/8)	71.002	-10.902	21.286	10.384	81.386
	(-1, -5/6, -5/6, -5/6)	-11.566	-11.142	18.420	7.278	-4.288		(1, 5/6, 5/6, 5/6)	71.014	-10.676	21.261	10.585	81.599
	(-1, -1, -3/4, -3/4)	-11.501	-12.790	19.905	7.115	**-4.386**		(1, 1, 3/4, 3/4)	71.028	-12.365	22.870	10.505	81.533
	(-1, -1, -1, -1/2)	-11.322	-15.305	22.244	**6.939**	-4.383		(1, 1, 1, 1/2)	71.103	-15.514	25.758	**10.244**	**81.347**

TABLE I. The exchange E_x, the correlation E_c and the total E_{tot} energies of competing flavor-symmetry broken states, calculated using $\epsilon_{\rm BN} = 5.1$. The lowest E_{tot} and E_{xc} are marked bold at each ν. On the electron-doped side, the ground state prefers flavor paramagnetism for $|\nu| < 1.0$ and flavor polarization for $|\nu| \gtrsim 1.0$. On the hole-doped side, however, the ground state favors flavor paramagnetism for $|\nu| \lesssim 2.0$ and flavor polarization for $|\nu| \gtrsim 2.0$. Energies are in the unit of meV per moiré unit cell. The kinetic energy E_k, defined as $E_k = E_{band} - 2E_H$ where E_{band} and E_H are the band energy and the Hartree energy of the SCH quasi-particle bands respectively, is regularized by the kinetic energy of flavor paramagnetic state at each filling ν.

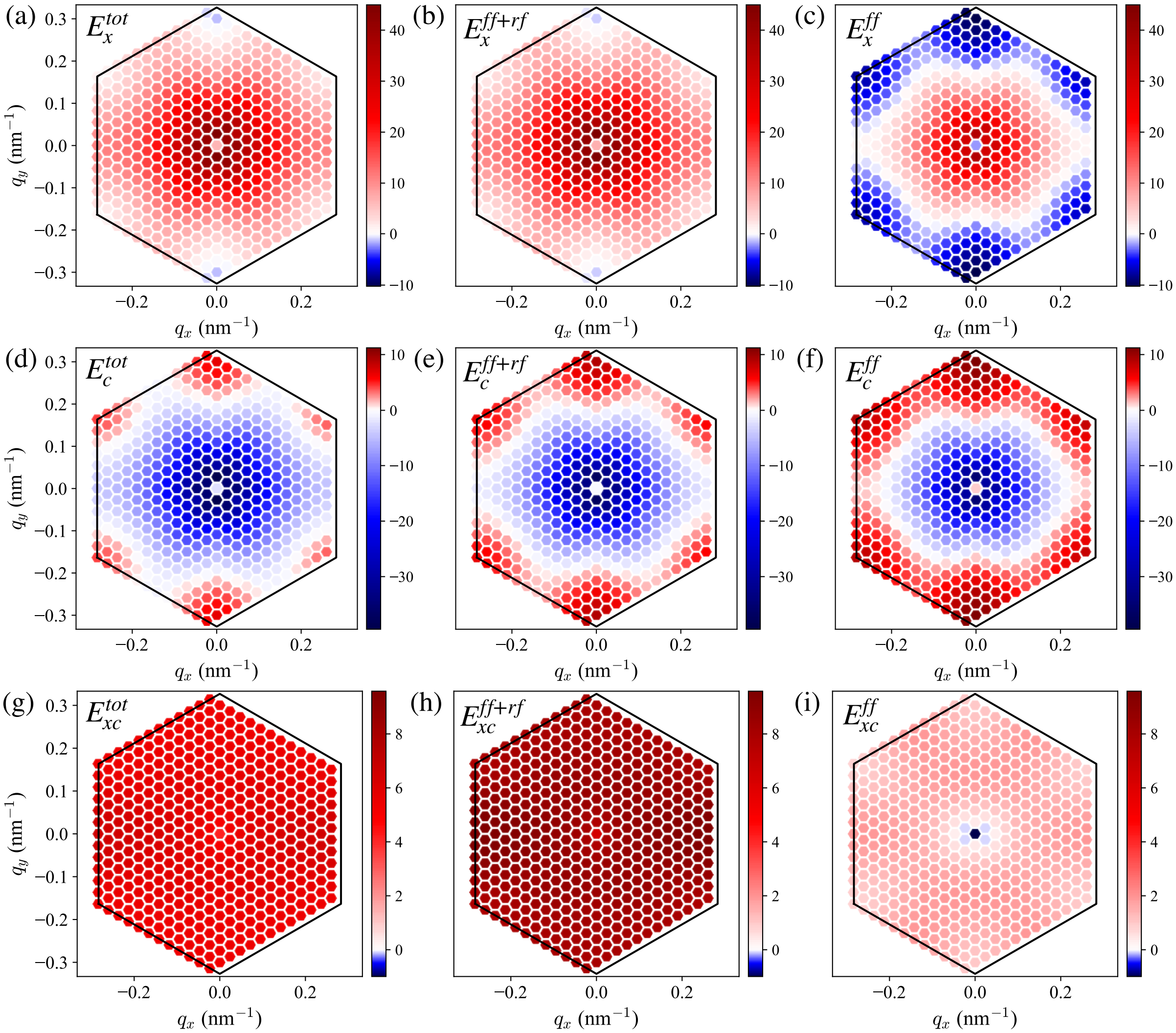


FIG. 11. Energies versus $\mathbf{q}$ of the paramagnetic state at $\nu = -1$. (a-c) The exchange energy including excitations between all SCH bands (E_x^{tot}), between flat bands and between flat and remote bands ($E_x^{\text{ff+rf}}$), between flat bands only (E_x^{ff}). (d-f) The correlation energy. (g-i) The xc energy.

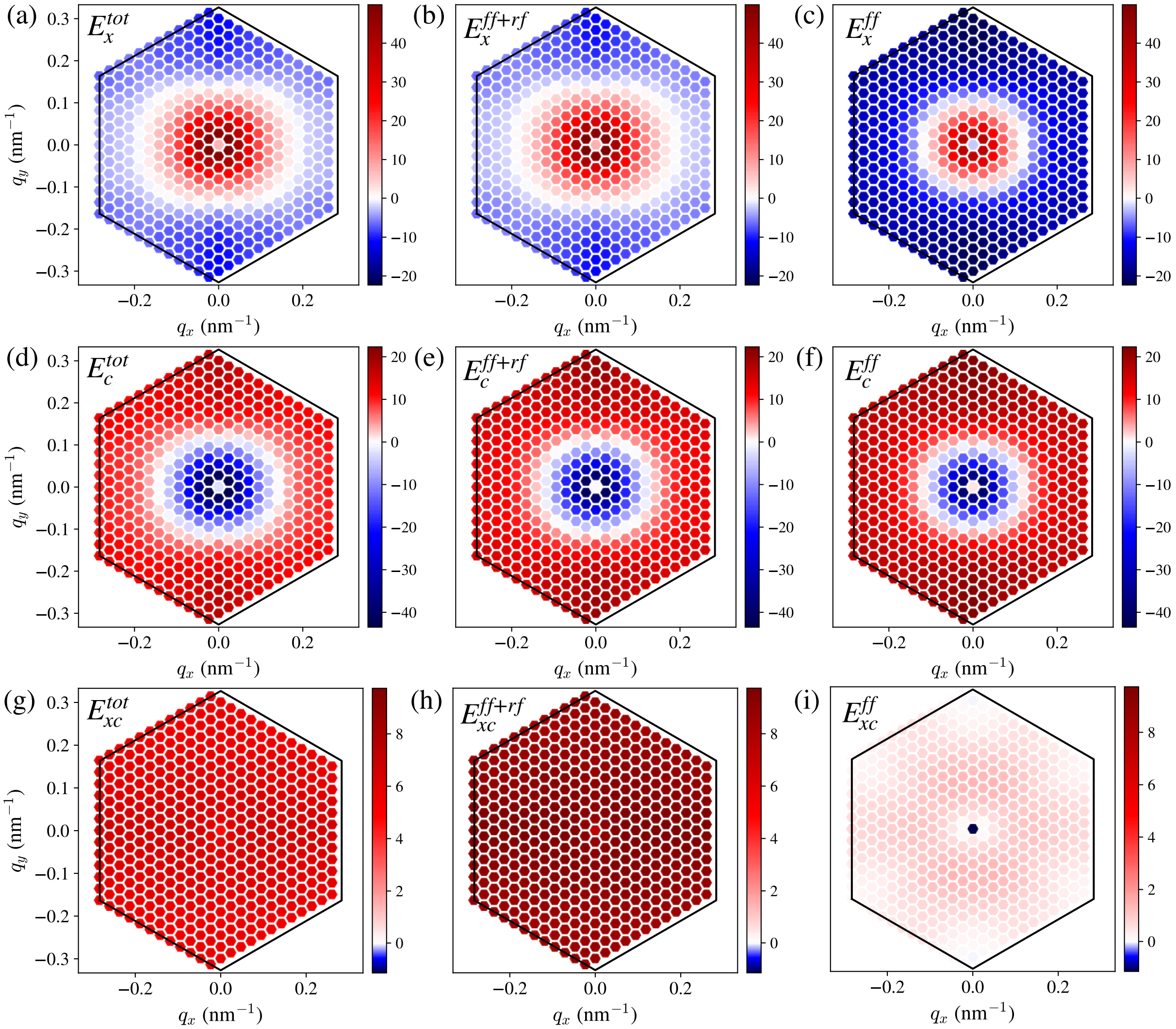


FIG. 12. Energies versus $\mathbf{q}$ of the polarized state at $\nu = -1$.

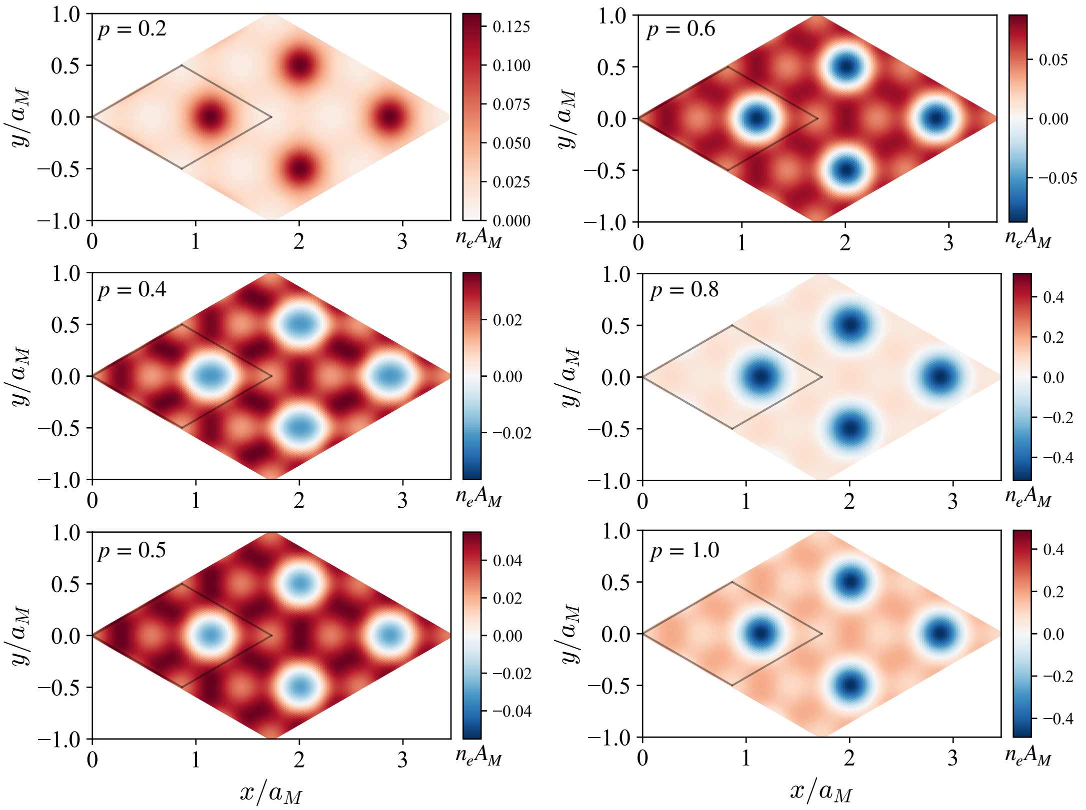


FIG. 13. Density distributions in real-space for $\nu = 0$, shown for various polarizations $p = 0.2, 0.4, 0.5, 0.6, 0.8, 1.0$. These correspond to Fig. 3(a) in the main text, where polarization p is defined. Displayed densities account for two flat band contributions and are depicted after subtraction of the $p = 0$ baseline. Compared to finite filling cases as shown in Fig. 14, densities for $\nu = 0$ are at least one order of magnitude smaller and therefore approximately uniform for all p. Density is in unit of $n_e A_\mathrm{M}$, where A_M is the area of moiré unit cell. The grey rhombus outlines the moiré unit cell.

ν	$(\nu_1, \nu_2, \nu_3, \nu_4)$	E_0	E_x	E_c	E_{xc}	E_{tot}	ν	$(\nu_1, \nu_2, \nu_3, \nu_4)$	E_0	E_x	E_c	E_{xc}	E_{tot}
-1.0	(-1/4, -1/4, -1/4, -1/4)	-4.119	25.398	-21.363	4.034	-0.084	1.0	(1/4, 1/4, 1/4, 1/4)	19.664	30.622	-26.822	3.800	23.464
	(-1/2, -1/4, -1/4, 0)	-3.980	23.401	-19.643	3.758	-0.221		(1/2, 1/4, 1/4, 0)	19.785	25.797	-21.969	3.828	23.613
	(-1/2, -1/2, 0, 0)	-3.809	21.189	-17.977	3.212	-0.597		(1/2, 1/2, 0, 0)	19.749	20.717	-17.055	3.662	23.411
	(-3/4, -1/4, 0, 0)	-3.768	17.368	-14.212	3.156	-0.612		(3/4, 1/4, 0, 0)	19.697	16.836	-13.436	3.400	23.096
	(-1, 0, 0, 0)	-3.500	6.145	-4.249	**1.897**	**-1.604**		(1, 0, 0, 0)	19.796	4.637	-2.015	**2.622**	**22.418**
-2.0	(-1/2, -1/2, -1/2, -1/2)	-3.919	41.452	-34.390	7.062	3.143	2.0	(1/2, 1/2, 1/2, 1/2)	43.469	41.064	-33.315	7.749	51.218
	(-1, -1/3, -1/3, -1/3)	-3.787	30.464	-23.639	6.826	3.039		(1, 1/3, 1/3, 1/3)	43.534	28.905	-21.296	7.609	51.144
	(-1, -1/2, -1/2, 0)	-3.106	27.850	-21.761	6.089	2.983		(1, 1/2, 1/2, 0)	43.903	26.511	-19.296	7.215	51.118
	(-1, -1, 0, 0)	-2.244	14.628	-10.070	**4.558**	**2.314**		(1, 1, 0, 0)	44.375	12.460	-6.424	**6.036**	**50.411**
-3.0	(-3/4, -3/4, -3/4, -3/4)	2.155	47.234	-38.426	8.808	10.963	3.0	(3/4, 3/4, 3/4, 3/4)	72.566	45.210	-33.703	11.507	84.073
	(-1, -2/3, -2/3, -2/3)	2.118	43.554	-34.843	8.711	**10.829**		(1, 2/3, 2/3, 2/3)	72.541	41.623	-30.052	11.570	84.111
	(-1, -1, -1/2, -1/2)	2.134	36.756	-27.537	9.219	11.353		(1, 1, 1/2, 1/2)	72.583	35.136	-23.817	11.320	83.903
	(-1, -1, -3/4, -1/4)	2.270	34.094	-25.249	8.846	11.116		(1, 1, 3/4, 1/4)	72.710	32.033	-20.865	11.168	**83.878**
	(-1, -1, -1, 0)	3.562	26.476	-18.317	**8.159**	11.721		(1, 1, 1, 0)	73.749	22.621	-12.401	**10.219**	83.968

TABLE II. Same as in Table I but with broken C_2T by using a massive Dirac Hamiltonian in the BM model. The mass terms in top and bottom layers are both 10 meV.

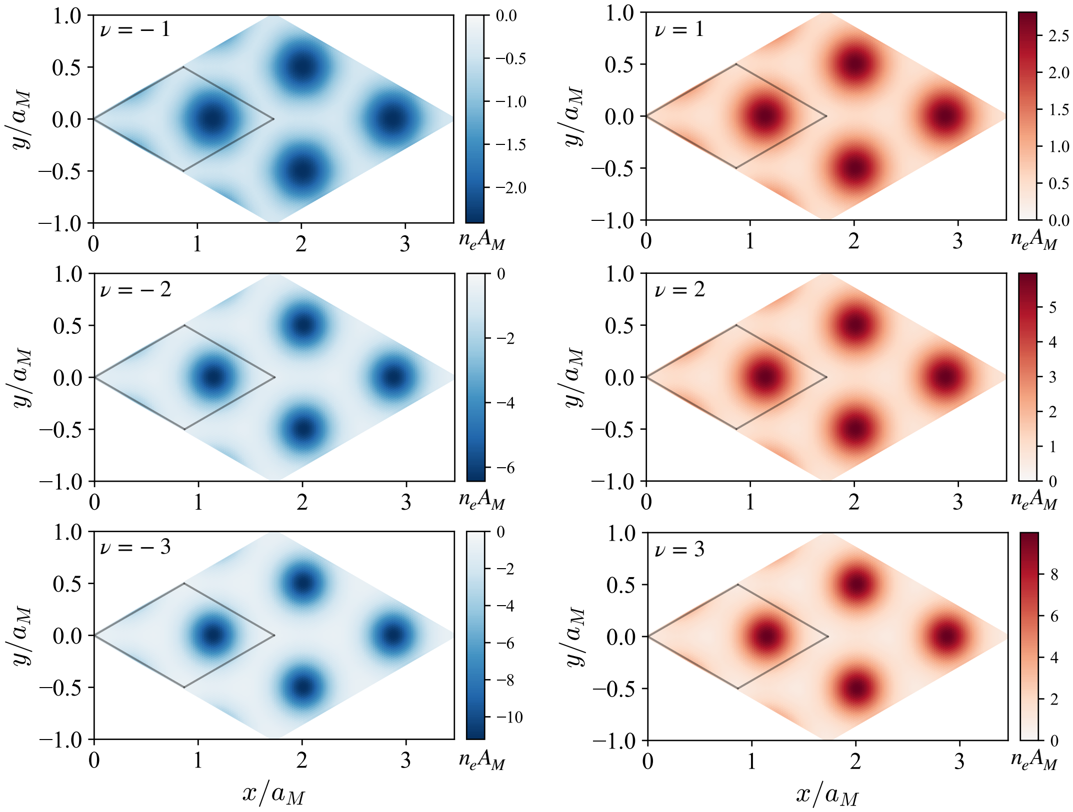


FIG. 14. Density distributions in real-space for the paramagnetic state, as a comparison to Fig. 13, displayed at different fillings $\nu = \pm1, \pm2, \pm3$.

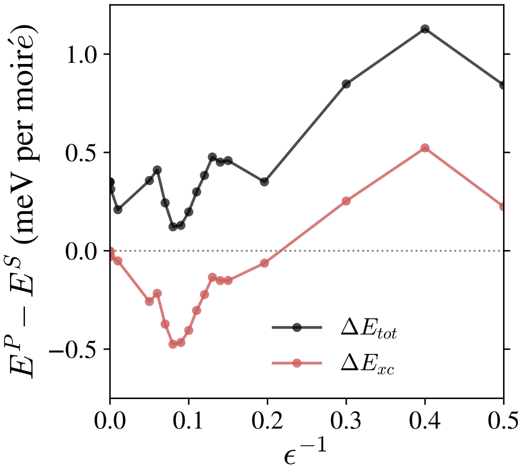


FIG. 15. The energy difference, defined as $\Delta E = E^P - E^S$, between flavor fully polarized state (P) and flavor paramagnetic state (S) at $\nu = -1$ as a function of Coulomb interaction strength ϵ^{-1}. ΔE_{tot} and ΔE_{xc} are energy difference of RPA total and xc energies respectively.

arXiv:2401.16384v2 [cond-mat.mes-hall] 7 Jun 2024

Theory of intrinsic acoustic plasmons in twisted bilayer graphene

Lorenzo Cavicchi,[1] Iacopo Torre,[2,3] Pablo Jarillo-Herrero,[4] Frank H. L. Koppens,[3,5] and Marco Polini[6,3]

[1]*Scuola Normale Superiore, I-56126 Pisa, Italy*
[2]*Departament de Física, Universitat Politècnica de Catalunya, Campus Nord B4-B5, 08034 Barcelona, Spain*
[3]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology,*
Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain
[4]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[5]*ICREA-Institució Catalana de Recerca i Estudis Avançats,*
Passeig de Lluís Companys 23, 08010 Barcelona, Spain
[6]*Dipartimento di Fisica dell'Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy*

(Dated: June 10, 2024)

We present a theoretical study of the intrinsic plasmonic properties of twisted bilayer graphene (TBG) as a function of the twist angle θ (and other microscopic parameters such as temperature and filling factor). Our calculations, which rely on the random phase approximation, take into account four crucially important effects, which are treated on equal footing: i) the layer-pseudospin degree of freedom, ii) spatial non-locality of the density-density response function, iii) crystalline local field effects, and iv) Hartree self-consistency. We show that the plasmonic spectrum of TBG displays a smooth transition from a strongly-coupled regime (at twist angles $\theta \lesssim 2^{\circ}$), where the low-energy spectrum is dominated by a weakly dispersive intra-band plasmon, to a weakly-coupled regime (for twist angles $\theta \gtrsim 2^{\circ}$) where an acoustic plasmon clearly emerges. This crossover offers the possibility of realizing tunable mid-infrared sub-wavelength cavities, whose vacuum fluctuations may be used to manipulate the ground state of strongly correlated electron systems.

I. INTRODUCTION

Parallel two-dimensional electron systems (P2DESs) have been at the center of a great deal of attention since they were theoretically proposed in 1975 as ideal setups for the study of superfluidity of spatially separated electrons and holes [1]. They have been experimentally fabricated by using two main experimental platforms: i) one based on GaAs/AlGaAs heterostructures realized by molecular beam epitaxy [2–5] and ii) one on atomically-thin 2D materials, such as graphene and transition-metal dichalcogenides (TMDs), produced by mechanical exfoliation [6]. These systems harbor a wide set of spectacular electrical phenomena, including Coulomb drag [7–12], exciton superfluidity in strong [13–17] and zero [18] magnetic fields, and broken symmetry states [19–24] driven by strong electron-electron interactions.

More recently, the many-body physics of P2DEs has been greatly enriched thanks to the discovery [25, 26] of correlated insulators and superconductors in twisted bilayer graphene (TBG). TBG [27–34] is a P2DES comprising two graphene sheets on top of each other, separated by a vertical distance d on the order on ≈ 0.3 nm, and rotated by a twist angle θ. In this system, inter-layer tunneling changes significantly as a function of θ, leading to a dramatic spectral reconstruction at a small, magic angle on the order of $\approx 1.1^{\circ}$ [35]. At this angle, the (moiré superlattice) Brillouin zone is covered by a pair of very weakly dispersing (so-called) "flat bands" centered on the charge neutrality point [34, 35]. The reduction of kinetic energy due to band flattening strengthens the role of electron-electron interactions and is believed to be responsible for the exciting many-body physics that has been experimentally unveiled (for recent reviews see, for example, Refs. [36, 37]).

P2DESs are also intriguing setups from the point of view of their plasmonic properties, which have been studied theoretically since the Eighties [38, 39]. Indeed, a single 2DES displays a *plasmon* mode [40], which, in the long wavelength $q \to 0$ limit, can be interpreted as a center-of-mass (COM) oscillation dispersing as $\omega_{\rm COM}(q) \propto \sqrt{q}$, as a function of the in-plane wave vector q. This mode is extremely well understood and its small-q behavior is highly constrained by 2D electrodynamics [40], posing practically no bounds on approximate theories for the 2D interacting many-particle problem. On the contrary, two P2DESs harbor an additional collective mode, which behaves very differently from the COM mode, depending on the amplitude of the inter-layer tunneling between the two layers where electrons roam. Let us consider a P2DES realized via a GaAs/AlGaAs double quantum well [2–5]. If the barrier between the two quantum wells is sufficiently strong, the inter-layer tunneling amplitude—which in these systems is well described by a constant quantity typically dubbed $\Delta_{\rm SAS}$, physically representing the splitting between the symmetric and anti-symmetric states in the two adjacent wells—is negligible. In this weak inter-layer tunneling (i.e. $\Delta_{\rm SAS} \to 0$) limit, the additional collective mode is acoustic [38, 39], i.e. $\omega(q) \propto q$ for $q \to 0$. Viceversa, in the limit of strong inter-layer tunneling, the additional collective mode is gapped [42], $\omega(q) \propto \Delta_{\rm SAS}$ for $q \to 0$. The many-body theory of this mode, either for $\Delta_{\rm SAS} = 0$ [39] or $\Delta_{\rm SAS} \neq 0$ [43], is much more subtle than that needed to describe the COM plasmon in a single 2DES. Gapless, acoustic plasmons exist also in graphene double layers and topological insulator thin films [41], provided that the two P2DESs there hosted are well isolated so that inter-layer tunneling can

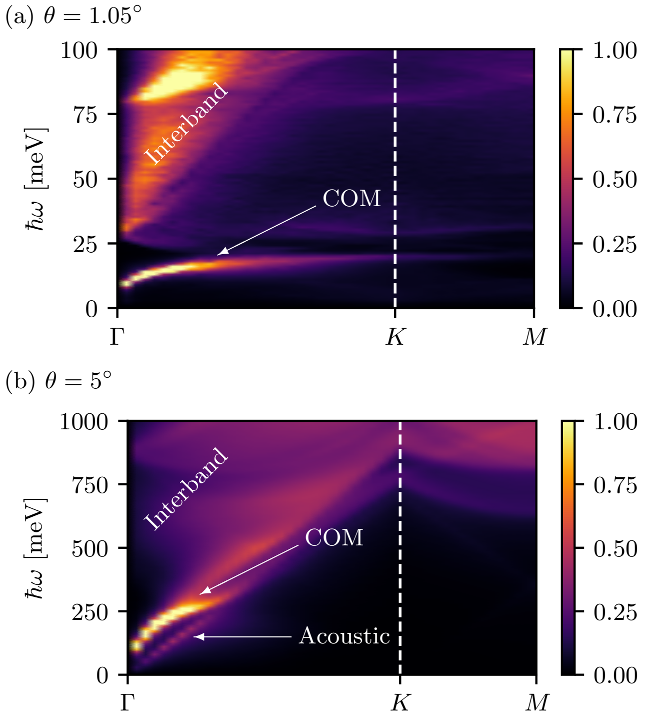


FIG. 1. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ as a function of $\boldsymbol{q}$ and ω. The dependence on $\boldsymbol{q}$ is displayed along the high-symmetry path Γ-K-M of the moiré BZ—see Fig. 2(b). Results in this plot refer to filling factor $\nu = +1$ and temperature $T = 5$ K. Panel (a) Results for $\theta = 1.05^\circ$ (chemical potential $\mu = 22$ meV). (b) Results for $\theta = 5^\circ$ (chemical potential $\mu = 256$ meV). In panel (b), an acoustic plasmon mode is clearly visible at low energies, just above the upper edge of the particle-hole continuum, i.e. $\omega = v^\star_\theta q$, $v^\star_\theta$ being the reduced Fermi velocity—see Eq. (28) below and also Section I of Ref. [63]. High-energy interband plasmons have been discussed at length in Refs. [45, 47, 75].

be neglected.

This Article focuses on a simple question. How is TBG "placed" in this general context? This question is motivated by the qualitative difference between the two interlayer tunneling Hamiltonians in the systems mentioned above, i.e. TBG and GaAs double quantum wells. While in the latter a constant tunneling $\Delta_{\rm SAS}$ works very well, in the former inter-layer tunneling is highly modulated in space on the moiré superlattice length scale. Moreover, TBG too consists of two layers and in principle should support two collective modes at low energies. However, at small twist angles near the magic angle, only one low-energy COM plasmon mode $\omega_{\rm COM}(q) \propto \sqrt{q}$ is seen in state-of-the-art theoretical calculations of the plasmonic modes of TBG [44–47]. Where is the acoustic plasmon mode?

The technical point is that in order to find an *intrinsic* acoustic plasmon in TBG [48], one needs to deal with the layer-pseudospin degree of freedom. This needs to be included into the theoretical treatment of the plasmonic response of TBG, while at the same time taking into account three other important physical effects, namely spatial non-locality of the density-density response function beyond the Drude limit [39, 41, 49], Hartree self-consistency [47, 50] and crystalline local field effects [51, 52].

Accurate theoretical predictions for the plasmonic modes of TBG are important for a variety of fundamental and applied reasons. On the one hand, plasmons in TBG have been suggested as potential candidates for the microscopic explanation of superconductivity [53]. On the other hand, plasmon polaritons in TBG (and many other twisted 2D materials either with itinerant carriers or long-lived phonon modes) enrich the polariton panorama [54], providing us with a system with ultra-slow acoustic plasmons—see Sect. V. Finally, since acoustic plasmons carry an electromagnetic field that is very well confined between the two layers [55–58], they may have important applications in the field of quantum nanophotonics [59] and cavity QED of strongly correlated electron systems [60–62].

This Article is organized as following. In Sect. II we introduce linear response theory for a P2DES consisting of two layers, formulating it for a system with in-plane Bloch translational invariance. In Sect. III we summarize the theoretical approach we have used in this work, which we dub "crystalline" random phase approximation, introducing local field effects and the experimental observable we focus on, i.e. the energy loss function. Section IV is devoted to a brief summary of the TBG continuum model Hamiltonian we rely on. Finally, in Sect. V we present our main numerical results. Section VI contains a brief summary and our main conclusions. Sections I-V of the Supplemental Material [63] contain a wealth of additional numerical results. In particular, Sect. IV deals with the role of an applied perpendicular electric field while Sect. V discusses the impact of heterostrain.

II. LINEAR RESPONSE THEORY FOR TWO-LAYER P2DESS

In this Section we summarize linear response theory (LRT) [40] for a P2DES consisting of two layers. The formalism outlined here will be employed below in Sect. III to evaluate the plasmonic spectrum of TBG.

The ordinary density-density response function for a single 2DES [40] can be easily extended to a P2DES consisting of two layers by using a 2×2 matrix formalism:

$$\begin{pmatrix}\delta n^{(1)}(\boldsymbol{q},\omega)\\ \delta n^{(2)}(\boldsymbol{q},\omega)\end{pmatrix} = \int \frac{d^2\boldsymbol{q}'}{(2\pi)^2}\begin{pmatrix}\chi^{(1,1)}_{\hat n_{\boldsymbol{q}}\hat n_{-\boldsymbol{q}'}}(\omega) & \chi^{(1,2)}_{\hat n_{\boldsymbol{q}}\hat n_{-\boldsymbol{q}'}}(\omega)\\ \chi^{(2,1)}_{\hat n_{\boldsymbol{q}}\hat n_{-\boldsymbol{q}'}}(\omega) & \chi^{(2,2)}_{\hat n_{\boldsymbol{q}}\hat n_{-\boldsymbol{q}'}}(\omega)\end{pmatrix} \times \begin{pmatrix}V^{(1)}_{\rm ext}(\boldsymbol{q}',\omega)\\ V^{(2)}_{\rm ext}(\boldsymbol{q}',\omega)\end{pmatrix} . \tag{1}$$

Here, $\delta n^{(1)}(\boldsymbol{q},\omega)$ and $\delta n^{(2)}(\boldsymbol{q},\omega)$ are the Fourier components of the densities in the two layers, which are linked to the Fourier components of the two external scalar

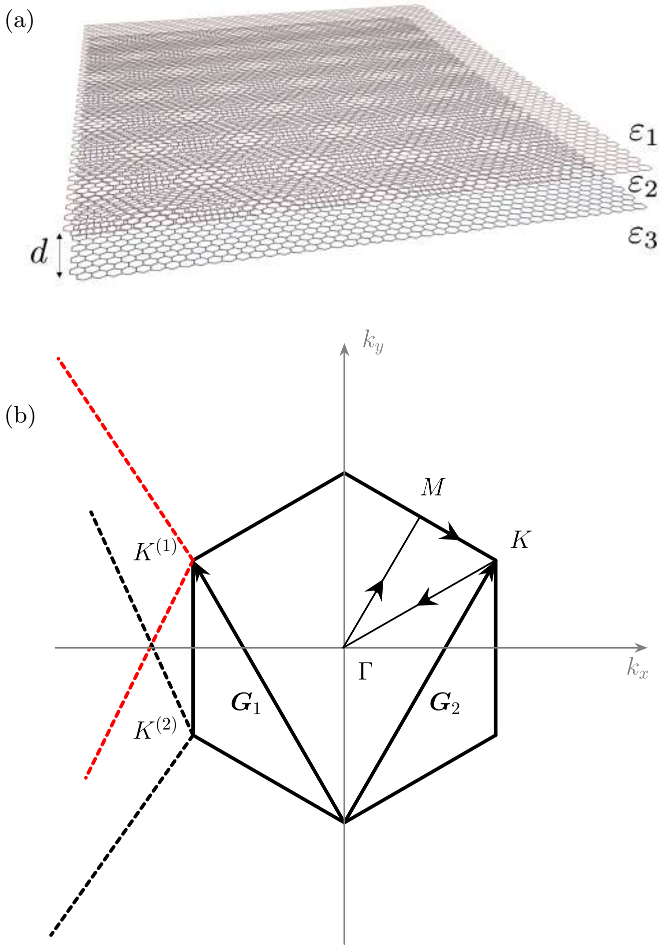


FIG. 2. (Color online) (a) Sketch of the setup studied in this work. TBG (spatial separation between the two graphene layers denoted by d) is embedded in a dieletric environment described by three isotropic and homogeneous dielectrics with dielectric constants, ε_1 (top), ε_2 (middle), and ε_3 (bottom). (b) The first moiré BZ of TBG. The red (black) dashed lines are the edges of the BZ of the graphene layer "1" ("2"), $K^{(1)}$ ($K^{(2)}$) being the corresponding K point. The path K-Γ-M-K is highlighted.

potentials $V^{(1)}_{\rm ext}(\boldsymbol{q}',\omega)$ and $V^{(2)}_{\rm ext}(\boldsymbol{q}',\omega)$ by a 2×2 linear-response matrix. Its matrix elements are the quantities $\chi^{(i,j)}_{\hat{n}_{\boldsymbol q}\hat{n}_{-\boldsymbol{q}'}}(\omega)$, where $i,j=1,2$ are layer indices. For the sake of simplicity, we start by neglecting intra- and inter-layer electron-electron interactions. In this case, the off-diagonal elements $\chi^{(1,2)}_{\hat{n}_{\boldsymbol q}\hat{n}_{-\boldsymbol{q}'}}(\omega)$ and $\chi^{(2,1)}_{\hat{n}_{\boldsymbol q}\hat{n}_{-\boldsymbol{q}'}}(\omega)$ are non-zero only because of inter-layer tunneling, which couples layer 1 with layer 2 and viceversa. Electron-electron interactions will be included below in Sect. III.

Good care needs to be exercised to correctly identify the layer-resolved density operators $\hat{n}^{(i)}_{\boldsymbol q}$ that lead to Eq. (1). The standard number density operator is defined by [40] $\hat{n}(\boldsymbol r)=\sum_{k=1}^{N}\delta(\boldsymbol r-\hat{\boldsymbol r}_k)$, where the sum runs over the $k=1\ldots N$ electrons. In a multi-layer structure, this operator is generalized to $\hat{n}^{(i)}(\boldsymbol r)=\hat{\Pi}^{(i)\dagger}\hat{n}(\boldsymbol r)\hat{\Pi}^{(i)}$. In the previous equation, $i=1,2$ denotes the layer index and $\hat{\Pi}^{(i)}$ is the projector operator onto the i-th layer. In the case of two layers the total density operator is $\hat{n}(\boldsymbol r)=\hat{\Pi}^{(1)\dagger}\hat{n}(\boldsymbol r)\hat{\Pi}^{(1)}+\hat{\Pi}^{(2)\dagger}\hat{n}(\boldsymbol r)\hat{\Pi}^{(2)}$. An explicit construction of the projector operators is given below in Sec. IV.

We now proceed to derive an expression for the quantity $\chi^{(i,j)}_{\hat{n}_{\boldsymbol q}\hat{n}_{-\boldsymbol{q}'}}(\omega)$, which applies to the case in which the P2DES is a crystal, i.e. a Bloch translationally-invariant system. In this case, the single-particle eigenstates are of the Bloch type, i.e. they are labeled by a crystal momentum $\boldsymbol k$ belonging to the first Brillouin Zone (BZ) and a band index λ. A Bloch state $|\boldsymbol k,\lambda\rangle$, with eigenvalue $\epsilon_{\boldsymbol k,\lambda}$, is explicitly given by:

$$\langle \boldsymbol r|\boldsymbol k,\lambda\rangle=\frac{1}{\sqrt{S}}\sum_{\boldsymbol G}u_{\boldsymbol G}(\boldsymbol k,\lambda)e^{i(\boldsymbol k+\boldsymbol G)\cdot\boldsymbol r}\ , \tag{2}$$

where S is the P2DES's area and $\boldsymbol G$ denotes the reciprocal lattice vectors of the crystal. Then, the elements $\chi^{(i,j)}_{\hat{n}_{\boldsymbol q}\hat{n}_{-\boldsymbol{q}'}}(\omega)$ of the non-interacting density-density response matrix can be expanded in a Bloch basis and the wave vectors $\boldsymbol q$ and $\boldsymbol q'$ appearing in Eq. (1) can differ at most by a reciprocal lattice vector (due to the periodicity of the lattice [40, 47]):

$$\begin{aligned}
&\chi^{(i,j)}_{\hat{n}_{\boldsymbol q+\boldsymbol G}\hat{n}_{-\boldsymbol q-\boldsymbol G'}}(\omega)=\\
&=g_{\rm s}\int_{\rm BZ}\frac{d^2\boldsymbol k}{(2\pi)^2}\sum_{\lambda,\lambda'}\frac{f_{\boldsymbol k,\lambda}-f_{\boldsymbol k+\boldsymbol q-\boldsymbol Q,\lambda'}}{\epsilon_{\boldsymbol k,\lambda}-\epsilon_{\boldsymbol k+\boldsymbol q-\boldsymbol Q,\lambda'}+\hbar\omega+i\eta}\\
&\times\langle\boldsymbol k,\lambda|\hat{n}^{(i)}_{\boldsymbol q+\boldsymbol G}|\boldsymbol k+\boldsymbol q-\boldsymbol Q,\lambda'\rangle\\
&\times\langle\boldsymbol k+\boldsymbol q-\boldsymbol Q,\lambda'|\hat{n}^{(j)}_{-\boldsymbol q-\boldsymbol G'}|\boldsymbol k,\lambda\rangle\ .
\end{aligned}\tag{3}$$

Here, $g_{\rm s}=2$ is a spin degeneracy factor, $f_{\boldsymbol k,\lambda}$ is the usual Fermi-Dirac distribution at chemical potential μ and temperature T,

$$f_{\boldsymbol k,\lambda}=\frac{1}{\exp[(\epsilon_{\boldsymbol k,\lambda}-\mu)/(k_{\rm B}T)]+1}\ , \tag{4}$$

and $\eta\to 0^+$ is a positive infinitesimal. A folding vector $\boldsymbol Q$ belonging to the reciprocal lattice has been introduced in Eq. (3) to ensure that $\boldsymbol k+\boldsymbol q$ remains in the first BZ.

III. "CRYSTALLINE" RANDOM PHASE APPROXIMATION

Plasmons are self-sustained density oscillations that emerge due to electron-electron interactions [40]. These need to be treated at some level of approximation. Here, we employ the time-dependent Hartree approximation [40], also known as random phase approximation (RPA), and focus our attention on the electron energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$. This quantity represents the probability of exciting the electronic system through the application of a scalar perturbation with wave vector $\boldsymbol{q}$ and energy $\hbar\omega$. $\mathcal{L}(\boldsymbol{q},\omega)$ contains valuable information about self-sustained charge oscillations, which appear as sharp peaks, as well as incoherent electron-hole pairs, which induce a broadening of the peaks or, more in general, produce a broadly distributed spectral weight in the $\boldsymbol{q}$-ω plane. The energy loss function can be in principle measured via electron energy loss spectroscopy [64] and scattering-type near-field optical spectroscopy (see, for example, Refs. [54–57] and references therein).

As stated in Sect. I, the loss function will be calculated by including *local field effects* (LFEs) [51, 52, 65–67], naturally arising out of the underlying crystalline nature of the system under study. This is very naturally accomplished by retaining the dependence of the quantity $\chi^{(i,j)}_{\hat{n}_{\boldsymbol{q}+\boldsymbol{G}}\hat{n}_{-\boldsymbol{q}-\boldsymbol{G}'}}(\omega)$ in Eq. (3) on the reciprocal lattice vectors $\boldsymbol{G}$, $\boldsymbol{G}'$.

Finally, many-body effects, in general, and plasmons, in particular, are sensitive to the dielectric environment surrounding the P2DES under investigation. In this Article, we assume that TBG is embedded between two homogeneous and isotropic dielectric media described by the dielectric constants ε_1 (top) and ε_3 (bottom)—see Fig. 2(a). The space between the layers is filled by a third homogeneous and isotropic dielectric characterized by a dielectric constant ε_2. In a typical experimental setup, the space between the layers is just a vacuum gap ($\epsilon_2 = 1$) and TBG is encapsulated between two slabs of hexagonal Boron Nitride (hBN), which is a homogeneous and *anisotropic* dielectric (therefore beyond the isotropic model introduced above). Such hBN slabs host hyperbolic phonon polariton modes [54], which strongly couple to plasmons [68]. We have therefore deliberately decided to neglect such plasmon-phonon polariton coupling in order to access, once again, the *intrinsic* plasmon modes of TBG. Including hBN polaritons into the theory is straightforward and can be accomplished by following for example the theory of Ref. [68].

The loss function can be calculated from the following expression:

$$\mathcal{L}(\boldsymbol{q},\omega) = -\operatorname{Im}\left\{\operatorname{Tr}_{\rm L}\left[\varepsilon(\boldsymbol{q},\omega)^{-1}\right]_{\boldsymbol{G}=\boldsymbol{0},\boldsymbol{G}'=\boldsymbol{0}}\right\} , \qquad (5)$$

where $\varepsilon(\boldsymbol{q},\omega)$ is the dynamical dielectric function, which, in the present case, is a matrix with respect to layer indices and reciprocal lattice vectors. The trace $\operatorname{Tr}_{\rm L}$ in Eq. (5) is intended to be over the layer-pseudospin degrees of freedom. We emphasize that, in order to evaluate the loss function via Eq. (5), the matrix $\varepsilon(\boldsymbol{q},\omega)$ needs to be inverted *before* a) the trace over the layer degrees of freedom is taken and b) the $\boldsymbol{G}=\boldsymbol{0}, \boldsymbol{G}'=\boldsymbol{0}$ element is selected.

Returning on the importance of LFEs, we remind the reader that the $\boldsymbol{G}=\boldsymbol{0}$, $\boldsymbol{G}'=\boldsymbol{0}$ element of the inverse of the dynamical dielectric matrix $\varepsilon(\boldsymbol{q},\omega)$ produces the so-called “macroscopic” dielectric function [65, 66] $\varepsilon_{\rm M}(\boldsymbol{q},\omega)$, which is defined through the following equation:

$$\varepsilon^{-1}_{\rm M}(\boldsymbol{q},\omega) \equiv \left[\varepsilon^{-1}(\boldsymbol{q},\omega)\right]_{\boldsymbol{G}=\boldsymbol{0},\boldsymbol{G}'=\boldsymbol{0}} . \qquad (6)$$

Inverting $\varepsilon(\boldsymbol{q},\omega)$ *first*, and *then* selecting the $\boldsymbol{G}=\boldsymbol{0}$, $\boldsymbol{G}'=\boldsymbol{0}$ element, brings to the macroscopic dielectric function contributions from non-zero reciprocal lattice vectors, i.e. $\boldsymbol{G}\neq\boldsymbol{0}$, $\boldsymbol{G}'\neq\boldsymbol{0}$. In solids, such LFEs are not negligible. As a result, the macroscopic field, which is the average of the microscopic field over a region larger than the lattice constant (but smaller than the wavelength) is not equivalent to the effective or local field that polarizes the charge in the crystal [65, 66]. This phenomenon is expected to be more relevant in systems with significant charge inhomogeneities, like moiré materials and TMDs [67, 69, 70]. In particular, modifications to the plasmon dispersion relation induced by LFEs tend to be important near BZ edges [67]. Importantly, the authors of Ref. [67] have recently shown that the inclusion of LFEs on the plasmon dispersion relation is crucial to probe correlated states in twisted hetero-bilayers of TMDs. More precisely, they argue that a loss function different from the one introduced in Eq. (5) and calculated by *tracing over the reciprocal lattice vectors* gives profound information about the many-body properties of the moiré material under investigation. While this is certainly true, standard plasmonic probes [54–57] usually access the response of the system to long-wavelength perturbations. Experimentally, therefore, the loss function defined in Eq. (5) seems the more appropriate one to interpret plasmonic experiments, as briefly pointed out by the authors of Ref. [67] too.

We now comment on the role of the layer degrees of freedom. At a first superficial glance, one may be puzzled by the definition of the loss function we gave above in Eq. (5) and, in particular, by its ability to display peaks at the collective modes of the layered structure. Indeed, in a layered structure, plasmon modes are calculated by looking at the zeroes of the *determinant* of the layer-resolved dielectric tensor [38, 39]. How can we reconcile these two seemingly different approaches to the collective modes of layered materials? The answer is that the trace of the inverse dielectric tensor with respect to the layer degrees of freedom is proportional to the reciprocal of the determinant over the same degrees of freedom, i.e. $\operatorname{Tr}_{\rm L}\left[\varepsilon(\boldsymbol{q},\omega)^{-1}\right]_{\boldsymbol{G}=\boldsymbol{0},\boldsymbol{G}'=\boldsymbol{0}} \propto 1/\det_{\rm L}\left[\varepsilon(\boldsymbol{q},\omega)\right]_{\boldsymbol{G}=\boldsymbol{0},\boldsymbol{G}'=\boldsymbol{0}}$. We therefore see that there is no contradiction between the usual approach [38, 39] and our loss-function based approach.

A. Approximate dynamical dielectric matrix

While the definition in Eq. (5) is totally general, we now need to introduce a necessarily approximate model for the dynamical dielectric matrix $\varepsilon(\boldsymbol{q},\omega)$, which includes electron-electron interactions.

In the RPA [40], we have

$$[\varepsilon(\boldsymbol{q},\omega)]^{(i,j)}_{\boldsymbol{G},\boldsymbol{G}'} = \delta^{(i,j)}\delta_{\boldsymbol{G},\boldsymbol{G}'} - e^2\sum_{\ell} L^{(i,\ell)}_{\boldsymbol{G}}(\boldsymbol{q})\chi^{(\ell,j)}_{\hat{n}_{\boldsymbol{q}+\boldsymbol{G}}\hat{n}_{-\boldsymbol{q}-\boldsymbol{G}'}}(\omega)\ , \tag{7}$$

where $L^{(i,j)}_{\boldsymbol{G}}(\boldsymbol{q}) = L^{(i,j)}(\boldsymbol{q}+\boldsymbol{G})$ is the Coulomb propagator relating the charge density fluctuations $\delta n^{(j)}_{\boldsymbol{q}+\boldsymbol{G}}(\omega)$ to the self-induced electrical potential, i.e. $W^{(i)}_{\boldsymbol{G}}(\boldsymbol{q},\omega) = e^2 L^{(i,j)}_{\boldsymbol{G}}(\boldsymbol{q})\delta n^{(j)}_{\boldsymbol{q}+\boldsymbol{G}}(\omega)$.

The quantities $L^{(i,j)}(q)$ are given by [41]:

$$L^{(1,1)}(q) = \frac{4\pi}{qD(q)}[(\varepsilon_2+\varepsilon_3)e^{qd} + (\varepsilon_2-\varepsilon_3)e^{-qd}]\ , \tag{8}$$

and

$$L^{(1,2)}(q) = L^{(2,1)}(q) = \frac{8\pi}{qD(q)}\varepsilon_2\ , \tag{9}$$

where

$$D(q) = (\varepsilon_1+\varepsilon_2)(\varepsilon_2+\varepsilon_3)e^{qd} + (\varepsilon_1-\varepsilon_2)(\varepsilon_2-\varepsilon_3)e^{-qd}\ . \tag{10}$$

The expression for the $L^{(2,2)}(q)$ component is obtained from Eq. (8) by interchanging ε_3 with ε_1. In the presence of hBN dielectrics, the Coulomb propagator acquires a frequency dependence [68], $L^{(i,j)}_{\boldsymbol{G}}(\boldsymbol{q},\omega)$, due to the strong dependence of the hBN dielectric permittivity tensor on frequency in the mid-infrared spectral range.

It is now time to pause for a moment and discuss about the statements we have made about the non-local nature of the calculations reported in this Article. In the so called "local approximation" for calculating the plasmon dispersion relation in a single 2DES, the density-density response function in equation (7) is approximated with its value in the so-called "dynamical limit" [40], i.e. in the limit $q \to 0$ and $\omega \gg v^*_{\rm F}q$, where $v^*_{\rm F}q$ represents the upper edge of the electron-hole continuum. This approximation is extremely well suited to calculate the leading order term of the dispersion relation $\omega_{\rm COM}(q)$ of the COM mode in the long-wavelength $q \to 0$ limit. However, it is very well known [39, 41] that such local approximation fails in predicting the correct acoustic plasmon dispersion, even in the long wavelength $q \to 0$ limit. This is why, in this Article, we have decided to retain the full dependence of $\chi^{(i,j)}_{\hat{n}_{\boldsymbol{q}+\boldsymbol{G}}\hat{n}_{-\boldsymbol{q}-\boldsymbol{G}'}}(\omega)$ in Eq. (3) on the wave vector $\boldsymbol{q}$, without making the local approximation (i.e. without taking the dynamical limit).

IV. TBG MODEL HAMILTONIAN AND HARTREE SELF-CONSISTENT THEORY

Before illustrating our numerical results, we would like to briefly summarize the single-particle band model we have used to describe TBG and the self-consistent Hartree procedure we have carried out to deal with the important ground-state charge density inhomogeneities displayed by TBG.

A. TBG bare-band model

The continuum model of TBG adopted in this work is the same as the one used in Ref. [47], which was first derived in Refs. [35] and [71].

Layer, sublattice, spin, and valley are the four discrete degrees of freedom characterizing single-electron states in TBG. We can take into account valley and spin degrees of freedom by a degeneracy factor $g = 4 = g_{\rm v}g_{\rm s}$, where the spin-degeneracy factor $g_{\rm s} = 2$ has been introduced earlier. The single-particle Hamiltonian of TBG is written in the layer/sublattice basis $\{|1A\rangle, |1B\rangle, |2A\rangle, |2B\rangle\}$ as:

$$\hat{\mathcal{H}}_0 = \begin{pmatrix} \hat{\mathcal{H}}^{(1)} & \hat{U} \\ \hat{U}^\dagger & \hat{\mathcal{H}}^{(2)} \end{pmatrix}\ . \tag{11}$$

The state $|\ell\tau\rangle$ refers to layer $\ell = 1, 2$ and sublattice index $\tau = A, B$, $\hat{\mathcal{H}}^{(\ell)}$ is the intra-layer Hamiltonian for layer ℓ, and the operator $\hat{U}$ describes inter-layer tunneling. For small twist angles, the moiré length scale $\sim a/\theta$ is much larger than the lattice parameter a of single-layer graphene. This allows us to replace $\hat{\mathcal{H}}^{(\ell)}$ by its $\boldsymbol{k}\cdot\boldsymbol{p}$ massless Dirac fermion limit. This low-energy expansion is done around one of the single layer valleys, $K^{(\ell)}/K'^{(\ell)}$:

$$\hat{\mathcal{H}}^{(\ell)} = v_{\rm D}\left[\mathcal{R}_\ell(\theta/2)(\hat{\boldsymbol{p}} \mp \hbar\boldsymbol{K}_\ell)\right]\cdot(\pm\sigma_x, -\sigma_y)\ . \tag{12}$$

Here, $(\pm\sigma_x, -\sigma_y)$ is a vector of 2×2 Pauli matrices (the ± sign referring to the K and K' valleys, respectively), $\hat{\boldsymbol{p}}$ is the momentum operator, $v_{\rm D} = \sqrt{3}|t|a/(2\hbar) \sim 1\times10^6$m/s is the Fermi velocity of single-layer graphene, $|t| = 2.78$ eV being the usual single-particle nearest-neighbor hopping. The vector $\boldsymbol{K}_\ell$ appearing in Eq. (12) is the position of single layer graphene's valley $K^{(\ell)}$ measured from the moiré BZ center Γ (Fig. 2 (b)):

$$\boldsymbol{K}_{1,2} = \frac{8\pi}{3a}\sin\left(\frac{\theta}{2}\right)\left(-\frac{\sqrt{3}}{2}, \pm\frac{1}{2}\right)\ . \tag{13}$$

The rotation matrix $\mathcal{R}_\ell(\theta/2)$ appearing in (12) is given by:

$$\begin{aligned}\mathcal{R}_{\ell=1,2}(\theta/2) &= \cos(\mp\theta/2)\mathbb{I}_{2\times2} - i\sin(\mp\theta/2)\sigma_y \\ &= \begin{pmatrix} \cos\theta/2 & \pm\sin\theta/2 \\ \mp\sin\theta/2 & \cos\theta/2 \end{pmatrix}\ .\end{aligned} \tag{14}$$

The convention adopted is such that $\theta_{\ell=1} = -\theta/2$ and $\theta_{\ell=2} = \theta/2$. The longitudinal displacement between the two layers is taken as zero in order to obtain the AB-Bernal stacking configuration for $\theta = 0$.

The $\hat{U}$ operator describes inter-layer hopping and is given by:

$$\begin{aligned}\hat{U} &= \begin{pmatrix} u_0 & u_1 \\ u_1 & u_0 \end{pmatrix} + e^{-i\frac{2\pi}{3} + i\boldsymbol{G}_1\cdot\hat{\boldsymbol{r}}}\begin{pmatrix} u_0 & u_1 e^{i\frac{2\pi}{3}} \\ u_1 e^{-i\frac{2\pi}{3}} & u_0 \end{pmatrix} + \\ &+ e^{i\frac{2\pi}{3} + i\boldsymbol{G}_2\cdot\hat{\boldsymbol{r}}}\begin{pmatrix} u_0 & u_1 e^{-i\frac{2\pi}{3}} \\ u_1 e^{i\frac{2\pi}{3}} & u_0 \end{pmatrix}\ ,\end{aligned} \tag{15}$$

where

$$\boldsymbol{G}_{1,2}=\frac{8\pi}{\sqrt{3}a}\sin\left(\frac{\theta}{2}\right)\left(\pm\frac{1}{2},\frac{\sqrt{3}}{2}\right)\ , \tag{16}$$

and u_0 (u_1) are the intra-sublattice (inter-sublattice) hopping parameters. In general $u_0 \neq u_1$. The difference between these two parameters can, in fact, take into account the lattice corrugation of TBG samples [71–74]. The intra- and inter-sublattice hopping energies might also be affected in value by possible stresses induced on the TBG sheet during the production phase. Recently [75] it has been shown experimentally that the difference between the intra- and inter-sublattice hopping parameters is in the range of $u_1 - u_0 \sim$ 30-60 meV. In this work, we take $u_1 = 97.5$ meV and $u_0 = 79.7$ meV. With this choice, we have $u_1 - u_0 \approx 20$ meV and the dimensionless parameter $u_0/u_1 \sim 0.8$ takes correctly into account relaxation effects [71]. Within the continuum model described by the single-particle Hamiltonian in Eq. (11), we can construct the projector operators onto the i-th layer $\hat{\Pi}^{(i)}$ by making explicit their action on the basis $|\ell\tau\rangle$:

$$\hat{\Pi}^{(i)}|\ell\tau\rangle = |i\tau\rangle\ . \tag{17}$$

In particular their matrix form is given explicitly by:

$$\hat{\Pi}^{(1)}=\begin{pmatrix}\hat{\mathbb{I}}_{2\times 2} & 0\\ 0 & 0\end{pmatrix}\ , \tag{18}$$

$$\hat{\Pi}^{(2)}=\begin{pmatrix}0 & 0\\ 0 & \hat{\mathbb{I}}_{2\times 2}\end{pmatrix}\ , \tag{19}$$

where $\hat{\mathbb{I}}_{2\times 2}$ is the identity operator acting on the sublattice index.

The chemical potential μ in Eq. (4) can be calculated by enforcing, as usual, particle-number conservation:

$$n=\delta n+n_0=g\sum_{\lambda}\int\frac{d^2\boldsymbol{k}}{(2\pi)^2}f^{\rm reg}_{\boldsymbol{k},\lambda}(\mu)\ . \tag{20}$$

Here, n_0 is the total electron density at the charge neutrality point (CNP) and δn is the electron density measured from the CNP. We stress that a *regularized* Fermi-Dirac distribution function $f^{\rm reg}_{\boldsymbol{k},\lambda}$ appears in Eq. (20). Indeed, since we are dealing with a continuum model, the number of bands is formally infinite below and above the CNP. In order to regularize the Dirac sea below the CNP, one needs to introduce the regularized Fermi-Dirac distribution function defined as following:

$$\begin{aligned}f^{\rm reg}_{\boldsymbol{k},\lambda}(\mu) &\equiv f^{\rm reg}(\epsilon_{\boldsymbol{k},\lambda}-\mu)=\\ &=f(\epsilon_{\boldsymbol{k},\lambda}-\mu)-\Theta(\epsilon_{\rm CNP}-\epsilon_{\boldsymbol{k},\lambda})\ ,\end{aligned} \tag{21}$$

where $\Theta(x)$ is the Heaviside step-function and $\epsilon_{\rm CNP}$ is the energy of the CNP.

With these conventions, the filling factor ν is defined by:

$$\nu\equiv\Omega_{\rm u.c.}\delta n\ , \tag{22}$$

where $\Omega_{\rm u.c.}=\frac{\sqrt{3}}{2}\left[\frac{a}{2\sin{(\theta/2)}}\right]^2$ is the area of the moiré unit cell. With this definition of the filling factor, one has $|\nu| < 4$ when the chemical potential is within the flat bands, at low temperatures.

B. Hartree self-consistency

Inhomogeneities in the ground-state charge density distribution of TBG create an inhomogeneous electrical potential that depends on the filling factor. To capture this effect, we need to add the so-called Hartree contribution $\hat{V}_{\rm H}$ to the bare TBG Hamiltonian $\hat{\mathcal{H}}_0$ [40, 47, 50]:

$$\hat{\mathcal{H}}=\hat{\mathcal{H}}_0+\hat{V}_{\rm H}[n_{\boldsymbol{G}}]\ , \tag{23}$$

where

$$\hat{V}_{\rm H}[n_{\boldsymbol{G}}]=\mathbb{I}_{4\times 4}\sum_{\boldsymbol{G}\neq\boldsymbol{0}}\frac{2\pi e^2}{\bar{\varepsilon}|\boldsymbol{G}|}n_{\boldsymbol{G}}e^{i\boldsymbol{G}\cdot\hat{\boldsymbol{r}}}\ . \tag{24}$$

Here, $\bar{\varepsilon}\equiv(\varepsilon_1+\varepsilon_3)/2$, $n_{\boldsymbol{G}}$ is the Fourier component of the ground-state electron density corresponding to the reciprocal lattice vector $\boldsymbol{G}$, and the identity matrix $\mathbb{I}_{4\times 4}$ is expressed in the same basis of states of the Hamiltonian, namely $\{|1A\rangle, |1B\rangle, |2A\rangle, |2B\rangle\}$.

The problem posed by Eqs. (23)-(24) needs to be solved self-consistently, i.e., one needs to solve the Hartree equation

$$\left(\hat{\mathcal{H}}_0+\hat{V}_{\rm H}[n_{\boldsymbol{G}}]\right)|\boldsymbol{k},\lambda\rangle=\epsilon_{\boldsymbol{k},\lambda}|\boldsymbol{k},\lambda\rangle\ , \tag{25}$$

together with the closure:

$$n_{\boldsymbol{G}}=g\sum_{\lambda}\int\frac{d^2\boldsymbol{k}}{(2\pi)^2}f^{\rm reg}_{\boldsymbol{k},\lambda}\langle\boldsymbol{k},\lambda|e^{-i\boldsymbol{G}\cdot\hat{\boldsymbol{r}}}|\boldsymbol{k},\lambda\rangle\ . \tag{26}$$

Note that, due to the real-space representation (2) of the Bloch eigenstates, we have:

$$\begin{aligned}&\langle\boldsymbol{k},\lambda|e^{-i\boldsymbol{G}\cdot\hat{\boldsymbol{r}}}|\boldsymbol{k},\lambda\rangle=\\ &=\frac{1}{S}\sum_{\boldsymbol{K},\boldsymbol{K}'}u^\dagger_{\boldsymbol{K}}(\boldsymbol{k},\lambda)u_{\boldsymbol{K}'}(\boldsymbol{k},\lambda)\int d^2\boldsymbol{r}e^{-i(\boldsymbol{G}+\boldsymbol{K}+\boldsymbol{k}-\boldsymbol{K}'-\boldsymbol{k})\cdot\boldsymbol{r}}\\ &=\sum_{\boldsymbol{K},\boldsymbol{K}'}u^\dagger_{\boldsymbol{K}}(\boldsymbol{k},\lambda)u_{\boldsymbol{K}'}(\boldsymbol{k},\lambda)\delta_{\boldsymbol{G}+\boldsymbol{K},\boldsymbol{K}'}\\ &=\sum_{\boldsymbol{K}}u^\dagger_{\boldsymbol{K}}(\boldsymbol{k},\lambda)u_{\boldsymbol{K}+\boldsymbol{G}}(\boldsymbol{k},\lambda)\ .\end{aligned} \tag{27}$$

Once the self-consistent problem has been solved, the Hartree eigenstates $|\boldsymbol{k},\lambda\rangle$ and eigenvalues $\epsilon_{\boldsymbol{k},\lambda}$ can be used in order to calculate the so-called Hartree density-density response [40] matrix. This is simply obtained by using Eq. (3), with the understanding that the two quantities $|\boldsymbol{k},\lambda\rangle$ and $\epsilon_{\boldsymbol{k},\lambda}$ in there need to be interpreted as self-consistently calculated Hartree quantities rather than single-particle, bare quantities.

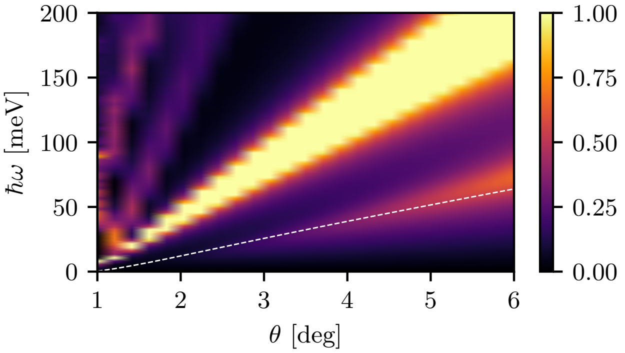


FIG. 3. (Color online) The energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ of TBG is plotted as a function of the twist angle θ and frequency ω. Results in this plot have been obtained by keeping fixed the wave number q and filling factor ν, i.e. $q = q_\theta \equiv 2|\boldsymbol{K}_{1,2}|/31$ (see main text) and $\nu = +1$. Bright bands correspond to plasmons peaks. The white dashed line indicates the upper edge of the particle-hole continuum, i.e. $\omega = v^\star(\theta)q_\theta$, $v^\star(\theta)$ being the reduced Fermi velocity, above which collective modes are well defined. The low-energy acoustic plasmon mode, which "tracks" the upper edge of the particle-hole continuum, disappears for $\theta \lesssim 2^\circ$.

V. NUMERICAL RESULTS

In this Section we present our main numerical results obtained with the theory outlined above. For the sake of definiteness, we set $\varepsilon_1 = \varepsilon_3 = 4.9$, $\varepsilon_2 = 1$, and $T = 5$ K.

The dielectric tensor and hence the loss function are obtained by using the calculated Hartree self-consistent bands and corresponding Bloch states. These calculations take into account the role of static screening in reshaping the electronic bands and redistributing in space the carrier density. The Hartree self-consistency effect on plasmons is more important at small twist angles, since in this regime the system displays larger charge inhomogeneities [47]. This is true also for the LFEs.

Fig. 1 shows the TBG loss function for filling factor $\nu = +1$ and two values of the twist angle θ, i.e. $\theta = 1.05^\circ$ in panel (a) and $\theta = 5^\circ$ in panel (b). This filling factor corresponds to a carrier density $n = 0.64\times10^{12}$ cm^{-2} for $\theta = 1.05^\circ$ and $n = 1.5\times10^{13}$ cm^{-2} for $\theta = 5^\circ$. Chemical potential values have been given in the caption of Fig. 1. Close to the magic angle, Fig. 1(a), flat bands centered at the CNP and separated by an energy gap from the higher-energy bands, lead to intrinsically undamped slow plasmons [46]. We clearly see this in Fig. 1(a), where a narrow, almost dispersion-less plasmon is present at energies on the order of ~ 20 meV. In general, we find that, at small twist angles, TBG hosts a standard intraband COM plasmon with a $\omega_{\rm COM}(q) \propto \sqrt{q}$ dispersion in the long-wavelength limit. No sign of other collective modes is seen at small values of θ, neither gapless [39, 41] nor gapped [42, 43]—further results are reported in Section II of Ref. [63].

This is not the case for larger values of the twist angle, as seen for example in Fig. 1(b) for $\theta = 5^\circ$. For this value of the twist angle, an acoustic plasmon is clearly visible. This mode lies just above the upper edge of the particle-hole continuum (Section I of Ref. [63]), which is identified by the line $\hbar\omega_\theta(q) = \hbar v^\star_\theta q$, $v^\star_\theta$ being the reduced Fermi velocity of the TBG Dirac cones [32]:

$$v^\star_\theta = v_{\rm D}\frac{1-3\alpha^2(\theta)}{1+6\alpha^2(\theta)}\ , \tag{28}$$

$\alpha(\theta) = u_1\left[\frac{8\pi}{\sqrt{3}a}\hbar v_{\rm D}\sin\left(\frac{\theta}{2}\right)\right]^{-1}$ being a dimensionless parameter that depends on the twist angle (the parameters $v_{\rm D}$ and u_1 have been introduced in Sect. IV A). For $\theta = 5^\circ$, the Fermi velocity (28) is $v^*_\theta \approx 7.99\times10^5$ m/s, while the acoustic plasmon velocity in Fig. 1(b) is $c_{\rm s} \approx 8.43\times10^5$ m/s. For the sake of comparison, we note that the acoustic plasmon velocity in two (tunnel-decoupled but Coulomb-coupled) graphene layers at a distance $d = 0.3$ nm is $c_{\rm s} \approx 1.2\times10^6$ m/s (and at the same density $n = 1.5\times10^{13}$ cm^{-2}) [41]. A reduced single-particle Fermi velocity in TBG leads to slower acoustic plasmons with respect to other graphene-related systems [41]. A plot illustrating the dependence of $c_{\rm s}$ on θ is reported in Section I of Ref. [63].

(Further numerical results are reported in Section II of Ref. [63]—where the plasmon dispersion relation obtained with the inclusion of the layer-pseudospin degree of freedom and LFEs is compared with that obtained by neglecting the latter—and Section III of Ref. [63]— where the dependence on the filling factor ν is discussed, for various twist angles. In Section II of Ref. [63], we note that the introduction of LFEs leads to a blue shift in the energy of the plasmon modes around the edge of the moiré BZ, as already found out in other systems [67, 69, 70]. This effect is even more pronounced at small twist angles. In Section III of Ref. [63], we observe, for a fixed value of θ, a weak dependence on ν. The impact of an applied perpendicular electric field and heterostrain on the plasmonic spectrum of TBG are discussed in Sects. IV and V of Ref. [63], respectively.)

Fig. 3 shows the loss function $\mathcal{L}(\boldsymbol{q},\omega)$ as a function of the twist angle θ and frequency ω. Results in this figure have been obtained by setting $q = q_\theta \equiv \xi|\boldsymbol{K}_{1,2}|$, where $|\boldsymbol{K}_1| = |\boldsymbol{K}_2|$ is the modulus of the θ-depending vector linking Γ to K in the moiré BZ—see Eq. (13)—and $\xi = 2/31 < 1$. The brightest feature in this figure corresponds to the usual COM plasmon while the lower-energy feature corresponds to the acoustic plasmon. At twist angles $\theta \lesssim 2^\circ$, the acoustic plasmon branch disappears. We conclude that, at small twist angles, low energies, and long wavelengths, TBG behaves effectively as a single 2DES with an ordinary COM plasmon. A weakly-damped out-of-phase acoustic plasmon appears only for twist angles larger than $\theta \approx 2^\circ$. As discussed in Sect I, this mode is typical of weakly-coupled double layers, where two spatially-separated 2DESs interact only through the long-range Coulomb interaction [39, 41]. The

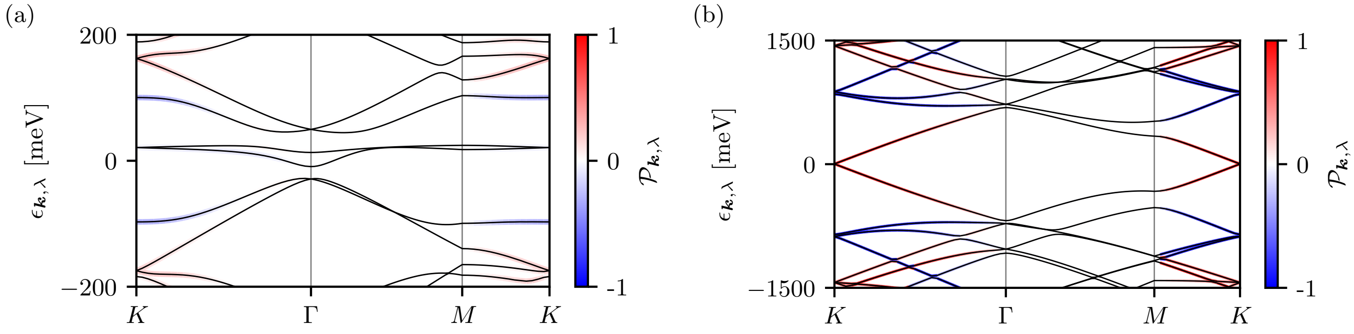


FIG. 4. (Color online) Layer polarization $\mathcal{P}_{\boldsymbol{k},\lambda}$ of the Hartree self-consistent eigenstates , superimposed on TBG energy bands calculated with Hartree self-consistency at filling factor $\nu = +1$. Panel (a) $\theta = 1.05^{\circ}$. Panel (b) $\theta = 5^{\circ}$. At lower angles the Hamiltonian eigenstates are less layer polarized, resulting in more hybridization and the suppression of the bi-layer acoustic plasmon mode. Bands are calculated at the K' valley.

gapless nature of the extra mode emerging for $\theta \gtrsim 2^{\circ}$ is reasonable since the moiré potential that couples the two layers does not open a gap at the K/K' points (Dirac cones are protected by symmetry).

Despite the apparent similarity with spatially-separated 2DESs, acoustic plasmons in TBG offer a qualitative difference: in the latter system, they emerge only for sufficiently large values of θ. In the former systems, instead, acoustic plasmons exist for all values of the macroscopic parameters, provided that the single-particle Fermi velocities in the two 2DESs are identical [39, 41].

Regarding damping of the TBG acoustic plasmon, let us recall that the upper edge of the particle-hole continuum in TBG is given by:

$$\hbar\omega_{\theta}(q_{\theta}) \equiv \hbar v^{*}_{\theta} q_{\theta} = \xi \frac{8\pi}{\sqrt{3}a} \hbar v_{\rm D} \frac{\sin^2(\theta/2) - 3\tilde{\alpha}^2}{\sin^2(\theta/2) + 6\tilde{\alpha}^2} \sin(\theta/2) \ , \tag{29}$$

where $\tilde{\alpha} = \alpha(\theta)/\sin(\theta/2)$ and $\alpha(\theta)$ has been introduced above in Eq. (28). If the plasmon dispersion lies *above* this threshold value, it is a well-defined (i.e. long lived) mode (at least within the RPA). Since the wave vector q is fixed at the value $q_{\theta} \equiv \xi|\boldsymbol{K}_{1,2}|$, the expression on the right hand side of Eq. (29) depends only on θ and is plotted in Fig. 3 (white dashed line) for small values of θ (up to $\theta = 6^{\circ}$). We clearly see that, for sufficiently large values of θ (i.e. $\theta \gtrsim 4^{\circ}$) the acoustic plasmon is a well-defined long-lived collective mode.

In order to better understand the disappearance of the acoustic mode for $\theta \lesssim 2^{\circ}$, we have calculated the layer polarization $\mathcal{P}_{\boldsymbol{k},\lambda}$ of the TBG Hartree self-consistent eigenstates $|\boldsymbol{k},\lambda\rangle$. This quantity is defined as [76]:

$$\mathcal{P}_{\boldsymbol{k},\lambda} \equiv \langle \boldsymbol{k},\lambda|\hat{\Pi}^{(1)}|\boldsymbol{k},\lambda\rangle - \langle \boldsymbol{k},\lambda|\hat{\Pi}^{(2)}|\boldsymbol{k},\lambda\rangle \ , \tag{30}$$

where $\hat{\Pi}^{(i)}$ is the projector operator onto the i-th layer introduced in Sec. IV, Eq. (17). Fig. 4 shows the layer polarization (color bar) at the K valley and for two values of the twist angle, i.e. $\theta = 1.05^{\circ}$—panel (a)—and $\theta = 5^{\circ}$—panel (b). For the latter value of the twist angle, the polarization is $|\mathcal{P}_{\boldsymbol{k},\lambda}| \approx 1$ for almost every value of the wave vector $\boldsymbol{k}$ and throughout all the bands. At $\theta = 1.05^{\circ}$, instead, we observe a very low layer polarization stemming from a strong inter-layer hybridization. It is this transition from high to low values of the layer polarization that, in our opinion, leads to the disappearance of the acoustic plasmon mode at twist angles $\theta \lesssim 2^{\circ}$.

VI. SUMMARY AND CONCLUSIONS

In this Article we have presented a theoretical study of the plasmonic response of twisted bilayer graphene as a function of the twist angle θ. Our theory treats on equal footing four important effects, namely the layer degree of freedom, non-local effects in the density-density response function beyond the dynamical long-wavelength limit, Hartree self-consistency, and crystalline local field effects.

We have found that at small values of the twist angle ($\theta \lesssim 2^{\circ}$) and in the low-energy long-wavelength limit, the 2D electron system in twisted bilayer graphene responds to a perturbation carrying wave vector q and energy $\hbar\omega$ as a single entity, displaying a center-of-mass mode $\omega_{\rm COM}(q) \propto \sqrt{q}$. This is in agreement with all earlier studies [44–47]. As the twist angle increases, however, inter-layer tunneling decreases and the layer-pseudospin becomes a quasi-good quantum number. For $\theta \gtrsim 2^{\circ}$, the layer-pseudospin degree of freedom needs to be taken into account and the plasmonic spectrum of the system displays a qualitatively different behavior. In this case, indeed, a weakly-damped acoustic plasmon mode appears, akin to the acoustic plasmon of other parallel 2D electron systems of historical importance [38, 39].

In the future it will be interesting to feed our results to an Eliashberg theory [77] of plasmon-mediated superconductivity in twisted bilayer graphene and to study the spatial distribution of chirality associated to this mode [78, 79].

ACKNOWLEDGMENTS

This work was supported by i) the European Union's Horizon 2020 research and innovation programme under grant agreement no. 881603 - GrapheneCore3 and the Marie Sklodowska-Curie grant agreement No. 873028 and ii) the Italian Minister of University and Research (MUR) under the "Research projects of relevant national interest - PRIN 2020" - Project no. 2020JLZ52N, title "Light-matter interactions and the collective behavior of quantum 2D materials (q-LIMA)". P.J.H. acknowledges support by the Gordon and Betty Moore Foundation's EPiQS Initiative through grant GBMF9463, the Fundacion Ramon Areces, and the ICFO Distinguished Visiting Professor program. F.H.L.K. acknowledges financial support from the ERC TOPONANOP (726001), the Government of Catalonia trough the SGR grant, the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (Ref. SEV-2015-0522) and Explora Ciencia (Ref. FIS2017- 91599-EXP), Fundacio Cellex Barcelona, Generalitat de Catalunya through the CERCA program, the Mineco grant Plan Nacional (Ref. FIS2016-81044-P), and the Agency for Management of University and Research Grants (AGAUR) (Ref. 2017-SGR-1656).

L.C. gratefully wishes to thank Matteo Ceccanti and Simone Cassandra for help with the graphics work.

[1] Y. E. Lozovik and V. I. Yudson, Feasibility of superfluidity of paired spatially separated electrons and holes; a new superconductivity mechanism, JETP Lett. **22**, 274 (1975).

[2] R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, Electron mobilities in modulation-doped semiconductor heterojunction superlattices, Appl. Phys. Lett. **33**, 665 (1978).

[3] M. J. Manfra, Molecular beam epitaxy of ultra-high quality AlGaAs/GaAs heterostructures: enabling physics in low-dimensional electronic systems, arXiv:1309.2717.

[4] Y. J. Chung, K. A. V. Rosales, K. W. Baldwin, K. W. West, M. Shayegan, and L. N. Pfeiffer, Working principles of doping-well structures for high-mobility two-dimensional electron systems, Phys. Rev. Materials **4**, 044003 (2020).

[5] Y. J. Chung, K. A. Villegas Rosales, K. W. Baldwin, P. T. Madathil, K. W. West, M. Shayegan, and L. N. Pfeiffer, Ultra-high-quality two-dimensional electron systems, Nat. Mater. **20**, 632 (2021).

[6] A. Geim and I. Grigorieva, Van der Waals heterostructures, Nature **499**, 419 (2013).

[7] M. B. Pogrebinskii, Drag effects in a two-layer system of spatially separated electrons and excitons, Sov. Phys. Semicond. **11**, 372 (1977).

[8] P. J. Price, Hot electron effects in heterolayers, Physica B **117**, 750 (1983).

[9] T. J. Gramila, J. P. Eisenstein, A. H. MacDonald, L. N. Pfeiffer, and K. W. West, Mutual friction between parallel two-dimensional electron systems, Phys. Rev. Lett. **66**, 1216 (1991).

[10] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, K. Watanabe, T. Taniguchi, and L. A. Ponomarenko, Strong Coulomb drag and broken symmetry in double-layer graphene, Nat. Phys. **8**, 896 (2012).

[11] A. G. Rojo, Electron-drag effects in coupled electron systems, J. Phys.: Condens. Matter **11**, R31 (1999).

[12] B. N. Narozhny and A. Levchenko, Coulomb drag, Rev. Mod. Phys. **88**, 025003 (2016).

[13] J. P. Eisenstein and A. H. MacDonald, Bose–Einstein condensation of excitons in bilayer electron systems, Nature **432**, 691 (2004).

[14] J. P. Eisenstein, Exciton condensation in bilayer quantum Hall systems, Annu. Rev. Condens. Matter Phys. **5**, 159 (2014).

[15] J.-J. Su and A. H. MacDonald, How to make a bilayer exciton condensate flow, Nat. Phys. **4**, 799 (2008).

[16] X. Liu, K. Watanabe, T. Taniguchi, B. I. Halperin, and P. Kim, Quantum Hall drag of exciton condensate in graphene, Nat. Phys. **13**, 746 (2017).

[17] J. I. A. Li, T. Taniguchi, K. Watanabe, J. Hone, and C. R. Dean, Excitonic superfluid phase in double bilayer graphene, Nat. Phys. **13**, 751 (2017).

[18] G. W. Burg, N. Prasad, K. Kim, T. Taniguchi, K. Watanabe, A. H. MacDonald, L. F. Register, and E. Tutuc, Strongly enhanced tunneling at total charge neutrality in double-bilayer graphene-WSe$_2$ heterostructures, Phys. Rev. Lett. **120**, 177702 (2018).

[19] B. Feldman, J. Martin, and A. Yacoby, Broken-symmetry states and divergent resistance in suspended bilayer graphene, Nat. Phys. **5**, 889 (2009).

[20] R. T. Weitz, M. T. Allen, B. E. Feldman, J. Martin, and A. Yacoby, Broken-symmetry states in doubly gated suspended bilayer graphene, Science **330**, 812 (2010).

[21] A. S. Mayorov, D. C. Elias, M. Mucha-Kruczynski, R. V. Gorbachev, T. Tudorovskiy, A. Zhukov, S. V. Morozov, M. I. Katsnelson, V. I. Fal'ko, A. K. Geim, and K. S. Novoselov, Interaction-Driven Spectrum Reconstruction in Bilayer Graphene, Science **333**, 860 (2011).

[22] F. Freitag, J. Trbovic, M. Weiss, and C. Schönenberger, Spontaneously gapped ground state in suspended bilayer graphene, Phys. Rev. Lett. **108**, 076602 (2012).

[23] W. Bao, J. Velasco Jr, F. Zhang, L. Jing, B. Standley, D. Smirnov, M. Bockrath, A. H. MacDonald, and C. N. Lau, Evidence for a spontaneous gapped state in ultraclean bilayer graphene, Proc. Nat. Acad. Sci. (USA) **109**, 10802 (2012).

[24] J. Velasco Jr., L. Jing, W. Bao, Y. Lee, P. Kratz, V. Aji, M. Bockrath, C. N. Lau, C. Varma, R. Stillwell, D. Smirnov, Fan Zhang, J. Jung, and A. H. MacDonald, Nat. Nanotech. **7**, 156 (2012).

[25] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero , Unconventional superconductivity in magic-angle graphene superlattices, Nature **556**, 43 (2018).

[26] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J D. Sanchez-Yamagishi, K. Watanabe, T.

Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, Nature **556**, 80 (2018).

[27] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Graphene bilayer with a twist: Electronic structure, Phys. Rev. Lett. **99**, 256802 (2007).

[28] S. Shallcross, S. Sharma, and O. A. Pankratov, Quantum interference at the twist boundary in graphene, Phys. Rev. Lett. **101**, 056803 (2008).

[29] E. J. Mele, Commensuration and interlayer coherence in twisted bilayer graphene, Phys. Rev. B **81**, 161405(R) (2010).

[30] G. Li, A. Luican, J. M. B. Lopes dos Santos, A. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei, Observation of Van Hove singularities in twisted graphene layers, Nat. Phys. **6**, 109 (2010).

[31] S. Shallcross, S. Sharma, E. Kandelaki, and O. A. Pankratov, Electronic structure of turbostratic graphene, Phys. Rev. B **81**, 165105 (2010).

[32] R. Bistritzer and A. H. MacDonald, Transport between twisted graphene layers, Phys. Rev. B **81**, 245412 (2010).

[33] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Continuum model of the twisted graphene bilayer, Phys. Rev. B **86**, 155449 (2012).

[34] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic, Flat bands in slightly twisted bilayer graphene: Tight-binding calculations, Phys. Rev. B **82**, 121407(R) (2010).

[35] R. Bistritzer and A. H. MacDonald, Moiré bands in twisted double-layer graphene, Proc. Natl. Acad. Sci. (USA) **108**, 12233 (2011).

[36] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, The marvels of moiré materials, Nat. Rev. Mater. **6**, 201 (2021).

[37] D. M. Kennes, M. Claassen, L. Xian, A. Georges, A. J. Millis, J. Hone, C. R. Dean, D. N. Basov, A. Pasupathy, and A. Rubio, Moiré heterostructures as a condensed-matter quantum simulator, Nat. Phys. **17**, 155 (2021).

[38] S. Das Sarma and A. Madhukar, Collective modes of spatially separated, two-component, two-dimensional plasma in solids, Phys. Rev. B **23**, 805 (1981).

[39] G. E. Santoro and G. F. Giuliani, Acoustic plasmons in a conducting double layer, Phys. Rev. B **37**, 937 (1988).

[40] G. F. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005).

[41] R. E. V. Profumo, R. Asgari, M. Polini, and A. H. MacDonald, Double-layer graphene and topological insulator thin-film plasmons, Phys. Rev. B **85**, 085443 (2012).

[42] S. Das Sarma and E. H. Hwang, Plasmons in coupled bilayer structures, Phys. Rev. Lett. **81**, 4216 (1998).

[43] S. H. Abedinpour, M. Polini, A. H. MacDonald, B. Tanatar, M. P. Tosi, and G. Vignale, Theory of the pseudospin resonance in semiconductor bilayers, Phys. Rev. Lett. **99**, 206802 (2007).

[44] T. Stauber, P. San-Jose, and L. Brey, Optical conductivity, Drude weight and plasmons in twisted graphene bilayers, New J. Phys. **15**, 113050 (2013).

[45] T. Stauber and H. Kohler, Quasi-flat plasmonic bands in twisted bilayer graphene, Nano Lett. **16**, 6844 (2016).

[46] C. Lewandowski and L. Levitov, Intrinsically undamped plasmon modes in narrow electron bands, Proc. Natl. Acad. Sci. (USA) **116**, 42 (2019).

[47] P. Novelli, I. Torre, F. H. L. Koppens, F. Taddei and M. Polini, Optical and plasmonic properties of twisted bilayer graphene: Impact of interlayer tunneling asymmetry and ground-state charge inhomogeneity, Phys. Rev. B **102**, 125403 (2020).

[48] We stress that "intrinsic" is an important keyword here. One can easily get an extrinsic low-energy acoustic mode by morphing the COM mode $\omega_{\rm COM}(q) \propto \sqrt{q}$ into an acoustic mode. Such morphing, which is achieved by placing TBG near a metal gate, has been clearly observed experimentally in the case of single-layer graphene: see, for example, Ref. [55] and P. Alonso-González, A.Y. Nikitin, Y. Gao, A. Woessner, M.B. Lundeberg, A. Principi, N. Forcellini, W. Yan, S. Vélez, A.J. Huber, K. Watanabe, T. Taniguchi, F. Casanova, L.E. Hueso, M. Polini, J. Hone, F. H. L. Koppens, and R. Hillenbrand, Ultra-confined acoustic THz graphene plasmons revealed by photocurrent nanoscopy, Nat. Nanotech. **12**, 31 (2017).

[49] Acoustic plasmons in TBG have been first discussed in pioneering work by T. Stauber, T. Low, and G. Gómez-Santos, Chiral response of twisted bilayer graphene, Phys. Rev. Lett. **120**, 046801 (2018). These authors, however, have carried out calculations by using a local approximation for the conductivity tensor of TBG. This approximation works very well in predicting the COM $\omega_{\rm COM}(q) \propto \sqrt{q}$ dispersion in the long-wavelength $q \to 0$ limit but fails to provide a reliable theory of the acoustic plasmon mode. For a detailed discussion, we refer the reader to Refs. 39–41.

[50] F. Guinea and N. R. Walet, Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers, Proc. Natl. Acad. Sci. (USA) **115**, 13174 (2018).

[51] A. Tomadin, F. Guinea, and M. Polini, Generation and morphing of plasmons in graphene superlattices, Phys. Rev. B **90**, 161406(R) (2014).

[52] A. Tomadin, M. Polini, and J. Jung, Plasmons in realistic graphene/hexagonal boron nitride moiré patterns, Phys. Rev. B **99**, 035432 (2019).

[53] G. Sharma, M. Trushin, O. P. Sushkov, G. Vignale, and S. Adam, Superconductivity from collective excitations in magic-angle twisted bilayer graphene, Phys. Rev. Research **2**, 022040(R) (2020).

[54] D. N. Basov, A. Asenjo-Garcia, P. James Schuck, X. Zhu, and A. Rubio, Polariton panorama, Nanophoton. **10**, 549 (2021).

[55] M. B. Lundeberg, Y. Gao, R. Asgari, C. Tan, B. Van Duppen, M. Autore, P. Alonso-González, A. Woessner, K. Watanabe, T. Taniguchi, R. Hillenbrand, J. Hone, M. Polini, and F. H. L. Koppens, Tuning quantum nonlocal effects in graphene plasmonics, Science **357**, 187 (2017).

[56] A. Woessner, A. Misra, Y. Cao, I. Torre, A. Mishchenko, M. B. Lundeberg, K. Watanabe, T. Taniguchi, M. Polini, K. S. Novoselov, and F. H. L. Koppens, Propagating plasmons in a charge-neutral quantum tunneling transistor, ACS Photon. **4**, 3012 (2017).

[57] D. Alcaraz Iranzo, S. Nanot, E. J. C. Dias, I. Epstein, C. Peng, D. K. Efetov, M. B. Lundeberg, R. Parret, J. Osmond, J.-Y. Hong, J. Kong, D. R. Englund, N. M. R. Peres, and F. H. L. Koppens, Probing the ultimate plasmon confinement limits with a van der Waals heterostructure, Science **360**, 291 (2018).

[58] A. Principi, E. van Loon, M. Polini, and M. I. Katsnel-

son, Confining graphene plasmons to the ultimate limit, Phys. Rev. B **98**, 035427 (2018).

[59] A. Reserbat-Plantey, I. Epstein, I. Torre, A. T. Costa, P. A. D. Gonçalves, N. A. Mortensen, M. Polini, J. C. W. Song, N. M. R. Peres, and F. H. L. Koppens, Quantum nanophotonics in two-dimensional materials, ACS Photon. **8**, 85 (2021).

[60] F. J Garcia-Vidal, C. Ciuti, and T. W. Ebbesen, Manipulating matter by strong coupling to vacuum fields, Science **373**, 6551 (2021).

[61] C. Genet, J. Faist, and T. W. Ebbesen, Inducing new material properties with hybrid light–matter states, Phys. Today **74**, 42 (2021).

[62] F. Schlawin, D. M. Kennes, and M. A. Sentef, Cavity quantum materials, Appl. Phys. Rev. **9**, 011312 (2022).

[63] See the Supplemental Material file for additional numerical results concerning: i) acoustic plasmons and the particle-hole continuum; ii) the impact of LFEs on the plasmonic spectrum; iii) the robustness of the plasmonic spectrum with respect to changes in the filling factor; iv) effects of a static, perpendicular electric field; v) heterostrain effects on plasmons.

[64] R. F. Egerton, Electron energy-loss spectroscopy in the TEM, Rep. Prog. Phys. **72**, 016502 (2009).

[65] S. L. Adler, Quantum theory of the dielectric constant in real solids, Phys. Rev. **126**, 413 (1962).

[66] N. Wiser, Dielectric constant with local field effects included, Phys. Rev. **129**, 62 (1963).

[67] M. Papaj and C. Lewandowski, Probing correlated states with plasmons, Sci. Adv. **9**, 3262 (2023).

[68] A. Tomadin, A. Principi, J. C. W. Song, L. S. Levitov, and M. Polini, Accessing phonon polaritons in hyperbolic crystals by angle-resolved photoemission spectroscopy, Phys. Rev. Lett. **115**, 087401 (2015).

[69] P. Cudazzo, M. Gatti, and A. Rubio, Local-field effects on the plasmon dispersion of two-dimensional transition metal dichalcogenides, New J. Phys. **15**, 125005 (2013).

[70] F. H. da Jornada, L. Xian, A, Rubio, and S. G. Louie, Universal slow plasmons and giant field enhancement in atomically thin quasi-two-dimensional metals, Nat. Commun. **11**, 1013 (2020).

[71] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Maximally localized Wannier orbitals and the extended Hubbard model for twisted bilayer graphene, Phys. Rev. X **8**, 031087 (2018).

[72] P. Lucignano, D. Alfè, V. Cataudella, D. Ninno, and G. Cantele, Crucial role of atomic corrugation on the flat bands and energy gaps of twisted bilayer graphene at the magic angle $\theta \sim 1.08^{\circ}$, Phys. Rev. B **99**, 195419 (2019).

[73] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, Exact continuum model for low-energy electronic states of twisted bilayer graphene, Phys. Rev. Research **1**, 013001 (2019).

[74] S. Carr, S. Fang, and E. Kaxiras, Electronic-structure methods for twisted moirë layers. Nat Rev. Mater. **5**, 748 (2020).

[75] N. C. H. Hesp, I. Torre, D. Rodan-Legrain, P. Novelli, Y. Cao, S. Carr, S. Fang, P. Stepanov, D. Barcons-Ruiz, H. Herzig Sheinfux, K. Watanabe, T. Taniguchi, D. K. Efetov, E. Kaxiras, P. Jarillo-Herrero, M. Polini, and F. H. L. Koppens, Observation of interband collective excitations in twisted bilayer graphene, Nat. Phys. **17**, 1162 (2021).

[76] C. De Beule, P. G. Silvestrov, M.-H. Liu, and P. Recher, Valley splitter and transverse valley focusing in twisted bilayer graphene, Phys. Rev. Research **2**, 043151 (2020).

[77] F. Marsiglio, Eliashberg Theory: A short review, Ann. Phys. **417**, 168102 (2020).

[78] T. Stauber, T. Low, and G. Gómez-Santos, Plasmon-enhanced near-field chirality in twisted van der Waals heterostructures, Nano Lett. **20**, 8711 (2020).

[79] X. Lin, Z. Liu, T. Stauber, G. Gómez-Santos, F. Gao, H. Chen, B. Zhang, and T. Low, Chiral plasmons with twisted atomic bilayers, Phys. Rev. Lett. **125**, 077401 (2020).

Supplemental Material for:
"Theory of intrinsic acoustic plasmons in twisted bilayer graphene"

Lorenzo Cavicchi,[1] Iacopo Torre,[2,3] Pablo Jarillo-Herrero,[4] Frank H. L. Koppens,[3,5] Marco Polini[6,3]

[1]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[2]*Departament de Física, Universitat Politècnica de Catalunya, Campus Nord B4-B5, 08034 Barcelona, Spain*
[3]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*
[4]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[5]*ICREA-Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys 23, 08010 Barcelona, Spain*
[6]*Dipartimento di Fisica dell'Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy*

In this Supplemental Material we present more numerical results for the energy loss function of TBG. We discuss further results concerning: i) acoustic plasmons and the particle-hole continuum; ii) the impact of LFEs on the plasmonic spectrum; iii) the robustness of the plasmonic spectrum with respect to changes in the filling factor; iv) effects of a static, perpendicular electric field; v) heterostrain effects on plasmons.

SECTION I: ACOUSTIC PLASMONS AND THE TBG PARTICLE-HOLE CONTINUUM

In this Section we show that the acoustic plasmon appearing in Fig. 1(b) of the main text lies above the TBG particle-hole continuum (and it is therefore undamped). As discussed in the main text, the upper edge of such continuum is identified by $\hbar\omega_\theta = \hbar v^*_\theta q$. In Figs. S1(a) and (b) we report the plasmon spectrum of TBG for $\theta = 5°$ and $\theta = 6°$, respectively. In each panel, the white dashed line represents the $\hbar\omega_\theta = \hbar v^*_\theta q$ line. We clearly see that, for both twist angles, the acoustic plasmon mode lies above the particle-hole continuum, although falls very close to it.

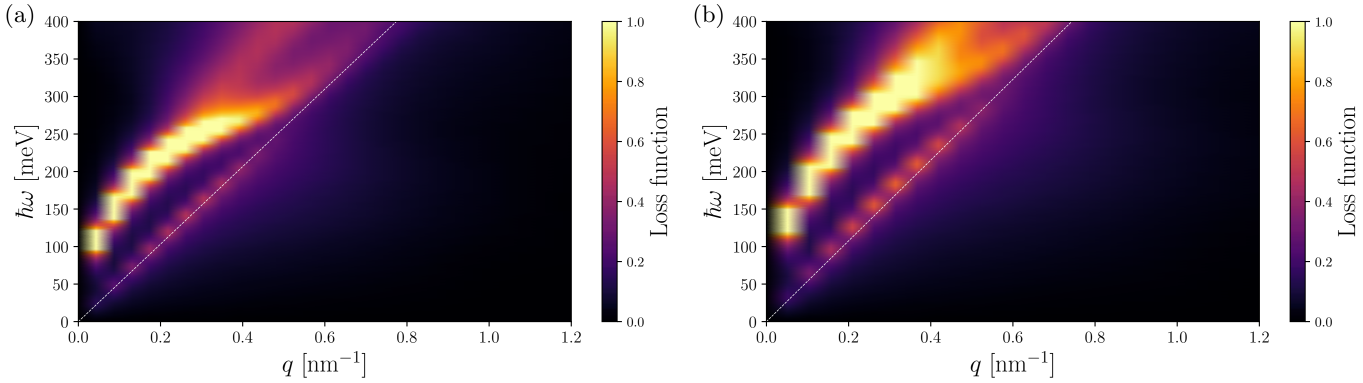


FIG. S1. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ is plotted as a function of $\boldsymbol{q}$ (along the $\Gamma - K$ high symmetry path) and ω for two values of the twist angle θ: $\theta = 5°$ in panel (a) and $\theta = 6°$ in panel (b). Results in this plot refer to filling factor $\nu = +1$ and temperature $T = 5$ K. The upper edge of the particle-hole continuum (which of course depends on θ), i.e. $\omega = v^\star_\theta q$, $v^\star_\theta$ being the reduced Fermi velocity (see Eq. (28) in the main text), is represented by a thin white dashed line.

A plot summarizing the dependence of $c_{\rm s}$ on θ is reported in Fig. S2.

SECTION II: IMPACT OF LFES ON THE PLASMONIC SPECTRUM

In this Section we discuss the role of LFEs on the plasmonic spectrum. Results presented in Figs. S3 and S4 have been obtained by setting $\varepsilon_1 = \varepsilon_3 = 4.9$, $\varepsilon_2 = 1$, and $T = 5$ K. In order to isolate the impact of LFEs, we have deliberately neglected Hartree corrections in producing the data reported in Figs. S3 and S4.

Fig. S3 compares the energy loss function (for $\theta = 1.35^\circ$ and various values of ν) in the local (i.e. $\boldsymbol{G} = \boldsymbol{G}' = \boldsymbol{0}$) approximation (panels in the right column) with that calculated by including LFEs (panels in the left columns). The impact of LFEs is most significant around the edges of the moiré BZ. This is especially true at low doping.

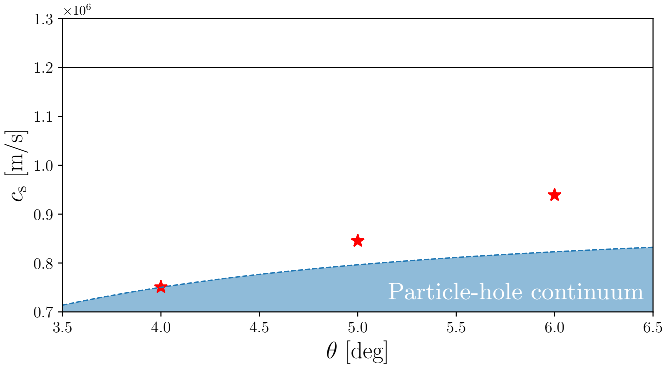


FIG. S2. (Color online) The sound velocity $c_{\rm s}$ (red stars) of the acoustic plasmon is plotted as a function of the twist angle θ. Results in this plot have been obtained by setting $\nu = +1$ and $T = 5$ K. The shaded region identifies the particle-hole continuum, whose upper edge coincides with $v_\theta^\star$ (see Eq. (28) in the main text). The horizontal solid line represents the acoustic plasmon velocity in two spatially-separated graphene layers [S41] at a distance $d = 0.3$ nm and total electron density of $n = 1.5 \times 10^{13}$ cm^{-2}.

A similar comparison is reported in Fig. S4 where the filling factor is fixed at $\nu = +1$ while the twist angle is varied. Increasing the angle leads to a reduction of the importance of LFEs corrections. As emphasized in the main text, for $\theta = 6^\circ$ we can clearly see the acoustic plasmon mode.

SECTION III: ROBUSTNESS OF THE PLAMONIC SPECTRUM WITH RESPECT TO CHANGES IN THE FILLING FACTOR

In this Section we study the robustness of the acoustic plasmon with respect to changes in the filling factor (exploring, in particular, higher values of ν, as compared to the main text).

We evaluated the energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$ with LFEs and Hartree corrections (while taking into account the layer-pseudospin degree of freedom) for $\theta = 1.05^\circ$ and $\theta = 5^\circ$, at filling factor $\nu = +2$. Results are shown in Figure S5. Similarly to Fig. 1 (which refers to $\nu = +1$), we clearly see an intrinsic acoustic plasmon mode for $\theta = 5^\circ$. Note that the peak in the energy loss function associated to the acoustic plasmon is weaker for $\nu = +2$ than $\nu = +1$.

Figure S6 shows the energy loss function for different values of ν and fixed values of the wave vector $\boldsymbol{q}$ taken along the Γ-K direction in the moiré BZ. Results in Fig. S6 have been obtained by neglecting the Hartree contribution. In Figure S7 we show similar results—for two values of θ, i.e. $\theta = 1.05^\circ$ and $\theta = 5^\circ$—but this time with the inclusion of the Hartree contribution. Note the very high level of particle-hole symmetry in the plasmonic spectrum and a significant suppression of the acoustic plasmon for $|\nu| > 2$.

SECTION IV: EFFECTS OF A STATIC, PERPENDICULAR ELECTRIC FIELD

In this Section, we discuss the effect of a static, perpendicular electric field $\boldsymbol{E}_z = E_0 \boldsymbol{z}$ on the plasmonic spectrum of TBG. Within the continuum model introduced in Sect. IV A of the main text, such an electric field is taken into account by adding to the Hamiltonian (11) the following contribution:

$$\hat{\mathcal{H}}_{\rm el} = U_{\rm el} \left(\hat{\Pi}^{(1)} - \hat{\Pi}^{(2)} \right) . \tag{S1}$$

Here, $\hat{\Pi}^{(i)}$ is the projector on the i-th layer defined in Eq. (17) of the main text and

$$U_{\rm el} = \frac{1}{2} e E_0 d \, , \tag{S2}$$

where e is the elementary charge and $d \approx 0.3$ nm the spatial separation between the two graphene layers.

We have evaluated the impact of Eq. (S1) on the energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$, by including LFEs and Hartree corrections. A summary of our main findings is reported in Fig. S8, where we show results obtained for $\theta = 1.05^\circ$—panels (a) and (b)—and $\theta = 5^\circ$—panels (c) and (d). For each of the two twist angles, we considered two values of

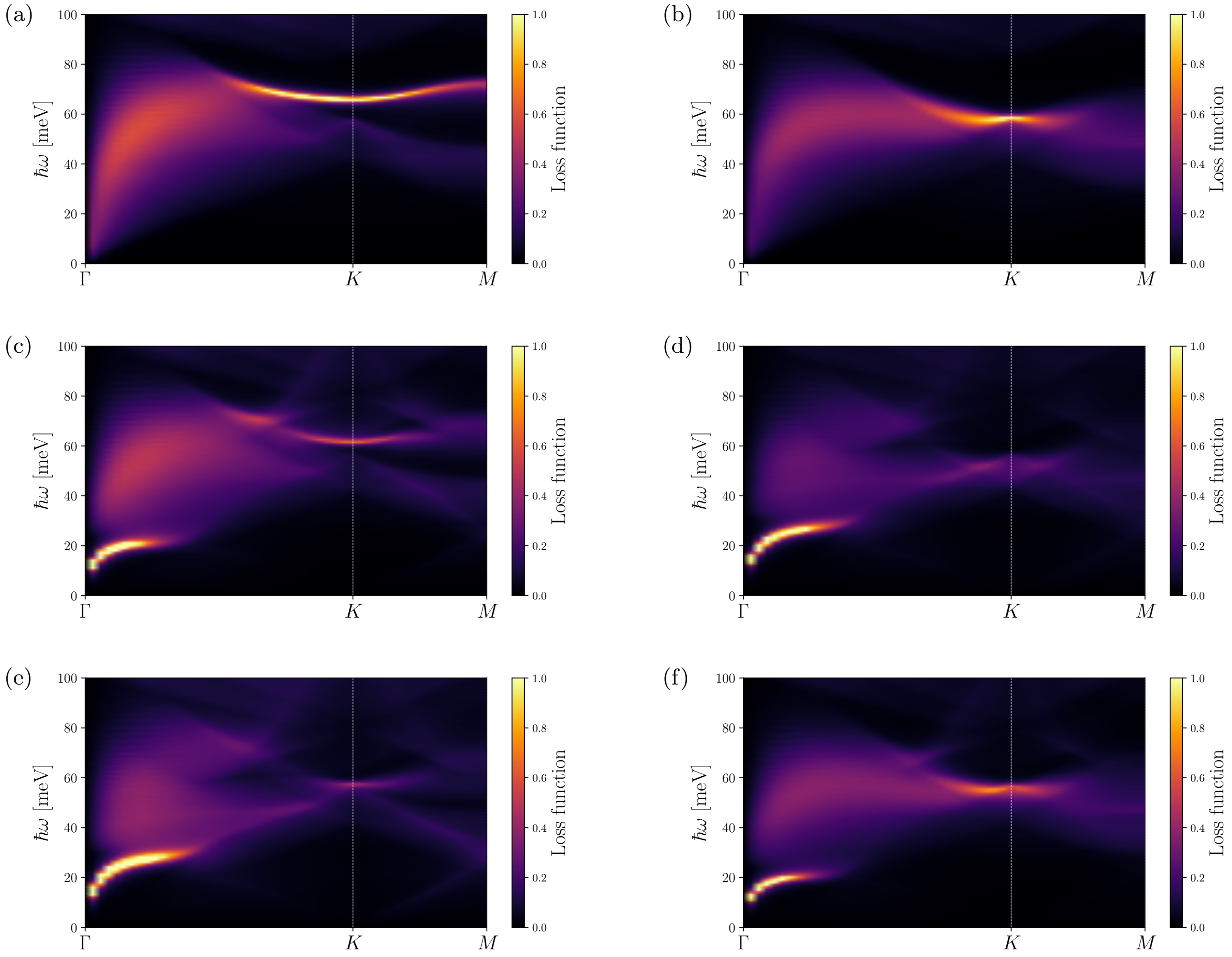


FIG. S3. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$ is plotted as a function of $\boldsymbol{q}$ and ω for $\theta = 1.35^{\circ}$ and various filling factors ν. Results shown in the panels on the left (right) column have been obtained by including (neglecting) LFEs. Panels (a)-(b): $\nu = 0$. Panels (c)-(d): $\nu = +1$. Panels (e)-(f): $\nu = +2$. On the horizontal axis we report $\boldsymbol{q}$ along the high-symmetry path Γ-K-M of the moiré BZ—see Fig. 2(b) in the main text.

the electric field, namely $E_0 = 0.5$ V/nm and $E_0 = 1$ V/nm which correspond to a potential energy of $U_{\rm el} \approx 84$ meV and $U_{\rm el} = 168$ meV, respectively. Clearly, the impact of the perpendicular electric field is more pronounced at small angles, as it suppresses the COM plasmon already for $E_0 = 0.5$ V/nm: see Fig. S8(a). At higher angles, instead, the plasmonic spectrum is only slightly modified, Fig. S8(c)-(d), the main effect of a large applied electric field being the suppression of the acoustic plasmon.

These results can be qualitatively explained as following. The applied perpendicular electric field polarizes TBG, leading to charge accumulation onto one of two layers and a consequent depletion of charge in the other layer. It is precisely this imbalance that is at the origin of the suppression of the COM mode at small twist angles, where the two layers are strongly tunnel-coupled, and of the acoustic plasmon at large twist angles, where the two layers are weakly tunnel-coupled.

SECTION V: HETEROSTRAIN EFFECTS ON PLASMONS

In this Section, we explore the effects of heterostrain on the plasmonic spectrum of TBG. We first briefly review how strain modifies the reciprocal lattice of the bilayer system. We then introduce the continuum model [S1] that describes heterostrained TBG.

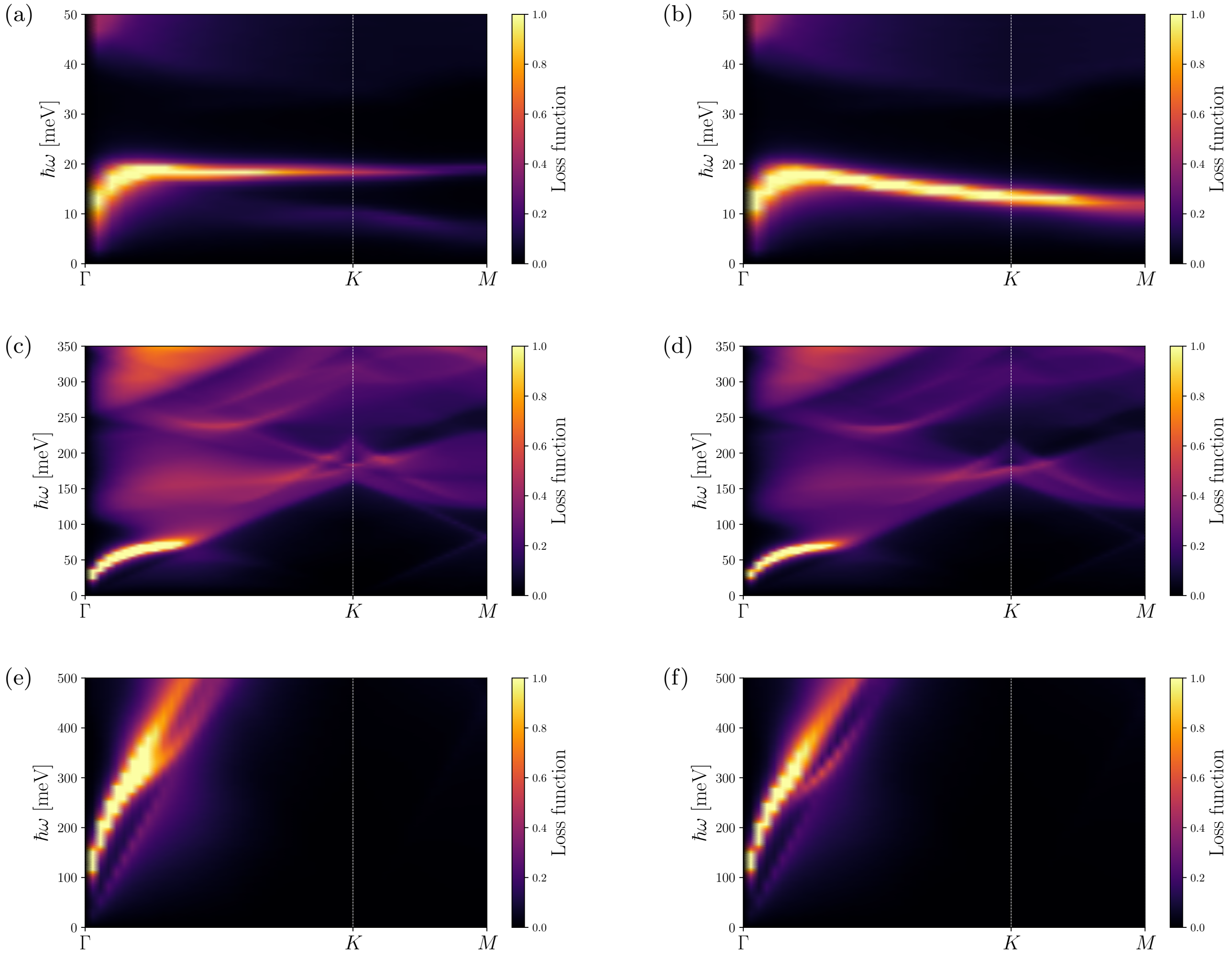


FIG. S4. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ is plotted as a function of $\boldsymbol{q}$ and ω for $\nu=+1$ and various values of the twist angle θ. Results in the left column (i.e. panels (a), (c), and (e)) have been obtained by including LFEs. Results in the right column (i.e. panels (b), (d), and (f)) have been obtained by neglecting LFEs. Panels (a)-(b): $\theta=1.05^{\circ}$. Panels (c)-(d): $\theta=2^{\circ}$. Panels (e)-(f): $\theta=6^{\circ}$. On the horizontal axis we report $\boldsymbol{q}$ along the high-symmetry path Γ-K-M of the moiré BZ—see Fig. 2(b) in the main text.

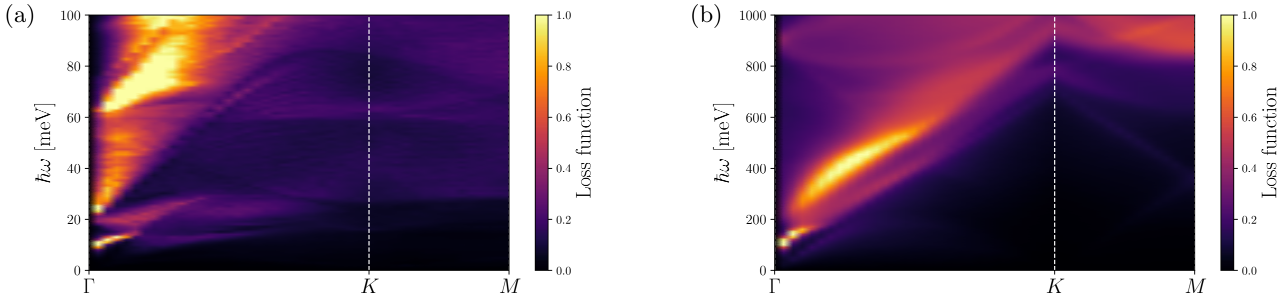


FIG. S5. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ is plotted as a function of $\boldsymbol{q}$ and ω for $\nu=+2$. Results in this figure have been obtained by taking into account both LFEs and Hartree corrections. Panel (a) $\theta=1.05^{\circ}$. Panel (b) $\theta=5^{\circ}$. On the horizontal axis we report $\boldsymbol{q}$ along the high-symmetry path Γ-K-M of the moiré BZ—see Fig. 2(b) in the main text.

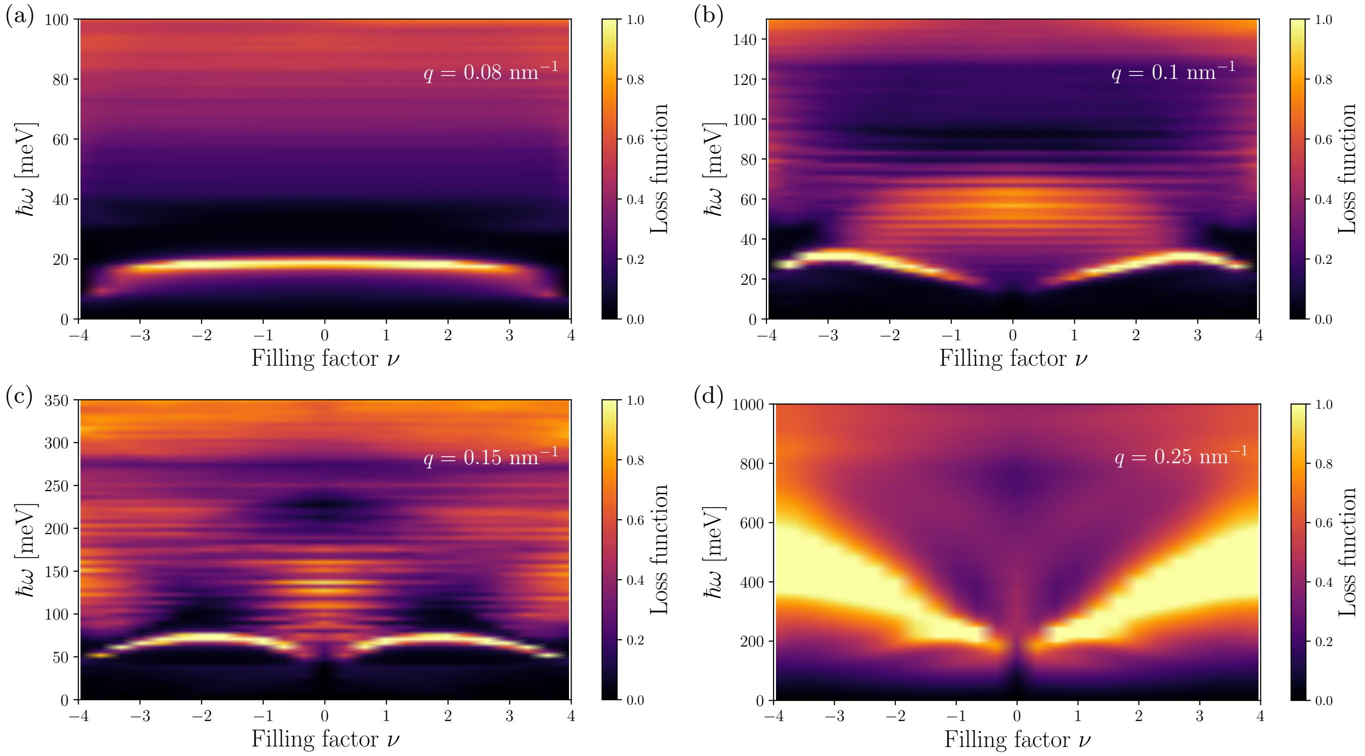


FIG. S6. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$ is plotted as a function of ω and filling factor ν. Each panel corresponds to a value of θ and q (the latter taken along the Γ-K direction of the moiré BZ). Results in this figure have been obtained by *neglecting* Hartree corrections. Panel (a) $\theta = 1.05^{\circ}$. Panel (b) $\theta = 1.35^{\circ}$. Panel (c) $\theta = 2^{\circ}$. Panel (d) $\theta = 5^{\circ}$.

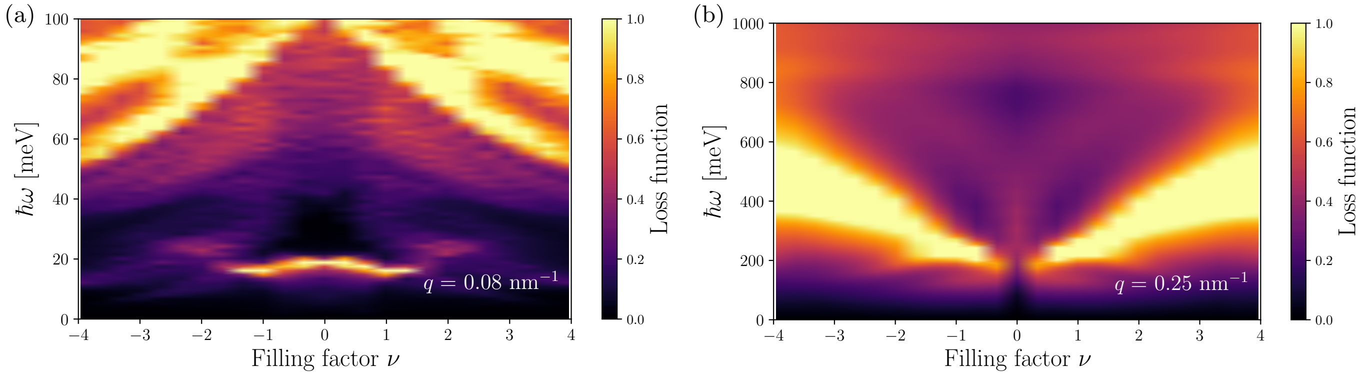


FIG. S7. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$ is plotted as a function of ω and filling factor ν. Each panel corresponds to a value of θ and q (the latter taken along the Γ-K direction of the moiré BZ). Results in this figure have been obtained by *including* Hartree corrections. Panel (a) $\theta = 1.05^{\circ}$. Panel (b) $\theta = 5^{\circ}$.

Heterostrain refers to relative strains between layers, and can be present in experimental TBG samples either because of unwanted interactions with the substrate or because intentionally applied and controlled by piezoelectrics. The properties of heterostrained TBG can be captured, in the small deformation and small rotation limit, by the following deformation matrices:

$$\mathcal{E}^{(\ell)} = \begin{pmatrix} \epsilon_{xx}^{(\ell)} & \epsilon_{xy}^{(\ell)} - (-)^{\ell}\theta/2 \\ \epsilon_{yx}^{(\ell)} + (-)^{\ell}\theta/2 & \epsilon_{yy}^{(\ell)} \end{pmatrix} . \tag{S3}$$

The strained geometry leads to a deformed moiré reciprocal lattice. This can be constructed according to the following

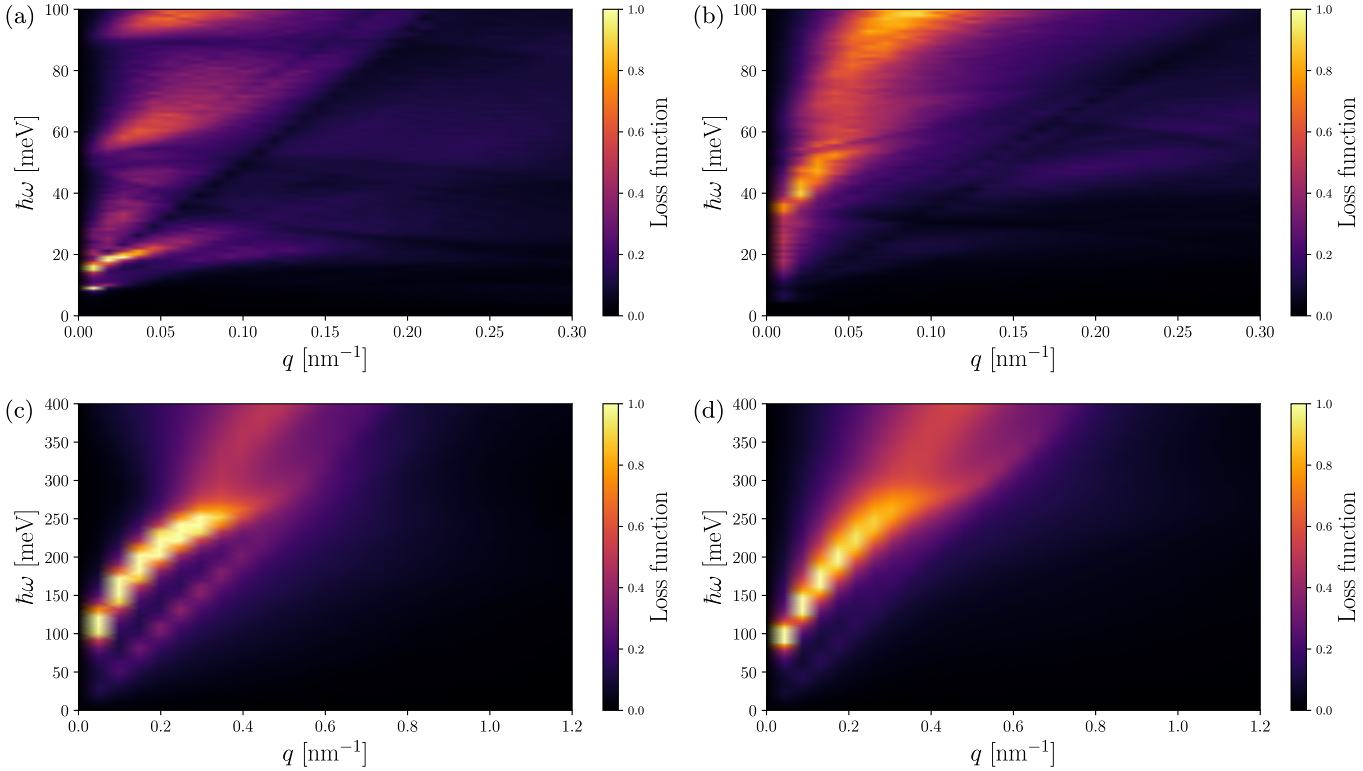


FIG. S8. (Color online) Dependence of the plasmonic spectrum on a static, perpendicular electric field. The TBG energy loss function $\mathcal{L}(\boldsymbol{q},\omega)$ is plotted as a function of $\boldsymbol{q}$ and ω for two values of $U_{\rm el}$: see Eq. (S2). On the horizontal axis we report $\boldsymbol{q}$ along the high-symmetry path Γ-K of the moiré BZ: see Fig. 2(b) in the main text. The left (right) column refers to $E_0 = 0.5$ V/nm ($E_0 = 1$ V/nm). Panels (a)-(b): $\theta = 1.05^\circ$. At this small twist angle, the applied electric field suppresses the COM mode. Panels (c)-(d): $\theta = 5^\circ$. At this larger value of the twist angle, instead, the main effect of the electric field is to suppress the acoustic mode.

equation,

$$\tilde{\boldsymbol{G}}_i = \mathcal{E}^{\rm T}\boldsymbol{g}_i \ , \tag{S4}$$

where $\boldsymbol{g}_1 = \frac{4\pi}{\sqrt{3}a}(\frac{\sqrt{3}}{2}, -\frac{1}{2})$ and $\boldsymbol{g}_2 = \frac{4\pi}{\sqrt{3}a}(0,1)$ are the unstrained and untwisted reciprocal lattice vectors and $\mathcal{E} \equiv \mathcal{E}^{(2)} - \mathcal{E}^{(1)}$ is the relative deformation matrix. The quantity $\tilde{\boldsymbol{G}}_i$ defines the strained moiré reciprocal lattice counterpart of the unstrained reciprocal lattice obtained by the vectors defined in the main text in Eq. (16).

In what follows, we further assume $\mathcal{E}^{(2)} = -\mathcal{E}^{(1)} = \frac{1}{2}\mathcal{E}$, as in Ref. [S1] and limit our investigation to uniaxial heterostrain. This type of heterostrain involves the application of stress predominantly along one direction of the bilayer system while leaving the perpendicular direction unstressed. With all these restrictions, the strain part of the relative deformation matrix $\mathcal{E}$ can be expressed with only three parameters: strain magnitude ϵ, strain direction ϕ, and Poisson ration $\nu_{\rm p}$, which takes the value $\nu_{\rm p} \approx 0.16$ in graphene. We find

$$\mathcal{E} = \mathcal{R}^{-1}(\phi)\begin{pmatrix} -\epsilon & 0 \\ 0 & \nu_{\rm p}\epsilon \end{pmatrix}\mathcal{R}(\phi) + \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \ , \tag{S5}$$

where the rotation matrix $\mathcal{R}(\phi)$ is given by

$$\mathcal{R}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \ . \tag{S6}$$

We now move on to describe the continuum model Hamiltonian for the uniaxial heterostrained TBG. The structure of the Hamiltonian operator is the same as we described in Sect. IV of the main text, with some modifications. Within the two-center approximation, the effect of strain on the intra-layer Hamiltonian can be described by an effective vector

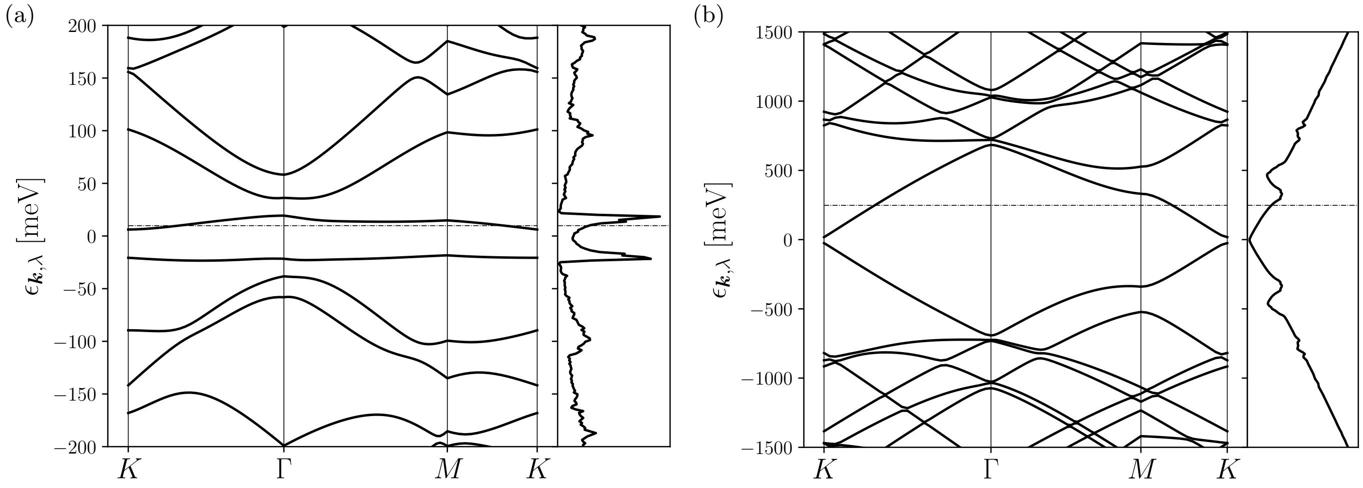


FIG. S9. (Color online) The energy bands of TBG and their corresponding density of states calculated by taking into account uniaxial heterostrain. On the horizontal axis we report the momentum $\boldsymbol{k}$ along the high-symmetry path K-Γ-M-K of the strained moiré BZ. In order to obtain these results we fixed $\varepsilon = 0.6\%$ and $\phi = 30^{\circ}$. In panel (a) the twist angle is $\theta = 1.05^{\circ}$, while in panel (b) $\theta = 5^{\circ}$. In both panels the dashed-dot line shows the chemical potential μ for filling factor $\nu = +1$ and temperature $T = 5$ K.

potential (gauge field) [S1, S2]:

$$\hat{\mathcal{H}}^{(\ell)}_{\xi} = v_{\rm D}\left[\left(\mathbb{I}+\mathcal{E}^{\rm T}\right)\left(\hat{\boldsymbol{p}} - \hbar\tilde{\boldsymbol{K}}_{\xi,\ell} + \xi\boldsymbol{A}_{\ell}\right)\right]\cdot\left(\xi\sigma_x, -\sigma_y\right) \ . \tag{S7}$$

Here $\xi = \pm$ is the valley index, $\tilde{\boldsymbol{K}}_{\xi,\ell}$ is the K point of the strained mBZ, and $\boldsymbol{A}_{\ell}$ is the effective vector potential defined as:

$$\boldsymbol{A}_2 = -\boldsymbol{A}_1 = \frac{\sqrt{3}}{4a}\beta\epsilon(1+\nu_{\rm p})\left(\cos(2\phi), \sin(2\phi)\right) \ , \tag{S8}$$

with $\beta \approx 3.14$ in graphene. Concerning inter-layer tunneling, it should be expressed in terms of the strained reciprocal lattice basis vectors introduced in Eq. (S4). With respect to the main text expression, cf. Eq. (15), we have a slightly different formula that takes into account the different reciprocal lattice basis:

$$\hat{U} = \begin{pmatrix} u_0 & u_1 \\ u_1 & u_0 \end{pmatrix} + e^{i\xi\tilde{\boldsymbol{G}}_1\cdot\hat{\boldsymbol{r}}}\begin{pmatrix} u_0 & u_1 e^{-i\frac{2\pi}{3}} \\ u_1 e^{i\frac{2\pi}{3}} & u_0 \end{pmatrix} + e^{i\xi\left(\tilde{\boldsymbol{G}}_1+\tilde{\boldsymbol{G}}_2\right)\cdot\hat{\boldsymbol{r}}}\begin{pmatrix} u_0 & u_1 e^{i\frac{2\pi}{3}} \\ u_1 e^{-i\frac{2\pi}{3}} & u_0 \end{pmatrix} \ . \tag{S9}$$

Moiré minibands as modified by uniaxial heterostrain are displayed in Fig. S9. Results in this figure refer to $\epsilon = 0.6\%$ and strain direction $\phi = 30^{\circ}$. The impact of uniaxial heterostrain on the energy loss function is illustrated in Figs. S10 and S11. In Fig. S10 the energy loss function is displayed for $\epsilon = 0.6\%$ and $\phi = 30^{\circ}$. Instead, in Fig. S11, we show results for the energy loss function at fixed momentum $q = \xi|\boldsymbol{K}_{1,2}|$, where $|\boldsymbol{K}_1| = |\boldsymbol{K}_2|$ is the modulus of the vector linking Γ to K in the moiré BZ (which depends on the strain magnitude ϵ). In Fig. S11 we used $\xi = 4/25$, and varied ϵ at fixed $\phi = 0^{\circ}$. In order to obtain these plots, Hartree contributions have been neglected, while we have retained LFEs. These results suggest that, at small twist angles, heterostrain largely suppresses the COM plasmon, which becomes extremely feeble and flat, while it pushes inter-band excitations toward lower energies: see Figs. S10(a) and S11(a). At larger angles, instead, the effect of strain is less severe, suggesting resilience of the acoustic plasmon for $\epsilon \lesssim 2.5\%$, as show in Figs. S10(b) and S11(b).

[S1] Z. Bi, N. F. Q. Yuan, and L. Fu, Designing flat bands by strain, Phys. Rev. B **100**, 035448 (2019).
[S2] N. N. T. Nam and M. Koshino, Lattice relaxation and energy band modulation in twisted bilayer graphene, Phys. Rev. B **96**, 075311 (2017).

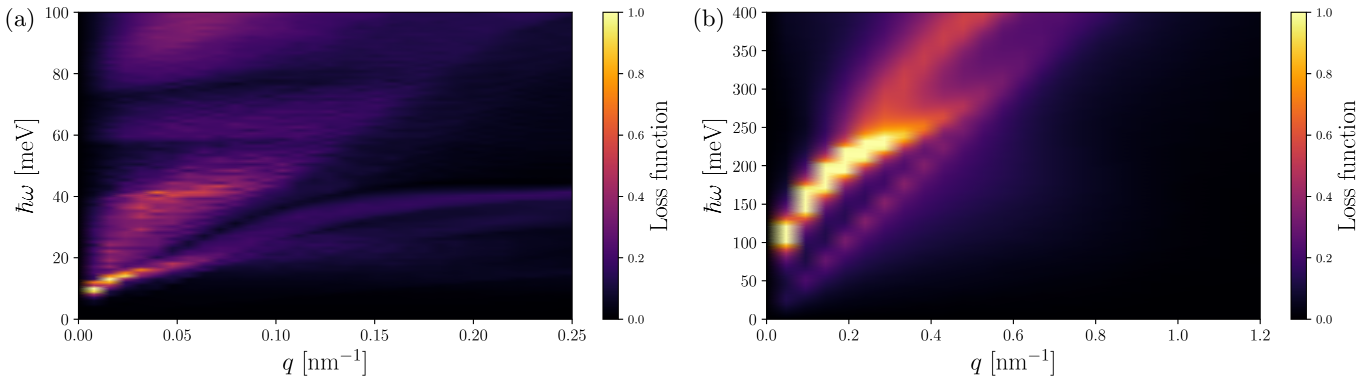


FIG. S10. (Color online) The impact of uniaxial heterostrain on the TBG energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$, which is plotted, as usual, as a function of $\boldsymbol{q}$ (along the high-symmetry path Γ-K of the strained mBZ) and ω. Results in this figure have been obtained by choosing $\epsilon = 0.6\%$ and $\phi = 30^{\circ}$. Panel (a): $\theta = 1.05^{\circ}$. We clearly see that heterostrain largely suppresses the COM plasmon while pushing the inter-band transitions toward lower energies. Panel (b): $\theta = 5^{\circ}$. In this case heterostrain does not affect the plasmonic spectrum as a direct comparison with e.g. Fig. S1 shows. In both panels the filling factor is fixed at $\nu = +1$ and temperature is $T = 5$ K.

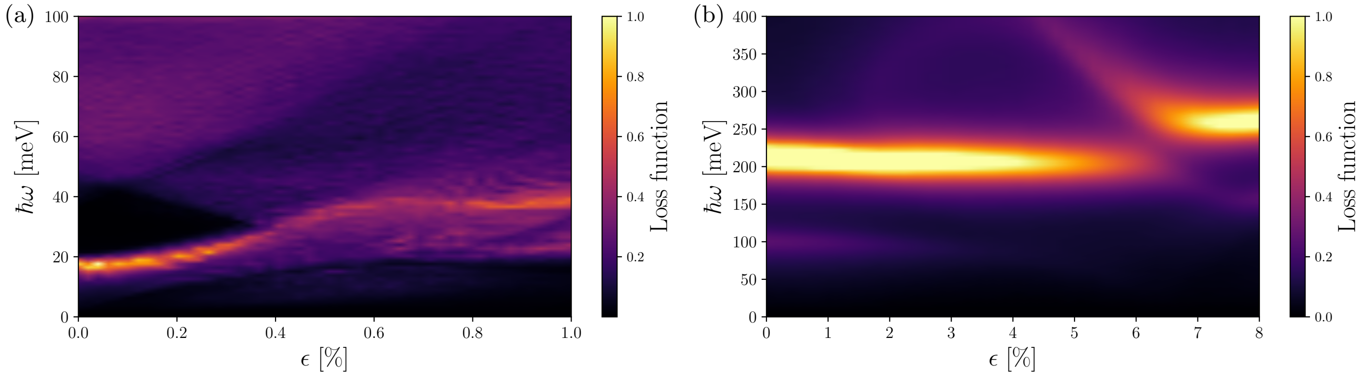


FIG. S11. (Color online) The TBG energy loss function $\mathcal{L}(\boldsymbol{q}, \omega)$ is plotted as a function of $\hbar\omega$ and strain magnitude ϵ. Results in this figure have been obtained by setting $q = \xi |\boldsymbol{K}_{1,2}|$ (see Sect. V), $\phi = 0^{\circ}$, $\nu = +1$, and $T = 5$ K. Panel (a): $\theta = 1.05^{\circ}$. We clearly see that for $\epsilon \gtrsim 0.4\%$ the COM mode is suppressed and merges with inter-band plasmons. Panel (b): $\theta = 5^{\circ}$. We clearly see that the acoustic plasmon, which in this figure is the feeble feature at $\hbar\omega \approx 100$ meV, is visible for $\epsilon \lesssim 2.5\% - 3.0\%$.

Interlayer electron-hole friction in tunable twisted bilayer graphene semimetal

D.A. Bandurin,[1,2] A. Principi,[3] I.Y. Phinney,[1] T. Taniguchi,[4] K. Watanabe,[5] and P. Jarillo-Herrero[1]

[1] *Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *Department of Material Science and Engineering, National University of Singapore, 117575 Singapore*
[3] *School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom*
[4] *International Center for Materials Nanoarchitectonics, National Institute of Material Science, Tsukuba 305-0044, Japan*
[5] *Research Center for Functional Materials, National Institute of Material Science, Tsukuba 305-0044, Japan*

arXiv:2208.05659v1 [cond-mat.mes-hall] 11 Aug 2022

Charge-neutral conducting systems represent a class of materials with unusual properties governed by electron-hole (e-h) interactions. Depending on the quasiparticles statistics, band structure, and device geometry these semimetallic phases of matter can feature unconventional responses to external fields that often defy simple interpretations in terms of single-particle physics. Here we show that small-angle twisted bilayer graphene (SA-TBG) offers a highly-tunable system in which to explore interactions-limited electron conduction. By employing a dual-gated device architecture we tune our devices from a non-degenerate charge-neutral Dirac fluid to a compensated two-component e-h Fermi liquid where spatially separated electrons and holes experience strong mutual friction. This friction is revealed through the T^2 resistivity that accurately follows the e-h drag theory we develop. Our results provide a textbook illustration of a smooth transition between different interaction-limited transport regimes and clarify the conduction mechanisms in charge-neutral SA-TBG.

Low-dimensional electron-hole (e-h) systems have recently emerged as an important platform in which to explore many-body quantum phenomena. In such systems, strong Coulomb interaction among electrons and holes can give rise to a plethora of exotic quantum phases whose inventory encompasses superfluids[1,2], correlated density wave states[3,4], excitonic insulators[5,6], and Wigner crystals[4,7], to name a few. Particularly interesting interacting e-h mixtures are hosted by graphene and its bilayer. Graphene-based devices enabled the discovery of novel non-trivial effects governed by e-h interactions: from the Wiedemann-Franz law violation[8] and the anomalous Coulomb drag[9–14] to the quantum critical conductivity[15–17] and giant thermal diffusivity[18]. Central in these effects is the dominance of momentum-conserving e-h collisions over other momentum-relaxing scattering processes brought upon by graphene's weak electron-phonon coupling and low disorder[19]. As a result, the behavior of graphene's e-h plasma at elevated temperatures T, often referred to as Dirac fluid, resembles that of interacting relativistic fluids governed by the laws of (relativistic) hydrodynamics[8,19–22]. Since hydrodynamics offers a natural framework by which to probe the long-wavelength behavior of strongly-interacting fluids, experiments on model platforms, such as graphene, can give insights for observations in more exotic quantum phases of matter[23,24], substantiating the interest in the field.

So far, the hydrodynamic behavior of interacting e-h plasmas in mono- and bilayer graphene (MLG and BLG respectively) was explored deep in the non-degenerate limit ($E_\mathrm{F} \ll k_\mathrm{B}T$, where E_F is the Fermi energy, k_B is the Boltzmann constant) and relied on thermal[8,25,26], light-[18] or current-driven[27] excitation of e-h pairs. The ambipolar hydrodynamics in the degenerate regime ($E_\mathrm{F} \gg k_\mathrm{B}T$) as well as its genesis from the Boltzmann phase have at present remained inaccessible. This inaccessibility stems from the fact that the conduction and valence band extrema in MLG and BLG coincide in momentum space and thus the e-h system can only be realized through the smearing of the charge neutrality point (NP); adding more carriers into the system converts the neutral Dirac fluid into a unipolar Fermi liquid (FL)[19]. In this work, we introduce biased SA-TBG as a convenient system in which to explore a smooth crossover between the Dirac fluid regime and the regime of degenerate e-h FL. In the latter case, we demonstrate that frequent momentum-conserving (yet velocity-relaxing) e-h collisions are the limiting factor for the SA-TBG conductivity.

We start by exploring the single-particle band structure of SA-TBG which is folded within a reduced Brillouin zone (BZ)[30] due to superlattice periodicity (Fig. 1a-b). At small energies, it resembles that of MLG but is characterized by a decreased Fermi velocity v_F. Like the BZ of MLG, the reduced BZ of SA-TBG is hexagonal and comprises two minivalleys located at the k_m and k'_m high symmetry points. These coincide with the K points of the two decoupled graphene sheets[30]. A prominent feature of the SA-TBG is that, away from the magic angle ($\theta \gtrsim 1.3°$), one can selectively populate its minivalleys with charge carriers of opposite types using a perpendicular displacement field, D, (Fig. 1b)[3,31–35]. Electrostatic calculations[32] for $D = 1$ V/nm, reveal that such a strong D, readily achievable in experiments, can result in the formation of relatively large electron and hole Fermi surfaces in the k_m and k'_m minivalleys, respectively. Quantitatively, in each minivalley, the Fermi temperature, T_F, exceeds room T, as in normal FLs (Fig. 1c dashed line). On the contrary, charge-neutral SA-TBG at $D = 0$ is half-filled up to the Dirac point where the Fermi surfaces shrink to two points and where the Dirac fluid emerges at elevated T[8,19]. This tunability enables the exploration

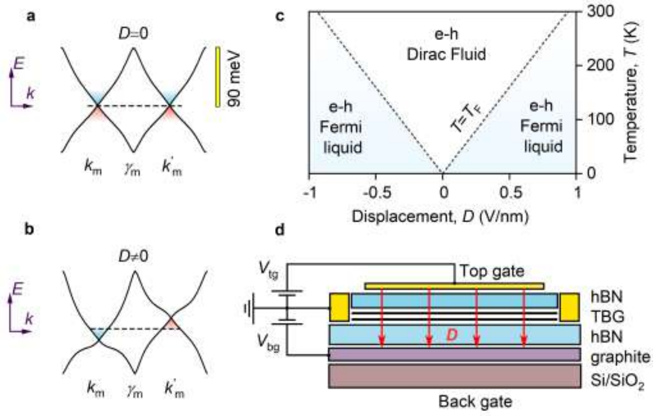


FIG. 1. **Biased SA-TBG. a-b,** Calculated single-particle band structure for 1.65° SA-TBG[28,29]. At low-energies, two Dirac cones are formed in the vicinity of the k_m and k'_m points (a); when $D \neq 0$, the cones are shifted with respect to each other (b). The horizontal dashed lines represent the Fermi level in the neutral SA-TBG. **c,** Phase diagram for the charge-neutral e-h mixture in SA-TBG mapped onto a $T-D$ plane. Dashed lines: the dependence of T_F in each minivalley on D for $n=0$. **d,** Schematic of the dual-gated encapsulated SA-TBG device.

of e-h plasma at the crossover between the Dirac fluid and FL regimes in standard transport experiments as we schematically illustrate on the $D-T$ diagram in Fig 1c.

To probe such a crossover, we fabricated a dual-gated multi-terminal Hall bar made out of $\theta \approx 1.65°$ SA-TBG encapsulated between two relatively thin ($<$ 100 nm thick) slabs of hexagonal boron nitride (hBN). At this angle, the SA-TBG is characterized by enhanced interaction strength and a reduced v_F, but is far enough from the magic angle (1.1°) that it allows for appreciable interlayer polarization[3,33]. The device was produced by a combination of tear-and-stack[36–38] and hot release[39] methods, and had a width of 2 μm (Inset of Fig. 2b) (Supplementary Section 1). The dual-gated configuration (Fig. 1d) allowed us to control the interlayer displacement $D/\epsilon_0 = (C_\mathrm{bg}V_\mathrm{bg} - C_\mathrm{tg}V_\mathrm{tg})/2$, and the total externally-induced carrier density, $n = (C_\mathrm{bg}V_\mathrm{bg} + C_\mathrm{tg}V_\mathrm{tg})/e$, where $C_\mathrm{tg,bg}$ are the top and bottom gate capacitance per unit area, ϵ_0 is the dielectric permittivity of vacuum, and e is the electron charge.

Figure 2a shows an example of the longitudinal resistivity, ρ_{xx}, dependence on V_bg and V_tg in a form of 2D map, measured in our SA-TBG and reveals its characteristic behavior. Namely, the map consists of three diagonal lines: central - that denotes the global neutrality, and two side diagonals, labeled as BI, that reflect the full filling of the first miniband where the single-particle band insulator emerges[37,38,40]. The BI lines allow for an accurate determination of the twist angle[37,38,40]. Below, we will only focus on the region in the vicinity of the global neutrality and away from the van Hove singularity.

Figure 2b shows the $\rho_\mathrm{xx}(n)$ dependence of our SA-TBG device measured at $D=0$ and $T=4.2$ K (the curve is measured along the blue trace in the map from Fig. 2a). At $D=0$, $\rho_\mathrm{xx}(n)$ exhibits a sharp peak and reaches 2.7 kΩ at $n=0$, a standard behavior for SA-TBG devices. The peak width is only $\delta n \simeq \times 10^{10}$ cm^{-2} that indicates low charge inhomogeneity provided by the graphite gate[41]. Upon doping, $\rho_\mathrm{xx}(n)$ rapidly decreases and already at 10^{12} cm^{-2} drops to 30 Ω which translates to the 1.7 μm mean free path, obtained from the standard Drude model. At liquid helium T, we also observed negative transfer resistance measured in the bend geometry (Supplementary Information 2), an indicative of the micrometre-scale ballistic transport[42,43]. These observations highlight an exceptional quality of our encapsulated SA-TBG device critical for further exploration of interaction-dominated transport at elevated T as we now proceed to discuss.

With the application of D, the transport properties of neutral SA-TBG change drastically (Fig. 2b, red curve). ρ_xx at the NP drops by more than an order of magnitude and becomes comparable to that of doped SA-TBG (cf. ρ_xx at 10^{12} cm^{-2}). This qualitative behavior remains unchanged upon increasing T (Fig. 2c). Namely, at $T=20$ K the NP resistivities measured at zero and finite D differ by more than an order of magnitude. The drop of ρ_xx with increasing D signals parallel conduction of two minivalleys when each of them is doped away from their NPs.

We further studied the temperature dependence of our sample's resistivity. Figures 3a-b shows $\rho_\mathrm{xx}(n)$ dependencies for varying T for the case of zero (a) and finite (b) D respectively. Away from the NP ($n=0$), ρ_xx grows with increasing T for both D values, indicating characteristic behavior of doped graphene sheets. On the contrary, at the NP, ρ_xx exhibits a very different behavior for the two cases. Namely, at $D=0$, ρ_xx drops rapidly when T is raised from 4.2 to 40 K (inset of Fig. 3a), whereas at $D=0.7$ V/nm, ρ_xx shows a clear metallic trend: the resistivity increases with increasing T (inset of Fig. 3b).

It is now instructive to normalize all measured $\rho_\mathrm{xx}(T)$ dependencies to their lowest T value in order to compare the functional forms of the $T-$dependencies in different cases. At $T=40$ K, the zero-D resistivity of the SA-TBG device is less than a half of its 4.2 K value; further increase of T leads to a very slow ascending trend of $\rho_\mathrm{xx}(T)$. At the same T and $D=0.7$ V/nm, ρ_xx experiences more than two times increase and keeps growing with increasing T following approximately an $a+bT^2$ dependence, where a and b are constants (dashed black line in the inset of Fig. 3b). To compare, we have also measured the resistivity of a BLG device of comparable quality as a function of n and T (Fig. 3c). At the NP, ρ_xx is practically unaffected by the T variation (Fig. 3c) over the entire range of T in our experiments.

The above observations clearly point to the difference in the conductivity mechanisms of these three bilayer systems at their NPs. The weak insulating behavior of charge-neutral SA-TBG at zero D resembles that of

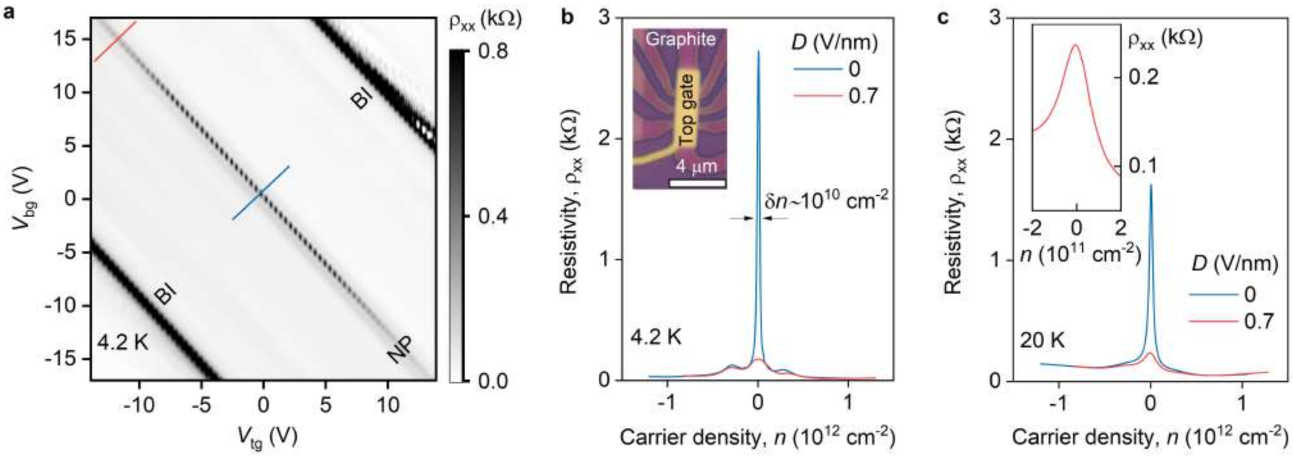


FIG. 2. **Effect of displacement on the transport properties of the SA-TBG.** **a,** $\rho_{\rm xx}$ as a function of $V_{\rm bg}$ and $V_{\rm tg}$ measured in the 1.65° SA-TBG device. Blue and red lines correspond to the $(V_{\rm tg}, V_{\rm bg})$ points where $D = 0$ and $D = 0.7$ V/nm respectively. **b,** $\rho_{\rm xx}(n)$ traces for $D = 0$ and $D = 0.7$ V/nm measured at $T = 4.2$ K. Inset: Optical photograph of an encapsulated SA-TBG device. **c,** Same as (b) but for $T = 20$ K. Inset: zoomed-in region of the NP vicinity for $D = 0.7$ V/nm.

MLG: the resistivity drops as a result of the thermal activation of electrons and holes[8]. A further increase of T leads to the enhanced scattering between electron and hole non-degenerate sub-systems hosted by SA-TBG leading to an increase of the resistivity. In contrast, the flat $T-$dependence of the BLG has been recently attributed to the perfect balance between the amount of thermally activated e-h pairs facilitating conductivity, and the e-h scattering that impedes the electrical current[22,26,44]. The peculiar T^2 growth of the resistivity in compensated SA-TBG at finite D has not been observed previously. Below we show that this effect stems from the e-h friction[45,46] in this degenerate ambipolar system.

To demonstrate this, we solve the steady-state Boltzmann equation for e-h hole mixture in SA-TBG; the details are given in Supplementary Information. In the limit of temperatures much smaller than $T_{\rm F}$, the resistivity due to e-h scattering reads

$$\rho_{\rm D} \simeq \frac{8\pi\alpha_{\rm ee}^2 g(\bar{q}_{\rm TF})}{3ne^2 v_{\rm F}^2\hbar}(k_{\rm B}T)^2\ . \qquad (1)$$

where n is the particle density in each minivalley, $g(\bar{q}_{\rm TF}) = 3(\bar{q}_{\rm TF} - 1) + (4 - 3\bar{q}_{\rm TF}^2){\rm arccoth}(1 + \bar{q}_{\rm TF})$ and $\bar{q}_{\rm TF} = N_{\rm f}\alpha_{\rm ee}$ is the Thomas-Fermi screening wavevector in units of the Fermi wavevector. Here, $\alpha_{\rm ee} = e^2/(2\pi\epsilon_0(\epsilon_r + 1)\hbar v_{\rm F})$ is the effective fine-structure constant of Dirac fermions, ϵ_r is a dielectric constant accounting for screening due to far bands and external dielectrics, $N_{\rm f}$ is the number of flavors, and $\hbar$ is the reduced Planck constant. Hereafter we set $\epsilon_r = 3.9$, as for graphene deposited on hBN. The total resistivity is then $\rho = \rho_0 + \rho_{\rm D}$, where ρ_0 is the zero-temperature resistivity due to momentum-non-conserving scattering processes. We also note that, as the minivalleys are predominantly formed from the energy bands of different graphene sheets, electrons and holes reside in the upper or lower graphene layers depending on the D direction[3,32,33], and thus $\rho_{\rm D}$ can be interpreted as the resistivity due to the interlayer e-h friction.

In Fig. 3d we compare the results of our calculations with $\rho_{\rm xx}(T)$ found experimentally. To this end, we plot the experimentally found resistivity excess, $\Delta\rho = \rho_{\rm xx}(T) - \rho_{\rm xx}(4.2\ K)$, and theoretically obtained $\rho_{\rm D}(T)$. For the latter, we used an electrostatic model that accounts for screening effects to calculate the Fermi energy in each minivalley[32], as well as the experimentally determined twist angle. Using that, for $\theta = 1.65°$, $v_{\rm F} \simeq 5 \times 10^5$ m/s (as determined from the continuum model of SA-TBG[28–30]), for $D = 0.7$ V/m we estimate the carrier density $n = 1.3 \times 10^{15}$ m^{-2}. Experimental data follows closely the expected $\mathcal{B}T^2$ dependence with $\mathcal{B} \simeq 0.062$ Ω/K^2 with some tendency to sub-quadratic dependence at higher T (inset of Fig. 3b). This deviation from the T^2 scaling can be attributed to the thermal smearing of the distribution function that leads to the exit of the SA-TBG e-h system from the degenerate state. Indeed, at $n = 1.3 \times 10^{11}$ cm^{-2}, the Fermi temperature of the 1.65° SA-TBG is of the order of 220 K.

Next, we analyze $\rho_{\rm D}(T)$ dependencies expected for other θ. We find that, at fixed carrier density, the resistivity due to e-h scattering depends on θ only through its dependence on the electron Fermi velocity $v_{\rm F}$. The latter controls the values of both the Fermi energy $\varepsilon_{\rm F}$ and effective fine-structure constant $\alpha_{\rm ee}$. The relationship between the Fermi velocity and twist angle can be obtained from a continuum model of SA-TBG[30]. In the inset of Fig. 3 we plot the ratio $\mathcal{B}(v_{\rm F})/\mathcal{B}(v_{\rm F}^{\rm g})$ for a carrier density $n = 4 \times 10^{14}$ m^{-2} as a function of θ. Here, $v_{\rm F}^{\rm g}$ is the Fermi velocity of MLG, while $\mathcal{B}(v_{\rm F})$ is defined from $\rho_{\rm D} = \mathcal{B}(v_{\rm F})T^2$. At $\theta > 3°$ the e-h drag would result in a 10 times smaller prefactor of the T^2- resistivity

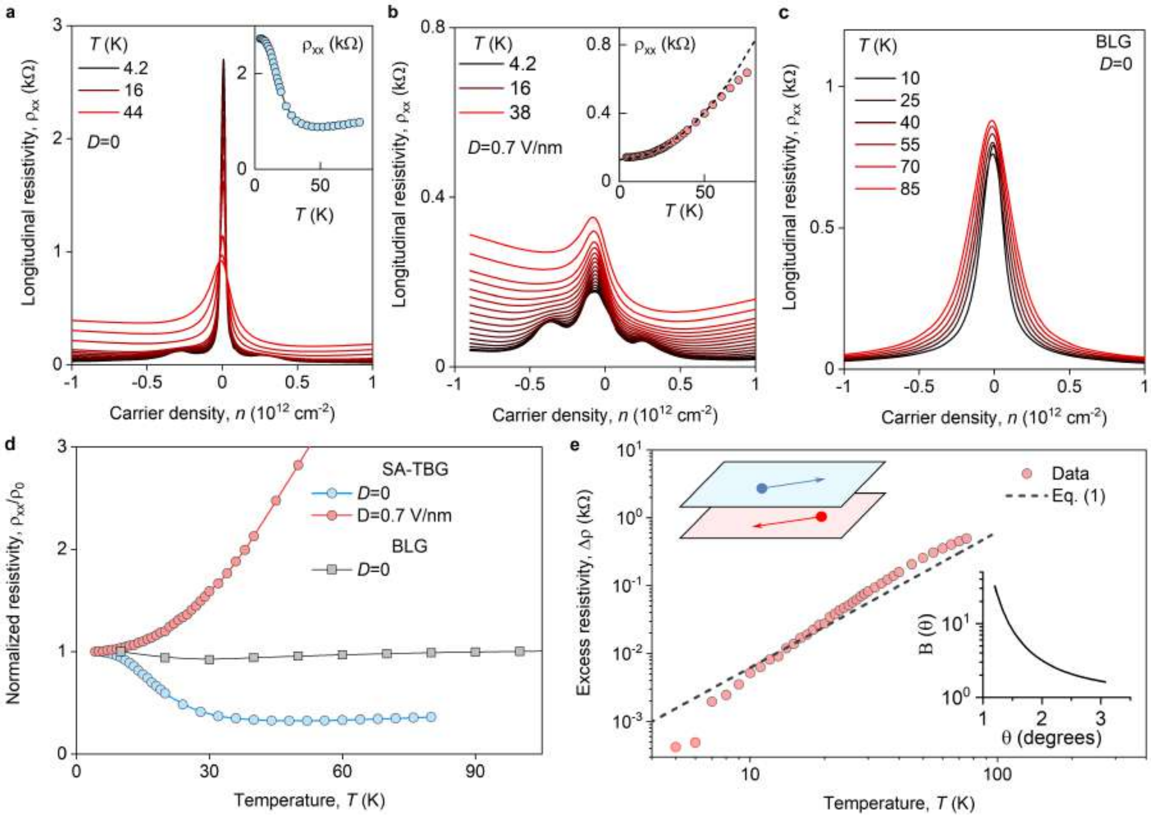


FIG. 3. **Temperature dependence of the SA-TBG resistivity a,** $\rho_{\mathrm{xx}}(n)$ for different T for the case of $D=0$. Inset: $\rho_{\mathrm{xx}}(T)$ at the NP and $D=0$. **b,** Same as (a) but for $D=0.7$ V/nm. Inset: $\rho_{\mathrm{xx}}(T)$ at the compensation point ($n=0$) and $D=0.7$ V/nm. Dashed line: guide for the eye that represents the $a+bT^2$ dependence. **c,** $\rho_{\mathrm{xx}}(n)$ for BLG at $D=0$. **d,** Resistivity as a function of T for the charge-neutral SA-TBG at $D=0$ (blue) and $D=0.7$ V/nm (red) and for BLG at $D=0$ (grey). The data is normalized to the lowest-T value of $\rho_{\mathrm{xx}}(n)$: 4.2 K for SA-TBG and 10 K for BLG. **e,** $\Delta\rho=\rho_{\mathrm{xx}}(T)-\rho_{\mathrm{xx}}(4.2\ K)$ as a function of T measured at $D=0.7$ V/nm and $n=0$ (symbols). Note, $\Delta\rho(T)$ exhibits somewhat faster $T-$dependence at $T<15$ K. This apparent behavior is spurious and is related to the subtraction operation of the $\rho_0=\rho_{\mathrm{xx}}(4.2\ K)$ from the experimental dataset rather than ρ_{xx} at $T\to 0$. Solid line: theoretical dependence, eq. (1). Upper left inset: schematic illustration of the interlayer e-h friction in SA-TBG at finite D. Lower right inset: Prefactor $\mathcal{B}$ as a function of twist angle, θ.

with respect to that observed in the present experiment on SA-TBG. Indeed, at large θ, the layers are fully decoupled and $v_{\mathrm{F}} \approx v_{\mathrm{F}}^{\mathrm{g}}$. On the contrary, close to the magic angle, the layers are hybridized so that control of individual minivallyes via electrostatic means cannot be realized.

It would be instructive to put our observations in the context of electron transport in semimetals. Depending on quasiparticle statistics, band structure details, device geometry and interaction strength, seemingly alike semimetallic e-h systems can display very different physical properties and regimes of transport. For example, in charge-neutral MLG, frequent collisions between thermally activated electrons and holes impede electrical currents while leaving thermal ones untouched, causing a staggering breakdown of the Wiedeman-Franz law. In this system the Lorentz ratio, i.e. the ratio between the thermal conductivity and its electrical counterpart, is found to be greatly enhanced[8]. On the contrary, in degenerate compensated semimetals such as WP_2 or Sb the Lorentz ratio has been found to be suppressed[47]. Despite their semimetallic nature, which would imply violations of the Wiedeman-Franz law akin to those observed in graphene[48], the behavior of these materials closely resembles that of conventional unipolar systems[49], where carriers of a single type transport both charge and heat. All these seemingly contradictory observations have stimulated a debate over the effect of quasiparticle statistics, band structure and many-body interactions on the thermal and electrical properties of these charge-neutral material platforms[48,50,51]. A definitive resolution of these long lasting puzzles is made especially difficult by the fact that completely different behaviors are observed in different systems and regimes, and therefore a thorough comparison between them becomes challenging. The behavior of SA-TBG observed in this work thus makes it a highly-tunable platform for the exploration of different semimetallic regimes on an equal footing, allowing for a gradual transition between them.

To conclude, we have shown that SA-TBG offers a highly-tunable semimetalic system in which to explore physics at the crossover between the charge-neutral Dirac fluid and compensated e-h FL. In the latter case we found strong Coulomb friction between spatially separated electron and hole subsystems that resulted in the T^2- growth of the resistivity. Finally, we have developed a theory for e-h scattering in SA-TBG and found that its predictions are close to the experimental observations. It would be further interesting to explore transport and thermal properties of e-h FLs in other polarizable layered systems with heavier charge carriers such as twisted double bilayer graphene[3] or twisted transition metal dichalcogenides[6] as well as to explore collective modes in such e-h mixtures[52–54].

---

*Correspondence to: dab@nus.edu.sg, alepr85@gmail.com and pjarillo@mit.edu.

ACKNOWLEDGEMENTS

This work was supported by AFOSR grant FA9550-21-1-0319 and the Gordon and Betty Moore Foundation's EPiQS Initiative through Grant GBMF9463 to P.J.-H. D.A.B. acknowledges the support from MIT Pappalardo Fellowship. I.Y.P acknowledges support from the MIT undergraduate research opportunities program and the Johnson and Johnson research scholars program. K.W. and T.T. acknowledge support from JSPS KAKENHI (Grant Numbers 19H05790, 20H00354 and 21H05233). A.P. acknowledges support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS) and from the Leverhulme Trust under the grant RPG-2019-363. The authors thank Clement Collignon, Alexey Berdyugin and Dmitry Svintsov for productive discussions.

DATA AVAILABILITY

All data supporting this study and its findings are available within the article and its Supplementary Information or from the corresponding authors upon reasonable request.

AUTHOR CONTRIBUTIONS

D.A.B. and A.P. conceived and designed the study. D.A.B. and I.Y.P. fabricated and measured the devices. T.T. and K.W. grew high-quality hBN crystals. A.P. developed the theory. P.J.H. supervised the project.

COMPETING INTERESTS

The authors declare no competing interests.

[1] J. P. Eisenstein and A. H. MacDonald, Nature **432**, 691 (2004).

[2] A. Perali, D. Neilson, and A. R. Hamilton, Phys. Rev. Lett. **110**, 146803 (2013).

[3] P. Rickhaus, F. K. de Vries, J. Zhu, E. Portoles, G. Zheng, M. Masseroni, A. Kurzmann, T. Taniguchi, K. Watanabe, A. H. MacDonald, T. Ihn, and K. Ensslin, Science **373**, 1257 (2021).

[4] M. Zarenia, D. Neilson, and F. M. Peeters, Scientific Reports **7**, 11510 (2017).

[5] Y. Jia, P. Wang, C.-L. Chiu, Z. Song, G. Yu, B. Jäck, S. Lei, S. Klemenz, F. A. Cevallos, M. Onyszczak, N. Fishchenko, X. Liu, G. Farahi, F. Xie, Y. Xu, K. Watanabe, T. Taniguchi, B. A. Bernevig, R. J. Cava, L. M. Schoop, A. Yazdani, and S. Wu, "Evidence for a monolayer excitonic insulator," (2020), arXiv:2010.05390 [cond-mat.mes-hall].

[6] Z. Zhang, E. C. Regan, D. Wang, W. Zhao, S. Wang, M. Sayyad, K. Yumigeta, K. Watanabe, T. Taniguchi, S. Tongay, M. Crommie, A. Zettl, M. P. Zaletel, and F. Wang, "Correlated interlayer exciton insulator in double layers of monolayer wse2 and moiré ws2/wse2," (2021), arXiv:2108.07131 [cond-mat.mes-hall].

[7] S. Rakhmanov, Zhurnal Ehksperimental'noj i Teoreticheskoj Fiziki **75**, 160 (1978).

[8] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, Science **351**, 1058 (2016).

[9] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, and L. A. Ponomarenko, Nature Physics **8**, 896 (2012).

[10] J. I. A. Li, T. Taniguchi, K. Watanabe, J. Hone, A. Levchenko, and C. R. Dean, Phys. Rev. Lett. **117**, 046802 (2016).

[11] J. C. W. Song, D. A. Abanin, and L. S. Levitov, Nano Letters **13**, 3631 (2013).

[12] M. Titov, R. V. Gorbachev, B. N. Narozhny, T. Tudorovskiy, M. Schütt, P. M. Ostrovsky, I. V. Gornyi, A. D. Mirlin, M. I. Katsnelson, K. S. Novoselov, A. K. Geim, and L. A. Ponomarenko, Phys. Rev. Lett. **111**, 166601 (2013).

[13] M. Schütt, P. M. Ostrovsky, M. Titov, I. V. Gornyi, B. N. Narozhny, and A. D. Mirlin, Phys. Rev. Lett. **110**, 026601 (2013).

[14] J. C. W. Song and L. S. Levitov, Phys. Rev. Lett. **111**, 126601 (2013).

[15] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[16] M. Müller, J. Schmalian, and L. Fritz, Phys. Rev. Lett. **103**, 025301 (2009).

[17] P. Gallagher, C.-S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, Science **364**, 158 (2019).

[18] A. Block, A. Principi, N. C. H. Hesp, A. W. Cummings, M. Liebel, K. Watanabe, T. Taniguchi, S. Roche, F. H. L. Koppens, N. F. van Hulst, and K.-J. Tielrooij, "Observation of giant and tuneable thermal diffusivity of dirac fluid at room temperature," (2020), arXiv:2008.04189 [cond-mat.mes-hall].

[19] A. Lucas and K. C. Fong, Journal of Physics: Condensed Matter **30**, 053001 (2018).

[20] D. Svintsov, V. Vyurkov, S. Yurchenko, T. Otsuji, and V. Ryzhii, Journal of Applied Physics **111**, 083715 (2012).

[21] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

[22] D. Y. H. Ho, I. Yudhistira, N. Chakraborty, and S. Adam, Phys. Rev. B **97**, 121404 (2018).

[23] C. Cao, E. Elliott, J. Joseph, H. Wu, J. Petricka, T. Schäfer, and J. E. Thomas, Science **331**, 58 (2011).

[24] E. Shuryak, Progress in Particle and Nuclear Physics **53**, 273 (2004), heavy Ion Reaction from Nuclear to Quark Matter.

[25] Y. Nam, D.-K. Ki, D. Soler-Delgado, and A. F. Morpurgo, Nature Physics **13**, 1207 (2017).

[26] C. Tan, D. Y. H. Ho, L. Wang, J. I. A. Li, I. Yudhistira, D. A. Rhodes, T. Taniguchi, K. Watanabe, K. Shepard, P. L. McEuen, C. R. Dean, S. Adam, and J. Hone, Science Advances **8**, eabi8481 (2022).

[27] A. I. Berdyugin, N. Xin, H. Gao, S. Slizovskiy, Z. Dong, S. Bhattacharjee, P. Kumaravadivel, S. Xu, L. A. Ponomarenko, M. Holwill, D. A. Bandurin, M. Kim, Y. Cao, M. T. Greenaway, K. S. Novoselov, I. V. Grigorieva, K. Watanabe, T. Taniguchi, V. I. Fal'ko, L. S. Levitov, R. K. Kumar, and A. K. Geim, (2021), arXiv:2106.12609 [cond-mat.other].

[28] S. Fang, S. Carr, Z. Zhu, D. Massatt, and E. Kaxiras, "Angle-dependent *ab initio* low-energy hamiltonians for a relaxed twisted bilayer graphene heterostructure," (2019), arXiv:1908.00058 [cond-mat.mes-hall].

[29] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, Phys. Rev. Research **1**, 013001 (2019).

[30] R. Bistritzer and A. H. MacDonald, Proceedings of the National Academy of Science **108**, 12233 (2011).

[31] J. D. Sanchez-Yamagishi, T. Taychatanapat, K. Watanabe, T. Taniguchi, A. Yacoby, and P. Jarillo-Herrero, Phys. Rev. Lett. **108**, 076601 (2012).

[32] S. Slizovskiy, A. Garcia-Ruiz, A. I. Berdyugin, N. Xin, T. Taniguchi, K. Watanabe, A. K. Geim, N. D. Drummond, and V. I. Fal'ko, Nano Letters **21**, 6678 (2021).

[33] F. K. de Vries, J. Zhu, E. Portolés, G. Zheng, M. Masseroni, A. Kurzmann, T. Taniguchi, K. Watanabe, A. H. MacDonald, K. Ensslin, T. Ihn, and P. Rickhaus, Phys. Rev. Lett. **125**, 176801 (2020).

[34] A. I. Berdyugin, B. Tsim, P. Kumaravadivel, S. G. Xu, A. Ceferino, A. Knothe, R. K. Kumar, T. Taniguchi, K. Watanabe, A. K. Geim, I. V. Grigorieva, and V. I. Fal'ko, Science Advances **6**, eaay7838 (2020).

[35] I. Y. Phinney, D. A. Bandurin, C. Collignon, I. A. Dmitriev, T. Taniguchi, K. Watanabe, and P. Jarillo-Herrero, Phys. Rev. Lett. **127**, 056802 (2021).

[36] K. Kim, M. Yankowitz, B. Fallahazad, S. Kang, H. C. P. Movva, S. Huang, S. Larentis, C. M. Corbet, T. Taniguchi, K. Watanabe, S. K. Banerjee, B. J. LeRoy, and E. Tutuc, Nano Letters **16**, 1989 (2016).

[37] K. Kim, A. DaSilva, S. Huang, B. Fallahazad, S. Larentis, T. Taniguchi, K. Watanabe, B. J. LeRoy, A. H. MacDonald, and E. Tutuc, Proceedings of the National Academy of Sciences **114**, 3364 (2017).

[38] Y. Cao, J. Y. Luo, V. Fatemi, S. Fang, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Phys. Rev. Lett. **117**, 116804 (2016).

[39] D. G. Purdie, N. M. Pugno, T. Taniguchi, K. Watanabe, A. C. Ferrari, and A. Lombardo, Nature Communications **9**, 5387 (2018).

[40] Y. Kim, P. Herlinger, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, and J. H. Smet, Nano Letters **16**, 5053 (2016).

[41] A. A. Zibrov, E. M. Spanton, H. Zhou, C. Kometter, T. Taniguchi, K. Watanabe, and A. F. Young, Nature Physics **14**, 930 (2018).

[42] C. W. J. Beenakker and H. van Houten, Phys. Rev. Lett. **63**, 1857 (1989).

[43] A. S. Mayorov, R. V. Gorbachev, S. V. Morozov, L. Britnell, R. Jalil, L. A. Ponomarenko, P. Blake, K. S. Novoselov, K. Watanabe, T. Taniguchi, and A. K. Geim, Nano Letters **11**, 2396 (2011).

[44] G. Wagner, D. X. Nguyen, and S. H. Simon, Phys. Rev. Lett. **124**, 026601 (2020).

[45] Z. D. Kvon, E. B. Olshanetsky, D. A. Kozlov, E. Novik, N. N. Mikhailov, and S. A. Dvoretsky, Low Temperature Physics **37**, 202 (2011).

[46] V. L. Müller, Y. Yan, O. Kashuba, B. Trauzettel, M. Abdelghany, J. Kleinlein, W. Beugeling, H. Buhmann, and L. W. Molenkamp, Nano Letters **0**, null (0).

[47] J. Gooth, F. Menges, N. Kumar, V. Süβ, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nature Communications **9**, 4093 (2018).

[48] M. Zarenia, A. Principi, and G. Vignale, Phys. Rev. B **102**, 214304 (2020).

[49] A. Principi and G. Vignale, Phys. Rev. Lett. **115**, 056603 (2015).

[50] S. Li and D. L. Maslov, Phys. Rev. B **98**, 245134 (2018).

[51] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, V. Y. Kachorovskii, B. N. Narozhny, M. Schütt, and M. Titov, Phys. Rev. Lett. **114**, 156601 (2015).

[52] V. Andreeva, D. A. Bandurin, M. Luskin, and D. Margetis, Phys. Rev. B **102**, 205411 (2020).

[53] A. Principi, D. Bandurin, H. Rostami, and M. Polini, Phys. Rev. B **99**, 075410 (2019).

[54] Z. Sun, D. N. Basov, and M. M. Fogler, Proceedings of the National Academy of Sciences **115**, 3285 (2018).

[55] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, arXiv e-prints , arXiv:2008.12296 (2020), arXiv:2008.12296 [cond-mat.mes-hall].

[56] M. Ben Shalom, M. J. Zhu, V. I. Fal'ko, A. Mishchenko, A. V. Kretinin, K. S. Novoselov, C. R. Woods, K. Watanabe, T. Taniguchi, A. K. Geim, and J. R. Prance, Nature Physics **12**, 318 (2016).

[57] A. S. Mayorov, R. V. Gorbachev, S. V. Morozov, L. Britnell, R. Jalil, L. A. Ponomarenko, P. Blake, K. S. Novoselov, K. Watanabe, T. Taniguchi, and A. K. Geim, Nano Letters **11**, 2396 (2011).

SUPPLEMENTARY INFORMATION

S1. Device fabrication.

Our device consisted of hBN-encapsulated twisted bilayer graphene, which we fabricated using a combination of cut-and-stack[37,55] and hot release[39] methods. Monolayer graphene, few-layer graphite, and 30-80 nm-thick hBN crystals were mechanically exfoliated on a Si/SiO_2 substrate, and sizable, uniform flakes were selected using optical microscopy. Then, using a homemade transfer system with μm-accuracy and a polycarbonate (PC) membrane stretched over a small ($8\,\mathrm{mm}\times 8\,\mathrm{mm}\times 4\,\mathrm{mm}$) polydimethylsiloxane (PDMS) polymer block on a glass slide, we assembled hBN and graphite stacks on a Si/SiO_2 wafer. We first picked up hBN crystal at 50-70 °C. Then, the graphite was picked up at room temperature, and then the entire stack was "ironed" and then released on a clean Si/SiO_2 wafer at high temperatures ($160 - 170\,^\circ$C). After removing the polymer membrane, we annealed the hBN and graphite stack at 350 °C for 3 hours in argon/hydrogen atmosphere. We then assembled the hBN and twisted bilayer graphene stack using a "cut-and-stack" method described previously[37,38]. After picking up the top hBN and twisted graphene, we "ironed" the entire stack at room temperature. The three-layer stack was then released onto the previously fabricated and cleaned bottom hBN and graphite gate at roughly 160 C. After this point, we avoided heating the stack to reduce the possibility of twist angle relaxation. The final stack was inspected using dark-field microscopy and atomic force microscopy (AFM), and bubble- and blister-free areas were selected to use for Hall bars.

To fabricate the devices, we covered the heterostructures by a protective polymethyl-methacrylate (PMMA) resist and used electron beam lithography (EBL) to define contact regions. We then performed a mild O_2 plasma cleaning before using reactive ion etching (RIE) with a plasma generated from CHF_3 and O_2 gases to selectively etch away the hBN in the parts of the heterostructure unprotected by the lithographic mask[56]. 3 nm chromium and 50-70 nm gold was then evaporated into the contact regions via thermal evaporation at high vacuum. We repeat the same EBL and thermal evaporation procedures to define a metallic top gate (3 nm chromium and 30-40 nm gold). Finally, we repeat the same EBL and RIE procedures to define the final Hall bar geometry, using, in this case, a plasma generated by Ar, O_2 and CHF_3 gases.

S2. Signatures of ballistic transport in SA-TBG.

One of the standard ways to probe the presence of ballistic transport in mesoscopic devices is to measure the transfer resistance in the bend geometry illustrated in Fig. S1[42,57]. In this geometry, an electrical current, I_{1-2}, is passed between contacts 1 and 2 and the voltage drop, I_{3-4}, is measured between contacts 3 and 4. In the case of diffusive transport, this configuration yields a positive signal, $R_{3-4,1-2} = V_{3-4}/I_{1-2}$ because charge carriers flow along electric field lines. The bend geometry is then topologically identical to the conventional 4-point configuration used for resistivity measurements. On the contrary, if the charge carriers experience ballistic motion on a scale of the device width then nothing prevents them from reaching the opposite device boundary producing negative $R_{3-4,1-2}$. The negative sign of the transfer resistance measured in the bend geometry is usually considered as a benchmark of ballistic transport regime.

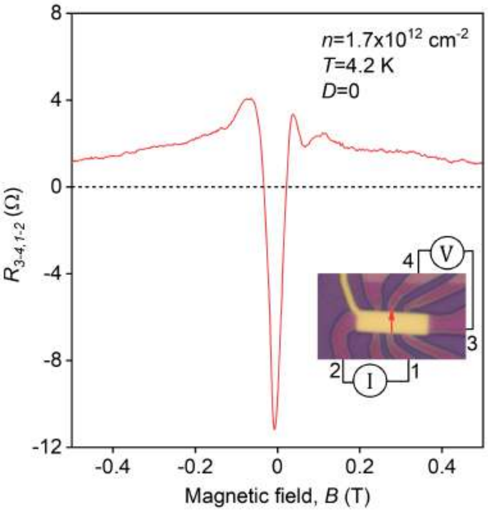


FIG. S1. **Micrometer-scale ballistic transport in encapsulated SA-TBG.** $R_{3-4,1-2}$ as a function of magnetic field measured in the geometry, shown in the inset, for given n, T and D.

We have performed such measurements in our sample and found that at liquid helium T, $R_{3-5,1-2}$ is negative over the wide range of n. Furthermore, the application of a magnetic field, causes the sign change of the measured signal, as the charge carriers are deflected from the straight trajectories. These observations highlight a high quality of our encapsulated sample, that is crucial for studies of interaction-dominated transport described in the main text.

S3. Theoretical calculations of the interaction-dominated resistivity in SA-TBG.

To derive the resistivity due to electron-electron interactions, we start from the Boltzmann equation for the space- and time-dependent fermion occupation function $f_{\boldsymbol{k},\lambda}(\boldsymbol{r},t)$,

$$\partial_t f_{\boldsymbol{k},\lambda}(\boldsymbol{r},t) + \boldsymbol{v}_{\boldsymbol{k},\lambda}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}} f_{\boldsymbol{k},\lambda}(\boldsymbol{r},t) - e\boldsymbol{E}\cdot\boldsymbol{\nabla}_{\boldsymbol{k}} f_{\boldsymbol{k},\lambda}(\boldsymbol{r},t) = \mathcal{I}_{\rm ee}[f_{\boldsymbol{k},\lambda}] \ , \tag{S2}$$

where $\boldsymbol{v}_{\boldsymbol{k},\lambda} = \boldsymbol{\nabla}_{\boldsymbol{k}}\varepsilon_{\boldsymbol{k},\lambda}$ is the velocity of a particle of Bloch wavevector $\boldsymbol{k}$ characterized by (band, spin, valley and layer) quantum numbers λ, $-e$ is the electron charge, and $\boldsymbol{E}$ is the (uniform and time-independent) external electric field. Finally, $\mathcal{I}_{\rm ee}[f_{\boldsymbol{k},\lambda}]$ is the collision integral of electron-electron interactions:

$$\begin{aligned}\mathcal{I}_{\rm ee}[f_{\boldsymbol{k},\lambda}] = {}& \frac{1}{\mathcal{A}^3}\sum_{\boldsymbol{k}_2,\lambda_2}\sum_{\substack{\boldsymbol{k}_3,\lambda_3\\ \boldsymbol{k}_4,\lambda_4}} W_{\rm ee}(\boldsymbol{k}_1,\lambda_1;\boldsymbol{k}_2,\lambda_2;\boldsymbol{k}_3,\lambda_3;\boldsymbol{k}_4,\lambda_4)\delta(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{k}_3-\boldsymbol{k}_4)\delta(\varepsilon_{\boldsymbol{k}_1,\lambda_1}+\varepsilon_{\boldsymbol{k}_2,\lambda_2}-\varepsilon_{\boldsymbol{k}_3,\lambda_3}-\varepsilon_{\boldsymbol{k}_4,\lambda_4})\\ &\times\left[f_{\boldsymbol{k}_1,\lambda_1}f_{\boldsymbol{k}_2,\lambda_2}(1-f_{\boldsymbol{k}_3,\lambda_3})(1-f_{\boldsymbol{k}_4,\lambda_4}) - (1-f_{\boldsymbol{k}_1,\lambda_1})(1-f_{\boldsymbol{k}_2,\lambda_2})f_{\boldsymbol{k}_3,\lambda_3}f_{\boldsymbol{k}_4,\lambda_4}\right] \ .\end{aligned} \tag{S3}$$

Within the Fermi-golden rule,

$$W_{\rm ee}(\boldsymbol{k}_1,\lambda_1;\boldsymbol{k}_2,\lambda_2;\boldsymbol{k}_3,\lambda_3;\boldsymbol{k}_4,\lambda_4) = \frac{2\pi}{\hbar}|V_{\rm ee}(\boldsymbol{k}_1-\boldsymbol{k}_3,\varepsilon_{\boldsymbol{k}_1,\lambda_1}-\varepsilon_{\boldsymbol{k}_3,\lambda_3})|^2\mathcal{D}(\boldsymbol{k}_1,\lambda_1;\boldsymbol{k}_3,\lambda_3)\mathcal{D}(\boldsymbol{k}_2,\lambda_2;\boldsymbol{k}_4,\lambda_4) \ , \tag{S4}$$

where $V(\boldsymbol{q},\omega) = v_{\boldsymbol{q}}/\epsilon(\boldsymbol{q},\omega)$, $v_{\boldsymbol{q}} = 2\pi e^2/q$ is the bare Coulomb interaction and $\epsilon(\boldsymbol{q},\omega) = 1 - v_{\boldsymbol{q}}\chi_{nn}(\boldsymbol{q},\omega)$ is the dielectric function. In Eq. (S4), we have defined the overlap between initial and final states as $\mathcal{D}(\boldsymbol{k},\lambda;\boldsymbol{k}',\lambda') = \left|\langle\boldsymbol{k},\lambda|\boldsymbol{k}',\lambda'\rangle\right|^2$, where $|\boldsymbol{k},\lambda\rangle$ is an eigenstate of the bare Hamiltonian. For future purposes we define $n_{\rm F}(x) = [\exp(x)+1]^{-1}$ and $n_{\rm B}(x) = [\exp(x)-1]^{-1}$ as the Fermi and Bose distribution, respectively.

We solve Eq. (S2) in the steady state and to linear order in the electric field by employing the following Ansatz:

$$f_{\boldsymbol{k},\lambda} = f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda}) - e\tau\frac{\boldsymbol{E}\cdot\boldsymbol{v}_{\boldsymbol{k},\lambda}}{k_{\rm B}T}f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})\left[1-f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})\right] \ , \tag{S5}$$

where $f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})$ is the equilibrium Fermi-Dirac distribution function at the temperature T. Plugging the Ansatz (S5) into Eq. (S2) and linearizing with respect $\boldsymbol{E}$ to we get

$$\begin{aligned}\left(-\frac{\partial f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})}{\partial\varepsilon_{\boldsymbol{k},\lambda}}\right)\boldsymbol{v}_{\boldsymbol{k},\lambda}\cdot\boldsymbol{E} = {}& -\frac{2\pi\tau}{\hbar\mathcal{A}^3k_{\rm B}T}\sum_{\boldsymbol{k}_2,\lambda_2}\sum_{\substack{\boldsymbol{k}_3,\lambda_3\\ \boldsymbol{k}_4,\lambda_4}}|V_{\rm ee}(\boldsymbol{k}_1-\boldsymbol{k}_3,\varepsilon_{\boldsymbol{k}_1,\lambda_1}-\varepsilon_{\boldsymbol{k}_3,\lambda_3})|^2\mathcal{D}(\boldsymbol{k}_1,\lambda_1;\boldsymbol{k}_3,\lambda_3)\mathcal{D}(\boldsymbol{k}_2,\lambda_2;\boldsymbol{k}_4,\lambda_4)\\ &\times\delta(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{k}_3-\boldsymbol{k}_4)\delta(\varepsilon_{\boldsymbol{k}_1,\lambda_1}+\varepsilon_{\boldsymbol{k}_2,\lambda_2}-\varepsilon_{\boldsymbol{k}_3,\lambda_3}-\varepsilon_{\boldsymbol{k}_4,\lambda_4})\\ &\times f^{(0)}(\varepsilon_{\boldsymbol{k}_1,\lambda_1})f^{(0)}(\varepsilon_{\boldsymbol{k}_2,\lambda_2})\left[1-f^{(0)}(\varepsilon_{\boldsymbol{k}_3,\lambda_3})\right]\left[1-f^{(0)}(\varepsilon_{\boldsymbol{k}_4,\lambda_4})\right]\\ &\times(\boldsymbol{v}_{\boldsymbol{k}_1,\lambda_1}+\boldsymbol{v}_{\boldsymbol{k}_2,\lambda_2}-\boldsymbol{v}_{\boldsymbol{k}_3,\lambda_3}-\boldsymbol{v}_{\boldsymbol{k}_4,\lambda_4})\cdot\boldsymbol{E} \ .\end{aligned} \tag{S6}$$

To determine the transport time τ, we multiply Eq. (S6) by $\boldsymbol{v}_{\boldsymbol{k},\lambda}$ and sum over all $\boldsymbol{k}$ and λ. We obtain $\tau = DI^{-1}$, where the Drude weight is

$$D = \frac{1}{2\mathcal{A}}\sum_{\boldsymbol{k},\lambda}\left(-\frac{\partial f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})}{\partial\varepsilon_{\boldsymbol{k},\lambda}}\right)|\boldsymbol{v}_{\boldsymbol{k},\lambda}|^2 \ , \tag{S7}$$

while

$$\begin{aligned}I = {}& -\frac{\pi}{16\hbar k_{\rm B}T\mathcal{A}^3}\sum_{\boldsymbol{q}}\int_{-\infty}^{\infty}d\omega\frac{|V_{\rm ee}(\boldsymbol{q},\omega)|^2}{\sinh^2\left(\dfrac{\omega}{2k_{\rm B}T}\right)}\sum_{\boldsymbol{k},\boldsymbol{k}'}\sum_{\lambda,\lambda'}\sum_{\eta,\eta'}\mathcal{D}(\boldsymbol{k},\lambda;\boldsymbol{k}-\boldsymbol{q},\lambda')\mathcal{D}(\boldsymbol{k}',\eta;\boldsymbol{k}'+\boldsymbol{q},\eta')\\ &\times\left[f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})-f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda}-\omega)\right]\left[f^{(0)}(\varepsilon_{\boldsymbol{k}',\eta})-f^{(0)}(\varepsilon_{\boldsymbol{k}',\eta}+\omega)\right]\\ &\times|\boldsymbol{v}_{\boldsymbol{k},\lambda}+\boldsymbol{v}_{\boldsymbol{k}',\eta}-\boldsymbol{v}_{\boldsymbol{k}-\boldsymbol{q},\lambda'}-\boldsymbol{v}_{\boldsymbol{k}'+\boldsymbol{q},\eta'}|^2\delta(\varepsilon_{\boldsymbol{k},\lambda}-\varepsilon_{\boldsymbol{k}-\boldsymbol{q},\lambda'}-\omega)\delta(\varepsilon_{\boldsymbol{k}',\eta}-\varepsilon_{\boldsymbol{k}'+\boldsymbol{q},\eta'}+\omega) \ .\end{aligned} \tag{S8}$$

The resistivity is therefore

$$\rho = \frac{I}{e^2D^2} \ . \tag{S9}$$

We now specialize this result to the case of twisted bilayer graphene. The two layers are kept at different potentials. Therefore, the two Dirac crossings at the points K and K' of the mini-Brillouin zone are shifted in energy in opposite directions. As such, one of the valleys is hole-doped, while the other is electron doped. When the temperature T is much smaller than the Fermi energy $\varepsilon_{\rm F}$ (which we assume to be equal in modulus for the two valleys), we find

$$D = \frac{N_{\rm f} v_{\rm F}^2}{2} \nu(\varepsilon_{\rm F}) \ , \tag{S10}$$

where $\nu(\varepsilon) = \varepsilon/(2\pi\hbar^2 v_{\rm F}^2)$ is the density of states of a single Dirac cone, and $N_{\rm f} = 8$ is the total number of (spin, valley and layer) fermion flavors. As Eq. (S8), in the low-temperature limit the function $1/\sinh^2[\omega/(2k_{\rm B}T)]$ strongly suppresses contributions at large ω. Expanding the integrand in the limit of $\omega \to 0$ to the leading order, we get

$$\begin{aligned} I \simeq \frac{\pi}{16\hbar k_{\rm B}T} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \int_{-\infty}^{\infty} d\omega \frac{\omega^2 |V_{\rm ee}(\boldsymbol{q},0)|^2}{\sinh^2\left(\dfrac{\omega}{2k_{\rm B}T}\right)} \int \frac{d^2\boldsymbol{k}}{(2\pi)^2} \int \frac{d^2\boldsymbol{k}'}{(2\pi)^2} \sum_{\lambda,\lambda'} \sum_{\eta,\eta'} \mathcal{D}(\boldsymbol{k},\lambda;\boldsymbol{k}-\boldsymbol{q},\lambda')\mathcal{D}(\boldsymbol{k}',\eta;\boldsymbol{k}'+\boldsymbol{q},\eta') \\ \times\, \delta(\varepsilon_{\boldsymbol{k},\lambda} - \varepsilon_{\boldsymbol{k}-\boldsymbol{q},\lambda'})\delta(\varepsilon_{\boldsymbol{k}',\eta} - \varepsilon_{\boldsymbol{k}'+\boldsymbol{q},\eta'}) \left(-\frac{\partial f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})}{\partial \varepsilon_{\boldsymbol{k},\lambda}}\right) \left(-\frac{\partial f^{(0)}(\varepsilon_{\boldsymbol{k}',\eta})}{\partial \varepsilon_{\boldsymbol{k}',\eta}}\right) |\boldsymbol{v}_{\boldsymbol{k},\lambda} + \boldsymbol{v}_{\boldsymbol{k}',\eta} - \boldsymbol{v}_{\boldsymbol{k}-\boldsymbol{q},\lambda'} - \boldsymbol{v}_{\boldsymbol{k}'+\boldsymbol{q},\eta'}|^2 \ . \end{aligned} \tag{S11}$$

We now observe that, in the limit of low temperature,

$$\left(-\frac{\partial f^{(0)}(\varepsilon_{\boldsymbol{k},\lambda})}{\partial \varepsilon_{\boldsymbol{k},\lambda}}\right) \to \delta(\varepsilon_{\boldsymbol{k},\lambda} - \varepsilon_{\rm F}) \ . \tag{S12}$$

Because of the four δ-functions in the integrand of Eq. (S11), all initial and final states are bound to the Fermi surface. Hence, the matrix element on its last line can be written as

$$\mathcal{M} \equiv |\boldsymbol{v}_{\boldsymbol{k},\lambda} + \boldsymbol{v}_{\boldsymbol{k}',\eta} - \boldsymbol{v}_{\boldsymbol{k}-\boldsymbol{q},\lambda'} - \boldsymbol{v}_{\boldsymbol{k}'+\boldsymbol{q},\eta'}|^2 \simeq \frac{v_{\rm F}^2 q^2}{k_{\rm F}^2} |s_\lambda - s_\eta|^2 \ . \tag{S13}$$

Here, $s_\lambda = \pm$ accounts for the direction of velocity with respect to momentum in a given mini-valley of the Brillouin zone. To obtain this result, we have neglected the possibility of inter-minivalley transitions. Hence, the particles labelled by λ and λ' (η and η') belong to the same mini-valley in the Brillouin zone. This means that they share the same band, valley, and layer index.

From Eq. (S13), it is clear that if all minivalleys are populated with the same type of carriers, the collision integral of electron-electron interactions vanishes. Since, in the present case, populations are unequal, *i.e.* $s_\lambda \neq s_\eta$ for some choices of λ and η. To be specific, for each of the $N_{\rm f}$ choices of λ, there are $N_{\rm f}/2$ possible choices for η such that $s_\lambda \neq s_\eta$. In these cases, $\mathcal{M} = 4v_{\rm F}^2 q^2/k_{\rm F}^2$. Therefore,

$$I \simeq \frac{\pi N_{\rm f}^2}{8\hbar k_{\rm B}T} \frac{v_{\rm F}^2}{k_{\rm F}^2} \int \frac{d^2\boldsymbol{q}}{(2\pi)^2} \int_{-\infty}^{\infty} d\omega \frac{\omega^2 |V_{\rm ee}(\boldsymbol{q},0)|^2 q^2}{\sinh^2\left(\dfrac{\omega}{2k_{\rm B}T}\right)} \Gamma^2(q) \ , \tag{S14}$$

where

$$\begin{aligned} \Gamma(q) &= \int \frac{d^2\boldsymbol{k}}{(2\pi)^2} \frac{1 + \cos(\varphi_{\boldsymbol{k}} - \varphi_{\boldsymbol{k}+\boldsymbol{q}})}{2} \delta(\varepsilon_{\boldsymbol{k}} - \varepsilon_{\rm F})\delta(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}} - \varepsilon_{\rm F}) \\ &\simeq \Theta(2k_{\rm F} - q)\nu(\varepsilon_{\rm F}) \frac{1}{\pi\hbar v_{\rm F} q} \sqrt{1 - \frac{q^2}{4k_{\rm F}^2}} \ . \end{aligned} \tag{S15}$$

After some lengthy algebra we find

$$I \simeq \frac{2\pi^2 N_{\rm f}^2}{3} \left(\frac{k_{\rm B}T}{\varepsilon_{\rm F}}\right)^2 \hbar v_{\rm F}^4 \nu^2(\varepsilon_{\rm F}) \alpha_{\rm ee}^2 \frac{3(N_{\rm F}\alpha_{\rm ee} - 1) + \left[4 - 3(N_{\rm F}\alpha_{\rm ee})^2\right] \operatorname{arccoth}(1 + N_{\rm F}\alpha_{\rm ee})}{2}, \tag{S16}$$

and therefore the scattering rate and resistivity read

$$\frac{1}{\tau} = \frac{\varepsilon_{\rm F}}{\hbar} \frac{2\pi N_{\rm f}}{3} \left(\frac{k_{\rm B}T}{\varepsilon_{\rm F}}\right)^2 \alpha_{\rm ee}^2 \frac{3(N_{\rm F}\alpha_{\rm ee} - 1) + \left[4 - 3(N_{\rm F}\alpha_{\rm ee})^2\right] \operatorname{arccoth}(1 + N_{\rm F}\alpha_{\rm ee})}{2}, \tag{S17}$$

and

$$\rho_{\mathrm{el}} = \frac{h}{e^2}\frac{4\pi}{3}\left(\frac{k_{\mathrm{B}}T}{\varepsilon_{\mathrm{F}}}\right)^2 \alpha_{\mathrm{ee}}^2 \frac{3(N_{\mathrm{F}}\alpha_{\mathrm{ee}}-1)+\left[4-3(N_{\mathrm{F}}\alpha_{\mathrm{ee}})^2\right]\mathrm{arccoth}(1+N_{\mathrm{F}}\alpha_{\mathrm{ee}})}{2}, \tag{S18}$$

respectively.

Giant chirality-induced spin polarization in twisted transition metal dichalcogenides

Guido Menichetti,[1, *] Lorenzo Cavicchi,[2] Leonardo Lucchesi,[1, 3] Fabio Taddei,[4, 2] Giuseppe Iannaccone,[3] Pablo Jarillo-Herrero,[5] Claudia Felser,[6] Frank H. L. Koppens,[7, 8] and Marco Polini[1, 7]

[1]*Dipartimento di Fisica dell'Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy*
[2]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[3]*Dipartimento di Ingegneria dell'Informazione dell'Università di Pisa, Via Girolamo Caruso 16, I-56122 Pisa, Italy*
[4]*NEST, Istituto Nanoscienze-CNR, Piazza S. Silvestro 12, I-56126 Pisa, Italy*
[5]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[6]*Max Planck Institute for Chemical Physics of Solids, Nöthnitzer Str. 40, Dresden 01187, Germany*
[7]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology,*
Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain
[8]*ICREA-Institució Catalana de Recerca i Estudis Avançats,*
Passeig de Lluís Companys 23, 08010 Barcelona, Spain

(Dated: January 24, 2024)

Chirality-induced spin selectivity (CISS) is an effect that has recently attracted a great deal of attention in chiral chemistry and that remains to be understood. In the CISS effect, electrons passing through chiral molecules acquire a large degree of spin polarization. In this work we study the case of atomically-thin chiral crystals created by van der Waals assembly. We show that this effect can be spectacularly large in systems containing just two monolayers, provided they are spin-orbit coupled. Its origin stems from the combined effects of structural chirality and *spin-flipping* spin-orbit coupling. We present detailed calculations for twisted homobilayer transition metal dichalcogenides, showing that the chirality-induced spin polarization can be giant, e.g. easily exceeding 50% for $MoTe_2$. Our results clearly indicate that twisted quantum materials can operate as a fully tunable platform for the study and control of the CISS effect in condensed matter physics and chiral chemistry.

Introduction.—In 1999, a team led by R. Naaman recognized that films made from chiral organic molecules scatter polarized electrons asymmetrically [1]. Spin polarization induced by chirality is an effect now universally known as chirality-induced spin selectivity (CISS) effect [2–6]. Almost ten years later, Rosenberg et al. [7] demonstrated that a polarized spin can induce an enantiospecific chemical process. It is now widely accepted [2–6] that the CISS effect has profound implications on asymmetric chemistry. Chiral molecules and, more in general, chiral materials [8] and bulk chiral crystals [9] can act as spin polarizers and filters and therefore influence processes associated with electron transfer, electron transport, and bond polarization through chiral structures.

To the best of our knowledge, though, a microscopic understanding of the molecular CISS effect is still lacking (for a recent summary we refer the reader to Ref. [10]). In particular, the role of spin-orbit coupling [10], geometric phases [11], electron correlations [12], electron-phonon coupling [13], dissipation [14], and the interactions between chiral molecules and/or chiral and magnetic substrates are not understood. This confusion on its microscopic origin limits the use of CISS in applications and makes it challenging to enhance it.

The point we want to raise in this work is that a thorough microscopic understanding of the CISS effect calls for a fully tunable platform where the dependence of the effect on a multitude of control parameters can be studied, both experimentally and theoretically. As we will demonstrate below, we believe that moiré superlattices obtained by twisting atomically-thin crystals with respect to each other [15–17] represent such a platform. Despite the great deal of interest that these systems have attracted since the discovery of superconductivity [18] and correlated insulating states [19] in twisted bilayer graphene (TBG) [20–28], very few studies have highlighted the fact that these systems are naturally chiral materials akin to "giant chiral molecules". The only qualitative difference between chiral molecules and twisted

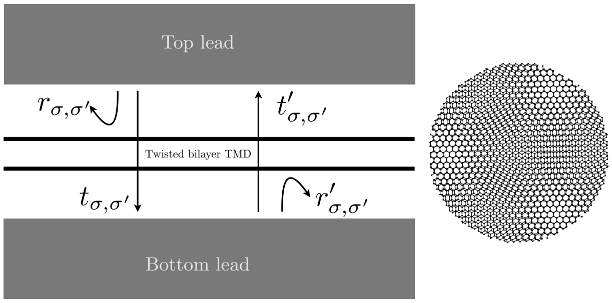


FIG. 1. Sketch of the two-terminal setup we have studied in this work. A twisted homobilayer TMD, developing along the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{y}}$ plane, is contacted by two semi-infinite leads (denoted by grey-shaded areas). The spatial separation between the two TMD layers in the vertical $\hat{\boldsymbol{z}}$ direction is d. The twist between the two layers is achieved by a counter-clockwise rotation of the top layer with respect to the bottom one by an angle θ. The resulting moiré lattice is plotted on the right. Black arrows show the spin-resolved reflection, $r_{\sigma,\sigma'}$, $r'_{\sigma,\sigma'}$, and transmission, $t_{\sigma,\sigma'}$, $t'_{\sigma,\sigma'}$, components of the scattering matrix S reported in Eq. (1).

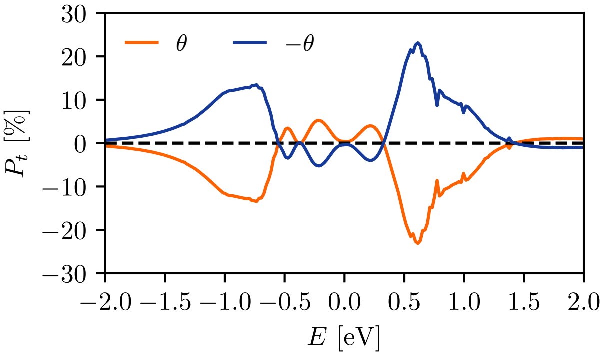


FIG. 2. (Color online) Numerical results for the energy dependence of the spin polarization $P_t = P_t(E)$, as defined in Eq. (5), of electrons transmitted from the top to the bottom lead. The calculated spin polarization has opposite sign for θ (orange curve) and $-\theta$ (blue curve). In the non-chiral (untwisted, $\theta = 0$) case, the spin polarization P_t vanishes (black dashed line). Results in this plot refer to parameters for twisted homobilayer MoTe$_2$ [49], $\theta = 6.01^\circ$, $\lambda_0 = 220$ meV, and $\lambda_{\rm BR} = 0$.

quasi-two-dimensional (2D) materials is that the latter have (a nearly exact) Bloch translational invariance.

The fact that twisted materials are chiral is not a mere mathematical statement but has experimental implications e.g. on the interaction between these materials and light. For example, natural optical activity [29], a hallmark of chiral materials, was experimentally discovered [30] in TBG at large twist angles well before exotic states of matter were found at the magic angle [18, 19]. TBG exhibits remarkably large circular dichroism [30, 31], up to a factor 100 stronger than for a layer of chiral molecules of similar thickness.

In this work, we demonstrate that the interplay between spin-orbit coupling (SOC) and chirality in twisted quantum materials gives rise to a giant CISS effect, *even for just two twisted monolayers*. More precisely, we consider a family of twisted moiré superlattices which are known to display strong SOC. These are twisted homobilayer transition metal dichalcogenides (TMDs) [32] such as twisted MoTe$_2$, which is currently attracting a great deal of attention because of the experimental discovery [33–36] of fractional Chern insulating states in zero magnetic field. These systems offer a wide range of tunable parameters, such as carrier density (which can be changed via e.g. the electrical field effect), twist angle (which can be changed at will, thereby inducing dramatic changes in the Bloch bands of these crystals), and more conventional ones, such as temperature and applied magnetic fields. Our interest here is on the study of the interplay between chiral orbital motion, spin-orbit coupling, and the spin degree of freedom. Quantifying the role of electron-electron interactions (e.g. exchange interactions [37]), which are strong in moiré materials, and quantum geometry [38–40] is well beyond the scope of the present work and is left for future work.

Setup and scattering matrix approach to the CISS effect.—In order to numerically extract information about the CISS effect in twisted homobilayer TMDs, we utilize the setup depicted in Fig. 1. It consists of a twisted homobilayer TMD contacted by two leads. One can inject carriers from the top lead, for example, which, after tunneling through the strongly spin-orbit coupled twisted homobilayer TMD, will be extracted from the bottom layer. As we will discuss momentarily, the reciprocity theorem [41] forbids the observation of a spin-polarized current through a standard linear-response magnetoresistance measurement in a two-terminal setup [42]. Nevertheless, the spin-polarized nature of electron scattering in our chiral van der Waals heterostructure can be detected through a spin-resolved scattering matrix approach [43–46].

The top and bottom leads in Fig. 1 offer asymptotic propagating states. We introduce: i) $\psi^{\rm out}_{{\rm T},\sigma}$ ($\psi^{\rm out}_{{\rm B},\sigma}$) as the asymptotic wave-function for electrons with spin $\sigma =\uparrow, \downarrow$, which are scattered out of the twisted homobilayer TMD into the top (bottom) lead; ii) $\psi^{\rm in}_{{\rm T},\sigma'}$ ($\psi^{\rm in}_{{\rm B},\sigma'}$) as the asymptotic wave-function for electrons with spin $\sigma' =\uparrow, \downarrow$, which are injected from the top (bottom) lead into the twisted homobilayer TMD. The latter can be viewed as a scatterer, whose transport properties are completely defined by its spin-resolved scattering matrix $S_{\sigma,\sigma'}$ [43]:

$$\begin{pmatrix} \psi^{\rm out}_{{\rm T},\sigma} \\ \psi^{\rm out}_{{\rm B},\sigma} \end{pmatrix} = \sum_{\sigma'=\uparrow,\downarrow} S_{\sigma,\sigma'} \begin{pmatrix} \psi^{\rm in}_{{\rm T},\sigma'} \\ \psi^{\rm in}_{{\rm B},\sigma'} \end{pmatrix} , \tag{1}$$

where

$$S_{\sigma,\sigma'} \equiv \begin{pmatrix} r_{\sigma,\sigma'} & t'_{\sigma,\sigma'} \\ t_{\sigma,\sigma'} & r'_{\sigma,\sigma'} \end{pmatrix} . \tag{2}$$

Here, $t_{\sigma,\sigma'}, r_{\sigma,\sigma'}$ ($t'_{\sigma,\sigma'}, r'_{\sigma,\sigma'}$) are the spin-resolved transmission and reflection amplitudes of electrons coming from the top (bottom) lead—see Fig. 1. All the quantities in Eqs. (1)-(2)—and below in Eqs. (3)-(5)—are functions of a single energy E, since we are assuming only elastic scattering mechanisms.

Starting from the entries of the scattering matrix, we introduce the following reflection $\rho_{\mathbb{r}}, \rho_{\mathbb{r}'}$ and transmission $\rho_{\mathbb{t}}, \rho_{\mathbb{t}'}$ probability matrices:

$$\begin{aligned} \rho_{\mathbb{r}} \equiv \mathbb{r}\mathbb{r}^\dagger \ , \quad \rho_{\mathbb{r}'} \equiv \mathbb{r}'(\mathbb{r}')^\dagger \\ \rho_{\mathbb{t}} \equiv \mathbb{t}\mathbb{t}^\dagger \ , \quad \rho_{\mathbb{t}'} \equiv \mathbb{t}'(\mathbb{t}')^\dagger \ . \end{aligned} \tag{3}$$

Here, $\mathbb{r}$, $\mathbb{r}'$, $\mathbb{t}$, and $\mathbb{t}'$ are matrices in spin space with matrix elements $r_{\sigma,\sigma'}$, $r'_{\sigma,\sigma'}$, $t_{\sigma,\sigma'}$, $t'_{\sigma,\sigma'}$, respectively.

We also define the following dimensionless spin conductances [44] for each scattering process:

$$\begin{aligned} \sigma_r \equiv {\rm Tr}[\hat{s}_z \rho_{\mathbb{r}}] \ , \quad \sigma_{r'} \equiv {\rm Tr}[\hat{s}_z \rho_{\mathbb{r}'}] \\ \sigma_t \equiv {\rm Tr}[\hat{s}_z \rho_{\mathbb{t}}] \ , \quad \sigma_{t'} \equiv {\rm Tr}[\hat{s}_z \rho_{\mathbb{t}'}] \ . \end{aligned} \tag{4}$$

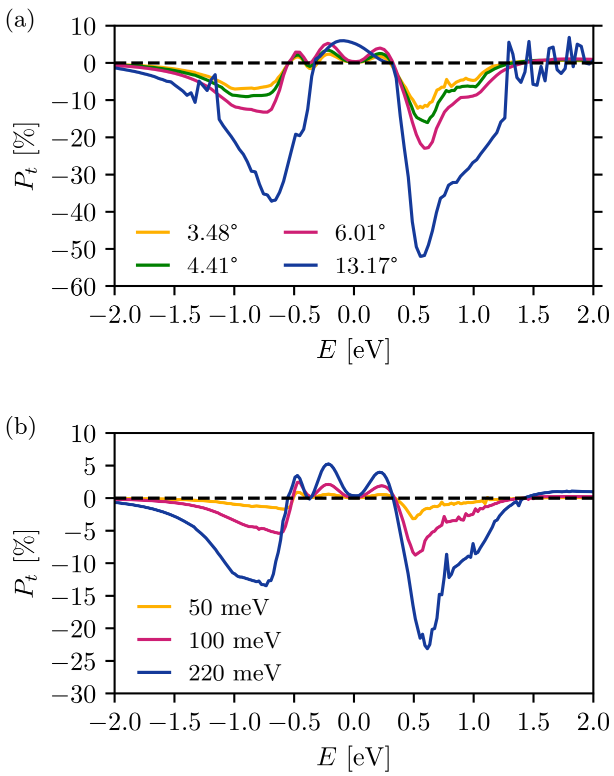


FIG. 3. (Color online) Panel (a) Results for the spin polarization $P_t(E)$ for different values of the twist angle θ ($\theta = 3.48°, 4.41°, 6.01°$, and $13.17°$). The black dashed line refers to the non-chiral $\theta = 0$ case. Panel (b) Results for the spin polarization $P_t(E)$ for different values of λ_0 ($\lambda_0 = 50$ meV, 100 meV, and 220 meV). In this panel, the angle has been fixed at $\theta = 6.01°$. The black dashed line refers to the case in which SOC is artificially turned off by setting $\lambda_0 = 0$. All the other parameters are the same as in Fig. 2.

where $\hat{s}_z$ is a Pauli matrix. The trace operation in Eq. (4) is of course intended over the spin degrees of freedom.

Once again, as demonstrated in great detail in Ref. [42], the reciprocity theorem [41] allows one to conclude that any linear-response electrical measurement in a two-terminal setup, even in the presence of auxiliary ferromagnets, is *insensitive* to the "spin conductances" introduced in Eq. (4). The setup depicted in Fig. 1 therefore should not be intended as a real experimental setup to infer information on the quantities in Eq. (4), but, rather, as a computational setup with asymptotic propagating states to numerically access the full scattering matrix.

The four quantities in Eq. (4) are not independent. Because of charge conservation (i.e. because of the unitarity of the full scattering matrix), $\sigma_r = -\sigma_{t'}$ and $\sigma_{r'} = -\sigma_t$. Therefore, in general, we have two independent spin conductances, e.g. σ_t and $\sigma_{t'}$. From now on, we focus only on σ_t for reasons that will become clear momentarily. In particular, we introduce the energy-dependent *chirality-induced spin polarization* $P_t(E)$ by properly normalizing the corresponding spin conductance:

$$P_t(E) \equiv \frac{\sigma_t(E)}{\mathrm{Tr}[\rho_{\mathfrak{t}}(E)]} \in [-1, 1] \ . \tag{5}$$

This quantity gives information on the strength of the CISS effect [42, 44] and vanishes when the scatterer is non-chiral.

Microscopic modelling of the scatterer.—For what concerns the microscopic modelling of the scatterer, i.e. the twisted homobilayer TMD, we employed a gapped Dirac-fermion model [32], by introducing a mass term in a tight-binding Hamiltonian for TBG [47, 48]. The details of the complete tight-binding Hamiltonian are reported in Sect. I of the Supplemental Material [49]. Here, given its crucial importance, we focus only on the tight-binding description of SOC in twisted homobilayer TMDs. Among all the possible SOC terms allowed by C_3 symmetry [50], we restrict our analysis to *spin-flipping* nearest-neighbor intra-layer hoppings, since nearest-neighbour and next-nearest-neighbour *spin-conserving* hopping terms yield a vanishing contribution to the CISS strength (5). The fact that the physics of CISS requires, at a general, fundamental level, spin-flip processes is in agreement with the conclusions of Ref. [42].

The SOC Hamiltonian we consider here is then given by [50]:

$$\hat{\mathcal{H}}^{(\ell)}_{\rm SOC} = \frac{2i}{3}\lambda'^{(\ell)}_0 \sum_{\sigma\neq\sigma'} \sum_{\langle m,n\rangle} \sum_{\tau,\tau'} [\hat{\boldsymbol{s}} \times \boldsymbol{d}_{m,\tau,n,\tau'}]_{\sigma,\sigma'} \, |m,\ell,\tau\rangle\langle n,\ell,\tau'| \otimes |\sigma\rangle \, \langle\sigma'| \ , \tag{6}$$

where $\ell = 1, 2$ is the layer index, $\lambda'^{(\ell)}_0 = [\lambda_0 - (-1)^\ell \lambda_{\rm BR}]$ is the SOC parameter (with units of energy), $\hat{\boldsymbol{s}} = (\hat{s}_x, \hat{s}_y, \hat{s}_z)$ is a vector of Pauli matrices, $|m,\ell,\tau\rangle \otimes |\sigma\rangle$ denotes the state of an electron with spin σ localized on site m and sublattice $\tau = A, B$ of layer ℓ (for further details, see Ref. [49]). Finally, $\boldsymbol{d}_{m,\tau,n,\tau'}$ is a dimensionless unit vector lying in the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{y}}$ plane, which points from lattice site n to the nearest-neighbor site m. The symbol $\langle \dots \rangle$ in one of the sums in Eq. (6) refers to the fact that the sum is restricted to nearest-neighbor pairs. One important observation is now in order. The SOC parameter λ_0 is an intrinsic intra-layer term—intrinsic in the sense that is also present in Bernal-stacked bilayer graphene [51]. The layer-dependent term, controlled by $\lambda_{\rm BR}$, stems instead from broken inversion symmetry. This latter term, which, in an ordinary Bernal-stacked bilayer graphene is due to a perpendicular applied electric field [51], here models the fact that twisted homobilayer TMDs do not have an inversion center, even in the absence of extrinsic external electric fields. Finally, we note that, in the case $\lambda_{\rm BR} = 0$, our setup is invariant under the simultaneous exchange of a) top and bottom leads and b) top and bottom layers of the twisted homobilayer TMD. In this case, one has $\sigma_{t'} = -\sigma_t$. For $\lambda_{\rm BR} \neq 0$, $\sigma_{t'} \neq -\sigma_t$.

While certainly approximate, this lightweight model of the scatterer treats on equal footing the structural chirality of the system (i.e. which stems from the non-zero twist angle θ contained in the intra- and inter-layer terms of the twisted TMD Hamiltonian reported in Ref. [49]) *and* the coupling between spin and orbital motion due the SOC term of the Hamiltonian, allowing us to perform a first exploration of their interplay in a chiral quasi-2D solid-state system.

Results.—We now present our main numerical results. We used the Kwant software package [52] to implement the tight-binding model Hamiltonian introduced above and calculate the spin-resolved $S_{\sigma,\sigma'}$ matrix of the two-terminal 3D transport setup presented in Fig. 1. Top and bottom leads are modelled as semi-infinite A-A stacked graphite leads. Each graphite lead is rotated in order to be aligned with the adjacent layer belonging to the twisted homobilayer TMD. The leads have Bloch translational invariance both along the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{y}}$ plane and the $\hat{\boldsymbol{z}}$ direction. The twisted homobilayer TMD has Bloch translational invariance in the $\hat{\boldsymbol{x}}$-$\hat{\boldsymbol{y}}$ plane. The spin-resolved scattering matrix $S_{\sigma,\sigma'}$ has been computed by carrying out integrals over the 2D moiré lattice Brillouin zone.

Fig. 2 displays the spin polarization $P_t(E)$, as defined in Eq. (5), of the electrons transmitted from the top to the bottom lead. These results refer to twisted homobilayer MoTe$_2$ [32] and a twist angle $\theta = 6.01°$. (Additional numerical results for twisted MoS$_2$, MoSe$_2$, WS$_2$, WSe$_2$, and WTe$_2$ are reported in Sect. II of Ref. [49].) For the SOC parameters, we have used the values $\lambda_0 = 220$ meV and $\lambda_{\rm BR} = 0$, which have been taken from Ref. [32]. We clearly see that the spin polarization exceeds 20% (see also Fig. 3(a), where much larger values have been obtained for larger twist angles), which is a giant value, given the atomic thickness of the scatterer. As expected, the calculated spin polarization has opposite sign for two opposite twist angles, θ and $-\theta$. Importantly, the spin polarization vanishes in the non-chiral $\theta = 0$ case.

The dependence of the chirality-induced spin polarization $P_t(E)$ on the twist angle θ is displayed in Fig. 3(a) for three values of θ. Notice that decreasing θ, from $\theta = 13.17°$ down to $\theta = 3.48°$, $|P_t(E)|$ decreases. This is not surprising since in the limit $\theta \to 0$ the system becomes non-chiral and $P_t(E)$ vanishes for all values of E, as seen in Fig. 2. For $\theta = 13.17°$, the spin polarization nearly reaches the impressive value of 60%. Fig. 3(b) shows the dependence of $P_t(E)$ on the magnitude of λ_0 (while keeping $\lambda_{\rm BR} = 0$). We note that increasing SOC leads to an overall increase of $|P_t(E)|$, pointing out the crucial role of this term in the twisted homobilayer TMD Hamiltonian for the emergence of chirality-induced spin polarization.

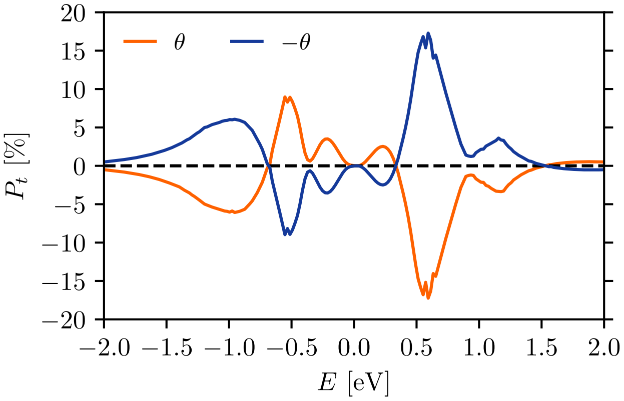


FIG. 4. (Color online) Same as in Fig. 2, but for a finite value of $\lambda_{\rm BR}$, $\lambda_{\rm BR} = 50$ meV. The black dashed line denotes the results for the non-chiral $\theta = 0$ case.

Finally, Fig. 4 shows $P_t(E)$ in the case of finite $\lambda_{\rm BR}$. The important thing to notice is the result for $\theta = 0$, i.e. $P_t(E) = 0\ \forall E$. At $\theta = 0$ and $\lambda_{\rm BR} \neq 0$, our twisted quasi-2D material reduces to a Bernal-stacked homobilayer TMD in the presence of a perpendicular applied electric field. The Hamiltonian of this system breaks inversion symmetry but lacks chirality, yielding *no* chirality-induced spin polarization.

In summary, we have shown that a chirality-induced spin polarization emerges in moiré materials with strong spin-orbit coupling. Surprisingly, we have found that the spin polarization can be giant even for just two twisted monolayers. We have presented a theory of the effect based on the microscopic Hamiltonian of twisted homobilayer transition metal dichalcogenides [32]. Numerical results have been obtained for twisted MoTe$_2$. In this particular case, the effect appears to be gigantic, provided that the twist angle is sufficiently large. As far as the experimental detection of the effect is concerned, one needs either to employ devices with more than two terminals [42] or transcend the linear-response regime [46] (if one insists on using a two-terminal setup).

During the preparation of this manuscript, we learned about a recent theoretical work reporting chirality-induced spin polarization in bulk (three-dimensional) inorganic crystals with homochiral crystal structures, such as tellurium and transition metal disilicide compounds [53].

Acknowledgements.—This work was partially supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 873028 - HYDROTRONICS and the FET project QUEFORMAL (Contract No. 829035). M.P., F.T. and G.I. acknowledge support by the MUR through the following PRIN projects: Project "q-LIMA" (Contract. No. 2020JLZ52N), Project "NEThEQS" (Contract No. 2022B9P8LN), and Project "FIVE2D" (Project No. 2017SRYEJH$_0$01). F.T. also acknowledges support by the Royal Society through the International Exchanges Scheme between the UK and Italy (Grant No. IEC R2 192166). P.J.H. acknowledges support by the Gordon and Betty Moore Foundation's EPiQS Initiative through grant GBMF9463, the Fundacion Ramon Areces, and the ICFO Distinguished Visiting Professor program. F.H.L.K. acknowledges financial support from the ERC TOPONANOP (726001), the Government of Catalonia trough the SGR grant, the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (Ref. SEV-2015-0522) and Explora Ciencia (Ref. FIS2017- 91599-EXP), Fundacio Cellex Barcelona, Generalitat de Catalunya through the CERCA program, the Mineco grant Plan Nacional (Ref. FIS2016-81044-P), and the Agency for Management of University and Research Grants (AGAUR) (Ref. 2017-SGR-1656).

We thank P. Novelli and B. van Wees for useful discussions and correspondence.

Computational resources have been provided by *computing@unipi*, a computing service provided by University of Pisa.

G. M., L. C., and L. L. and contributed equally to this work.

[1] Ray, K., Ananthavel, S. P., Waldeck, D. H. & Naaman, R. Asymmetric scattering of polarized electrons by organized organic films of chiral molecules. *Science* **283**, 814 (1999).

[2] Naaman, R. & Waldeck, D. H. Chiral-induced spin selectivity effect. *J. Phys. Chem. Lett.* **3**, 2178 (2012).

[3] Naaman, R. & Waldeck, D. H. Spintronics and chirality: spin selectivity in electron transport through chiral molecules. *Annu. Rev. Phys. Chem.* **66**, 263 (2015).

[4] Naaman, R., Paltiel, Y. & Waldeck, D. H. Chiral molecules and the electron spin. *Nat. Rev. Chem.* **3**, 250 (2019).

[5] Naaman, R., Paltiel, Y. & Waldeck, D. H. Chiral molecules and the spin selectivity effect. *J. Phys. Chem. Lett.* **11**, 3660 (2020).

[6] Yang, S. H., Naaman, R., Paltiel, Y. & Parkin, S. S. P. Chiral spintronics. *Nat. Rev. Phys.* **3**, 328 (2021).

[7] Rosenberg, R. A., Abu Haija, M. & Ryan, P. J. Chiral-selective chemistry induced by spin-polarized secondary electrons from a magnetic substrate. *Phys. Rev. Lett.* **101**, 178301 (2008).

[8] Waldeck, D. H., Naaman, R. & Paltiel, Y. The spin selectivity effect in chiral materials. *APL Mater.* **9**, 040902 (2021).

[9] Shiota, K., Inui, A., Hosaka Y., Amano, R., Ōnuki, Y., Hedo, M., Nakama, T., Hirobe, D., Ohe, J., Kishine, J., Yamamoto, Y. M., Shishido, H. & Togawa, Y. Chirality-induced spin polarization over macroscopic distances in chiral disilicide crystals. *Phys. Rev. Lett.* **127**, 126602 (2021).

[10] Evers, F., Aharony, A., Bar-Gill, N., Entin-Wohlman, O., Hedegård, P., Hod, O., Jelinek, P., Kamieniarz, G., Lemeshko, M., Michaeli, K., Mujica, V., Naaman, R., Paltiel, Y., Refaely-Abramson, S., Tal, O., Thijssen, J., Thoss, M., van Ruitenbeek, J. M., Venkataraman, L., Waldeck, D. H., Yan, B. & Kronik, L. Theory of chirality induced spin selectivity: progress and challenges. *Adv. Mater.* **34**, 2106629 (2022).

[11] Liu, Y., Xiao, J., Koo, J. & Yan, B. Chirality-driven topological electronic structure of DNA-like materials. *Nat. Mater.* **20**, 638 (2021).

[12] Fransson, J. Chirality-induced spin selectivity: the role of electron correlations. *J. Phys. Chem. Lett.* **10**, 7126 (2019).

[13] Fransson, J. Vibrational origin of exchange splitting and "chiral-induced spin selectivity". *Phys. Rev. B* **102**, 235416 (2020).

[14] Volosniev, A. G., Alpern, H., Paltiel, Y., Millo, O., Lemeshko, M. & Ghazaryan, A. Interplay between friction and spin-orbit coupling as a source of spin polarization. *Phys. Rev. B* **104**, 024430 (2021).

[15] Andrei, E. Y. & MacDonald, A. H. Graphene bilayers with a twist. *Nat. Mater.* **19**, 1265 (2020).

[16] Andrei, E. Y., Efetov, D. K., Jarillo-Herrero, P., MacDonald, A. H., Mak, K. F., Senthil, T., Tutuc, E., Yazdani, A. & Young, A. F. The marvels of moiré materials. *Nat. Rev. Mater.* **6**, 201 (2021).

[17] Kennes, D. M., Claassen, M., Xian, L., Georges, A., Millis, A. J., Hone, J., Dean, C. R., Basov, D. N., Pasupathy, A. & Rubio, A. Moiré heterostructures as a condensed-matter quantum simulator. *Nat. Phys.* **17**, 155 (2021).

[18] Cao, Y., Fatemi, V., Fang, S., Watanabe, K., Taniguchi, T., Kaxiras, E. & Jarillo-Herrero, P. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43 (2018).

[19] Cao, Y., Fatemi, V., Demir, A., Fang, S., Tomarken, S. L., Luo, J. Y., Sanchez-Yamagishi, J. D., Watanabe, K., Taniguchi, T., Kaxiras, E., Ashoori, R. C. & Jarillo-Herrero, P. Correlated insulator behaviour at half-filling in magic-angle graphene superlattices. *Nature* **556**, 80 (2018).

[20] Lopes dos Santos, J. M. B., Peres, N. M. R. & Castro Neto, A. H. Graphene bilayer with a twist: electronic structure. *Phys. Rev. Lett.* **99**, 256802 (2007).

[21] Shallcross, S., Sharma, S. & Pankratov, O. A. Quantum interference at the twist boundary in graphene. *Phys. Rev. Lett.* **101**, 056803 (2008).

[22] Mele, E. J., Commensuration and interlayer coherence in twisted bilayer graphene. *Phys. Rev. B* **81**, 161405(R) (2010).

[23] Li, G., Luican, A., Lopes dos Santos, J. M. B., Castro Neto, A. H., Reina, A., Kong, J. & Andrei, E. Y. Observation of Van Hove singularities in twisted graphene

layers. *Nat. Phys.* **6**, 109 (2010).

[24] Shallcross, S., Sharma, S., Kandelaki, E. & Pankratov, O. A. Electronic structure of turbostratic graphene. *Phys. Rev. B* **81**, 165105 (2010).

[25] Suárez Morell, E., Correa, J. D., Vargas, P., Pacheco, M. & Barticevic, Z. Flat bands in slightly twisted bilayer graphene: tight-binding calculations. *Phys. Rev. B* **82**, 121407(R) (2010).

[26] Bistritzer R. & MacDonald, A. H. Transport between twisted graphene layers. *Phys. Rev. B* **81**, 245412 (2010).

[27] Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proc. Natl. Acad. Sci. (USA)* **108**, 12233 (2011).

[28] Lopes dos Santos, J. M. B., Peres, N. M. R. & Castro Neto, A. H. Continuum model of the twisted graphene bilayer. *Phys. Rev. B* **86**, 155449 (2012).

[29] Landau, L. D. & Lifshitz, E. M. *Course of Theoretical Physics: Electrodynamics of Continuos Media* (Pergamon Press, New York, 1984).

[30] Kim, C. J., Sánchez-Castillo, A., Ziegler, Z., Ogawa, Y., Noguez, C. & Park, J. Chiral atomically thin films. *Nat. Nanotech.* **11**, 520 (2016).

[31] Suárez Morell, E., Chico, L. & Brey, L. Twisting dirac fermions: circular dichroism in bilayer graphene. *2D Mater.* **4**, 035015 (2017).

[32] Wu, F., Lovorn, T., Tutuc, E., Martin, I. & MacDonald, A. H. Topological insulators in twisted transition metal dichalcogenide homobilayers. *Phys. Rev. Lett.* **122**, 086402 (2019).

[33] Cai, J., Anderson, E., Wang, C., Zhang, C., Liu, X., Holtzmann, W., Zhang, Y., Fan, F., Taniguchi, T., Watanabe, K., Ran, Y., Cao, T., Fu, L., Xiao, D., Yao, W. & Xu, X. Signatures of fractional quantum anomalous Hall states in twisted $\mathrm{MoTe_2}$. *Nature* **622**, 63 (2023).

[34] Park, H., Cai, J., Anderson, E., Zhang, Y., Zhu, J., Liu, X., Wang, C., Holtzmann, W., Hu, C., Liu, Z., Taniguchi, T., Watanabe, K., Chu, J.-H., Cao, T., Fu, L., Yao, W., Chang, C.-Z., Cobden, D., Xiao, D. & Xu, X. Observation of fractionally quantized anomalous Hall effect. *Nature* **622**, 74 (2023).

[35] Zeng, Y., Xia, Z., Kang, K., Zhu, J., Knüppel, P., Vaswani, C., Watanabe, K., Taniguchi, T., Mak, K. F. & Shan, J. Thermodynamic evidence of fractional Chern insulator in moiré $\mathrm{MoTe_2}$. *Nature* **622**, 69 (2023).

[36] Xu, F., Sun, Z., Jia, T., Liu, C., Xu, C., Li, C., Gu, Y., Watanabe, K., Taniguchi, T., Tong, B., Jia, J., Shi, Z., Jiang, S., Zhang, Y., Liu, X. & Li, T. Observation of integer and fractional quantum anomalous Hall effects in twisted bilayer $\mathrm{MoTe_2}$. *Phys. Rev. X* **13**, 031037 (2023).

[37] Naaman, R., Waldeck, D. H. & Fransson, J. New perspective on electron transfer through molecules. *J. Phys. Chem. Lett* **13**, 11753 (2022).

[38] Provost, J. P. & Vallee, G. Riemannian structure on manifolds of quantum states, *Comm. Math. Phys.* **76**, 289 (1980).

[39] Xiao, D., Chang, M. C. & Niu, Q. Berry phase effects on electronic properties. *Rev. Mod. Phys.* **82**, 1959 (2010).

[40] Törmä, P., Peotta, S. & Bernevig, B. A. Superconductivity, superfluidity and quantum geometry in twisted multilayer systems. *Nat. Rev. Phys.* **4**, 528 (2022).

[41] Büttiker, M. Symmetry of electrical conduction. *IBM J. Res. Dev.* **32**, 317 (1988).

[42] Yang, X., van der Wal, C. H. & van Wees, B. J. Spin-dependent electron transmission model for chiral molecules in mesoscopic devices. *Phys. Rev. B* **99**, 024418 (2019).

[43] Nazarov, Y. V. & Blanter, Y. M. *Quantum Transport*, (Cambridge University Press, 2012).

[44] Wolf, Y., Liu, Y., Xiao, J., Park, N. & Yan, B. Unusual spin polarization in the chirality-induced spin selectivity. *ACS Nano* **16**, 18601 (2022).

[45] Xiao, J., Zhao, Y. & Yan, B. Nonreciprocal nature and induced tunneling barrier modulation in chiral molecular devices. Preprint at https://doi.org/10.48550/arXiv.2201.03623.

[46] A powerful, general approach to spin and charge transport through any (non-magnetic) chiral system has been proposed in Ref. [42] and by the authors of the following works: Yang, X., van der Wal, C. H. & van Wees, B. J. Detecting chirality in two-terminal electronic nanodevices. *Nano Lett.* **20**, 6148 (2020) and Yang, X. & van Wees, B. J., Linear-response magnetoresistance effects in chiral systems *Phys. Rev. B* **104**, 155420 (2021). In the framework of such phenomenological approach one can include for example dephasing and spin relaxation. Our aim, however, is to present quantitative numerical predictions for twisted 2D materials and this is why we have used the standard spin-resolved scattering matrix approach based on microscopic Hamiltonians.

[47] de Laissardière, G. T., Mayou, D. & Magaud, L. Numerical studies of confined states in rotated bilayers of graphene. *Phys. Rev. B* **86**, 125413 (2012).

[48] dos Santos, J. M. B. L., Peres, N. M. R. & Neto, A. H. C. Continuum model of the twisted graphene bilayer. *Phys. Rev. B* **86**, 155449 (2012).

[49] See the Supplemental Material file for additional details on the derivation of the tight-binding Hamiltonian and further numerical results.

[50] Kochan, D., Irmer, S. & Fabian, J. Model spin-orbit coupling Hamiltonians for graphene systems. *Phys. Rev. B* **95**, 165415 (2017).

[51] Konschuh, S., Gmitra, M., Kochan, D. & Fabian, J. Theory of spin-orbit coupling in bilayer graphene. *Phys. Rev. B* **85**, 115423 (2012).

[52] Groth, C. W., Wimmer, M., Akhmerov, A. R. & Waintal, X. Kwant: a software package for quantum transport. *New J. Phys.* **16**, 063065 (2014).

[53] Yang, Q., Li, Y., Felser, C. & Yan, B. Chirality induced spin selectivity in chiral crystals. Preprint at https://doi.org/10.48550/arXiv.2312.04366.

Supplemental Material for: “Chirality-Induced Spin Polarization in Twisted Bilayer Transition Metal Dichalcogenides”

Guido Menichetti,[1] Lorenzo Cavicchi,[2] Leonardo Lucchesi,[1, 3] Fabio Taddei,[4, 2] Giuseppe Iannaccone,[3] Pablo Jarillo-Herrero,[5] Claudia Felser,[6] Frank H. L. Koppens,[7, 8] Marco Polini[1, 7]

[1]*Dipartimento di Fisica dell’Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy*
[2]*Dipartimento di Ingegneria dell’Informazione dell’Università di Pisa, Via Girolamo Caruso 16, I-56122 Pisa, Italy*
[3]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[4]*NEST, Istituto Nanoscienze-CNR, Piazza S. Silvestro 12, I-56126 Pisa, Italy*
[5]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[6]*Max Planck Institute for Chemical Physics of Solids, Nöthnitzer Str. 40, Dresden 01187, Germany*
[7]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*
[8]*ICREA-Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys 23, 08010 Barcelona, Spain*

In this Supplemental Material we present more details on the tight-binding Hamiltonian we have used in our numerical calculations. We also present further numerical results concerning the chirality-induced spin-polarization for two of the most common families of TMDs.

SECTION I: MODEL HAMILTONIAN

Here we present more details on the tight-binding model we have used to describe electrons roaming in the twisted TMD homobilayer moiré superlattice.

We start from the tight binding model for TBG [S1–S3]. The basis of Bloch states is built from the p_z atomic orbitals of Carbon. We introduce the localized atomic orbitals centered at the point $\boldsymbol{d}_{\tau,\ell}+\boldsymbol{t}_{n,\ell}$, i.e.

$$\langle \boldsymbol{r}|n,\ell,\tau\rangle = \phi(\boldsymbol{r}-\boldsymbol{d}_{\tau,\ell}-\boldsymbol{t}_{n,\ell}) \ , \tag{S1}$$

where $\phi(\boldsymbol{r})$ is the wavefunction of a p_z orbital centered at the origin, $\boldsymbol{d}_{\tau,\ell}$ is the basis vector of the sublattice τ in layer ℓ, whereas the symbol $\boldsymbol{t}_{n,\ell}$ is a shorthand for

$$\boldsymbol{t}_{n,\ell} = n_1\tilde{\boldsymbol{t}}_{1,\ell} + n_2\tilde{\boldsymbol{t}}_{2,\ell} \quad \text{with } n_1,\, n_2 \in \mathbb{N} \ . \tag{S2}$$

The vectors $\tilde{\boldsymbol{t}}_{1/2,\ell}$ are primitive translation vectors of the graphene lattice in layer ℓ, and the sum over n should be intended as

$$\sum_n [\cdots] = \sum_{n1,n2\in\mathbb{N}} [\cdots] \ . \tag{S3}$$

The atomic orbitals are assumed to be orthogonalized according to

$$\langle n,\ell,\tau|n',\ell',\tau'\rangle = \delta_{n,n'}\delta_{\ell,\ell'}\delta_{\tau,\tau'} \ . \tag{S4}$$

In the two-center approximation, and retaining only the nearest-neighbour contributions, the intra-layer Hamiltonian of graphene in layer ℓ takes the form

$$\hat{H}^{(\ell)}_{\rm intra} = -t\sum_{\langle m,n\rangle}\sum_{\tau,\tau'} |m,\ell,\tau\rangle\langle n,\ell,\tau'|(1-\delta_{\tau,\tau'}) \ , \tag{S5}$$

where the energy t is given by

$$-t \equiv \int d\boldsymbol{r}\ \phi^*(\boldsymbol{r})V(\boldsymbol{r}-\boldsymbol{d}_{\tau,1})\phi(\boldsymbol{r}-\boldsymbol{d}_{\tau,1}) = \int d\boldsymbol{r}\ \phi^*(\boldsymbol{r})V(\boldsymbol{r}-\boldsymbol{d}_{\tau,2})\phi(\boldsymbol{r}-\boldsymbol{d}_{\tau,2}) \ , \tag{S6}$$

$V(\boldsymbol{r})$ being the spherically-symmetric potential of a Carbon atom centered at the origin. The sum over $\langle m,n\rangle$ runs over neighboring orbitals, i.e. the states $|m,\ell,\tau\rangle$ and $|n,\ell,\tau'\rangle$ in Eq. (S5) correspond to neighbouring orbitals.

In this work we have chosen the following primitive translation vectors

$$\tilde{\boldsymbol{t}}_{1/2,\ell=1} = \left(\mp\frac{a}{2}, \frac{a\sqrt{3}}{2}\right) \quad , \quad \tilde{\boldsymbol{t}}_{1/2,\ell=2} = \mathcal{R}(\theta)\left(\mp\frac{a}{2}, \frac{a\sqrt{3}}{2}\right) \ , \tag{S7}$$

where $\mathcal{R}(\theta)$ is the rotation matrix defined by:

$$\mathcal{R}\left(\theta\right) = \cos(\theta)\mathbb{I}_{2\times 2} - i\sin(\theta)\sigma_y = \begin{pmatrix} \cos(\theta) & \pm\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \ . \tag{S8}$$

As in the main text, layer 1 (top) is the reference layer, while layer 2 (bottom) is counter-clockwise rotated by an angle θ. In addition, the basis vectors are

$$\boldsymbol{d}_{\tau,\ell} = \begin{cases} \frac{a}{\sqrt{3}}\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right) \ , & \text{if layer} = 1 \text{ and sub-lattice} = B \ . \\ -\frac{a}{\sqrt{3}}\mathcal{R}(\theta)\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right) \ , & \text{if layer} = 2 \text{ and sub-lattice} = A \ . \\ \boldsymbol{0} \ , & \text{otherwise} \ . \end{cases} \tag{S9}$$

The choice of these translation and basis vectors is such that in the limit $\theta \to 0$ one obtains AB-stacked bilayer graphene.

We now describe the tunneling of electrons between orbitals in different layers. The inter-layer Hamiltonian can be written as

$$\hat{H}^{(\ell,\ell')}_{\text{inter}} = \sum_{n,n'}\sum_{\tau,\tau'} h_{\tau,\tau'}(\boldsymbol{d}_{\tau,\ell} + \boldsymbol{t}_{n,\ell} - \boldsymbol{d}_{\tau',\ell'} - \boldsymbol{t}_{n',\ell'})|n,\ell,\tau\rangle\langle n',\ell',\tau'| + \text{H.c.} \ . \tag{S10}$$

An empirical form of the transfer integral between two p_z orbitals of Carbon atoms in the Slater-Koster approximation is given by [S2]:

$$h_{\tau,\tau'}(\boldsymbol{R}_{n,n'}) = -t\exp\left(-\frac{|\boldsymbol{R}_{n,n'}| - a}{\lambda}\right)\frac{|\boldsymbol{R}_{n,n'}\cdot\boldsymbol{e}_{\parallel}|^2}{|\boldsymbol{R}_{n,n'}|^2} + t_{\perp}\exp\left(-\frac{|\boldsymbol{R}_{n,n'}| - d}{\lambda}\right)\frac{|\boldsymbol{R}_{n,n'}\cdot\boldsymbol{e}_{\perp}|^2}{|\boldsymbol{R}_{n,n'}|^2} \ . \tag{S11}$$

In Eq. (S11) we have introduced the short-hand notation $\boldsymbol{R}_{n,n'} \equiv \boldsymbol{d}_{\tau,\ell} + \boldsymbol{t}_{n,\ell} - \boldsymbol{d}_{\tau',\ell'} - \boldsymbol{t}_{n',\ell'}$, for which indeed $h_{\tau,\tau'}(\boldsymbol{d}_{\tau,\ell} + \boldsymbol{t}_{n,\ell} - \boldsymbol{d}_{\tau',\ell'} - \boldsymbol{t}_{n',\ell'}) = h_{\tau,\tau'}(\boldsymbol{R}_{n,n'})$. In Eq. (S11) we have also introduced the decay length λ and the inter-layer distance d. The hopping parameter $t_{\perp}$ is given by:

$$t_{\perp} = \int d\boldsymbol{r}\phi^*(\boldsymbol{r})V(\boldsymbol{r} - d\boldsymbol{e}_z)\phi(\boldsymbol{r} - d\boldsymbol{e}_z) \ , \tag{S12}$$

which is the transfer integral between two p_z orbitals that are one on top of the other, vertically displayed by a distance d.

In this work we study vertical transport through twisted homobilayer TMDs, whose Hamiltonian can be obtained by modifying the above tight-binding model into a twisted bilayer gapped graphene one [S4]. To this end, we reproduce the monolayer TMD energy gap Δ and carrier effective mass m^* by adding a staggered potential E_A, E_B on the sub-lattice sites and re-scaling the intra-layer nearest-neighbor hopping t, while retaining the lattice parameter of graphene a:

$$\hat{H}^{(\ell)}_{\text{intra}} = \sum_{m}\sum_{\tau} E_\tau|m,\ell,\tau\rangle\langle m,\ell,\tau| - \tilde{t}\sum_{\langle m,n\rangle}\sum_{\tau\tau'}|m,\ell,\tau\rangle\langle n,\ell,\tau'|(1-\delta_{\tau\tau'}) \ , \tag{S13}$$

where

$$\tilde{t} = \frac{\hbar}{\sqrt{3}a}\sqrt{\frac{2\Delta}{3m^*}} \ . \tag{S14}$$

The numerical results presented and discussed in the main text are obtained for the following choice of parameters: $a = 0.246$ nm, $\Delta = 1.1$ eV, and $m^* = 0.6\ m_{\text{e}}$, m_{e} being the bare electron mass in vacuum. The inter-layer hopping energy and distance appearing in the inter-layer Hamiltonian term (S11) are instead fixed to $t_{\perp} = 0.15$ eV and $d = 0.69$ nm, while the exponential decay length is kept equal to the one in TBG, i.e. $\lambda = 0.184\sqrt{3}a$ [S2]. These parameters have been chosen in order to effectively reproduce the energy bands of twisted MoTe$_2$.

In Fig. S1 we compare our results for the energy bands obtained from the above described tight-binding model—panel (a)—with the results of Ref. [S4], which were obtained from a continuum model Hamiltonian.

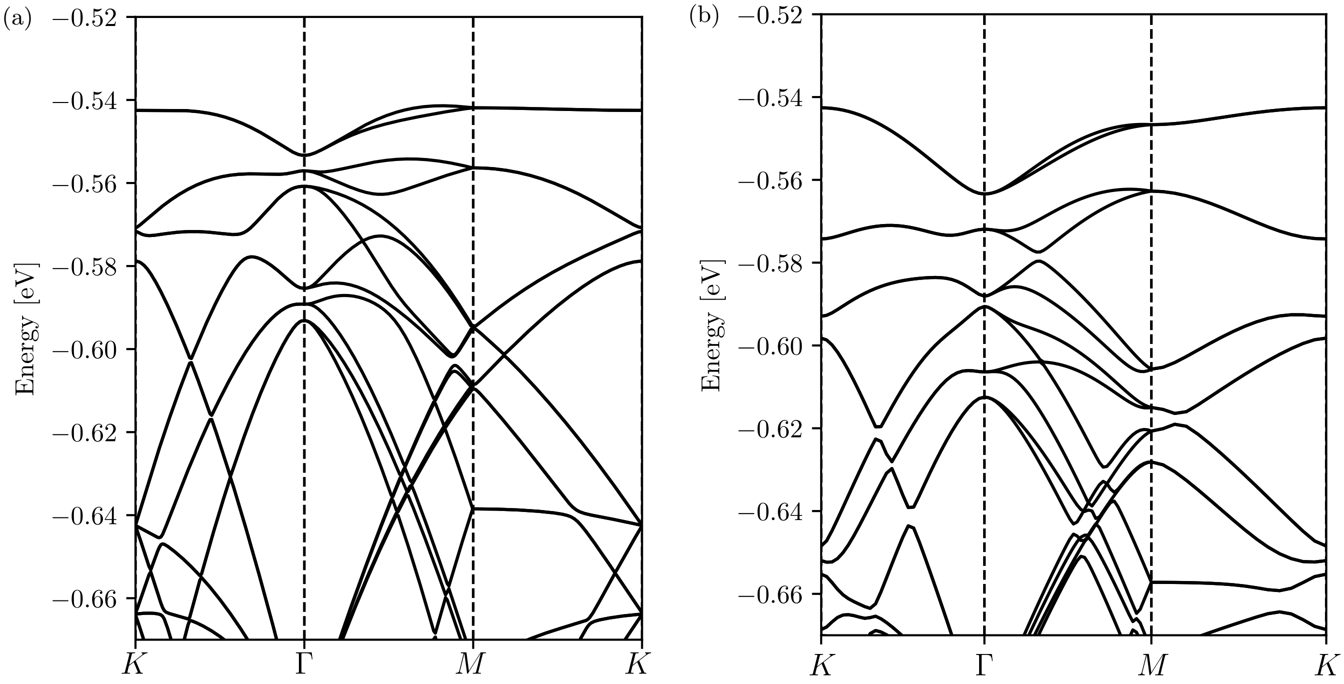


FIG. S1. Band structure of twisted $MoTe_2$ obtained from our tight-binding model—panel (a)—and from the continuum model of Ref. [S4]—panel (b). Results in this plot refer to a twist angle $\theta = 3.89^\circ$. The parameters of the continuum model Hamiltonian are extracted from Ref. [S4]. The dependence of the bands on the wave-vector $\boldsymbol{k}$ is displayed along the high-symmetry path K-Γ-M-K of the moiré Brillouin zone.

SECTION II: ADDITIONAL NUMERICAL RESULTS

In this Section we present additional numerical results. In the main text, indeed, we focussed on for the case of twisted $MoTe_2$. For the sake of completeness, in Fig. S2 we show the chirality-induced spin polarization P_t for two of the most common families of TMDs. In panel (a) we report results for twisted homobilayers composed by MoS_2, $MoSe_2$, and $MoTe_2$ (already presented in the main text). In panel (b) we do the same for WS_2, WSe_2, and WTe_2. Both panels have been calculated by setting $\theta = 13.17^\circ$.

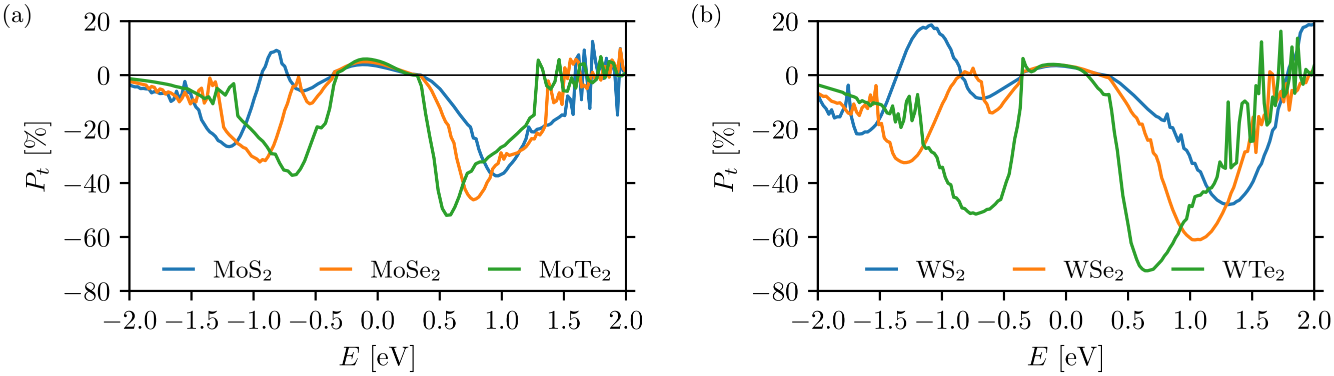


FIG. S2. (Color online) Results for the spin polarization $P_t(E)$ for two different families of twisted TMDs. Results in this figure refer to a single twist angle, i.e. $\theta = 13.17^\circ$. Panel (a) Results for MoX_2 with $X = S, Se$, and Te. Panel (b) Results for WX_2 with $X = S, Se$, and Te. In both panels the solid black line has been added as a guide to the eye. The parameters of the microscopic Hamiltonians used to produce these numerical results have been reported in Table S1.

The microscopic parameters of the Hamiltonian used to compute the spin polarization $P_t(E)$ in Fig. S2 are collected in Table S1. Each Hamiltonian contain the following parameters: the energy gap Δ, the hole-carrier effective mass m^* (expressed in units of the bare electron mass in vacuum), the SOC parameter λ_0 discussed in the main text, and the interlayer distance d. The Table reports also the SOC to energy gap ratio λ_0/Δ and the maximum of the absolute value of the chirality-induced spin-polarization:

$$|P_t|^{\max} \equiv \max_{E\in[-2,2]\text{eV}} |P_t(E)| \;, \tag{S15}$$

extracted from Fig. S2. As it can be inferred from Fig. S2, within a given TMD family, $|P_t|^{\max}$ is achieved in the material with the maximum value of λ_0/Δ.

	Δ [eV]	m^* [m_e]	λ_0 [meV]	d [nm]	λ_0/Δ	$\lvert P_t\rvert^{\max}$ [%]
MoS_2 [S5, S6]	1.62	0.58	255	0.62	0.16	37.3
$MoSe_2$ [S5, S6]	1.4	0.67	263	0.65	0.19	46.2
$MoTe_2$ [S4]	1.1	0.62	220	0.69	0.20	59.5
WS_2 [S5, S6]	1.74	0.42	459	0.61	0.26	48.0
WSe_2 [S5, S6]	1.43	0.45	446	0.64	0.31	61.1
WTe_2 [S5, S6]	0.86	0.41	401	0.69	0.47	72.5

TABLE S1. Numerical parameters employed to obtain the results shown in Fig. S2. Energy gap Δ (eV), hole-carrier effective mass m^* (m_e) in units of the bare electron mass in vacuum m_e, the SOC parameter λ_0 discussed in the main text, the interlayer distance d (in nm), the SOC to energy gap ratio λ_0/Δ, and $|P_t|^{\max}$ as defined in Eq. (S15). This last quantity has been extracted from Fig. S2 and refers to $\theta = 13.17°$.

In order to illustrate the role of the energy gap Δ on the quantity $P_t(E)$, in Fig. S3 we plot three different results for $P_t(E)$ in $MoTe_2$. Different curves in this figure have been obtained by artificially changing the energy gap Δ while keeping fixed all the other microscopic parameters of $MoTe_2$. The twist angle is fixed at $\theta = 6.01°$. It is evident that $|P_t|^{\max}$ increases with decreasing Δ (thereby increasing the value of the ratio λ_0/Δ).

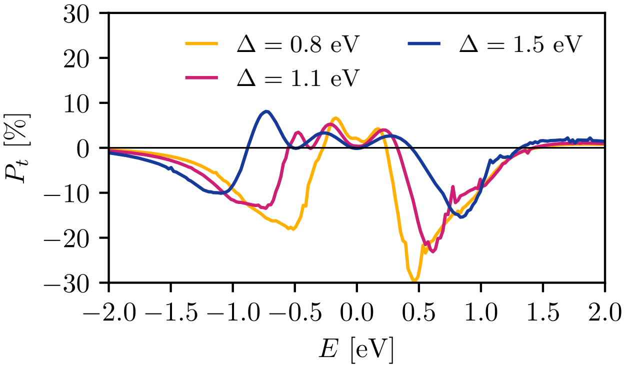


FIG. S3. (Color online) Chirality-induced spin polarization $P_t(E)$ in $MoTe_2$. Different colors refer to different values of Δ: $\Delta = 0.8$ eV (orange), $\Delta = 1.1$ eV (magenta), and $\Delta = 1.5$ eV (blue). All the results have been obtained by setting $\theta = 6.01°$.

We conclude this Section by showing the dependence of $|P_t|^{\max}$ on the twist angle θ in Fig. S4. Results in this plot refer to twisted $MoTe_2$. For the computed angles, a linear behavior is evident.

[S1] Lopes dos Santos, J. M. B., Peres, N. M. R. & Castro Neto, A. H. Graphene bilayer with a twist: electronic structure. *Phys. Rev. Lett.* **99**, 256802 (2007).

[S2] de Laissardière, G. T., Mayou, D. & Magaud L. Numerical studies of confined states in rotated bilayers of graphene. *Phys. Rev. B* **86**, 125413 (2012).

[S3] Novelli, P., Torre, I., Koppens, F. H. L., Taddei, F. & Polini, M. Optical and plasmonic properties of twisted bilayer graphene: Impact of interlayer tunneling asymmetry and ground-state charge inhomogeneity. *Phys. Rev. B* **102**, 125403 (2020).

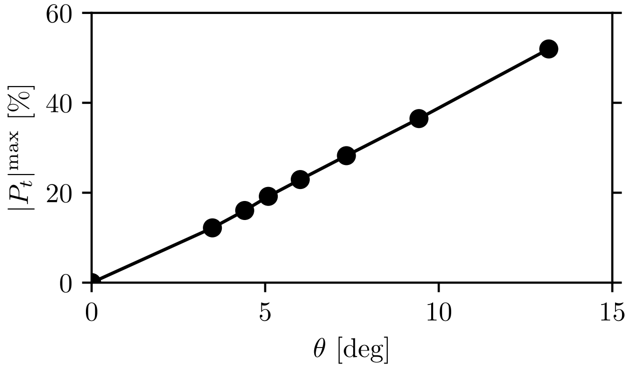


FIG. S4. The quantity $|P_t|^{\max}$ is plotted as a function of the twist angle θ. The black solid line is a guide to the eye. All the other parameters are the same as in Fig. 2 in the main text.

[S4] Wu, F., Lovorn, T., Tutuc, E., Martin, I. & MacDonald, A. H. Topological insulators in twisted transition metal dichalcogenide homobilayers. *Phys. Rev. Lett.* **122**, 086402 (2019).

[S5] Zibouche, N., Kuc, A. , Musfeldt, J. & Heine, T. Transition-metal dichalcogenides for spintronic applications. *Ann. Phys. (Berlin)* **526**, 395 (2014).

[S6] Kim, H.-G. & Choi, H. J. Thickness dependence of work function, ionization energy, and electron affinity of Mo and W dichalcogenides from DFT and GW calculations. *Phys. Rev. B* **103**, 085404 (2021).