Thermal transport in compensated semimetals: Effect of electron-electron scattering on Lorenz ratio

Mohammad Zarenia, Alessandro Principi, and Giovanni Vignale

Department of Physics and Astronomy, University of Missouri, Columbia, Missouri 65211, USA
and School of Physics, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

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.

I. 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 nondegenerate 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 ($\rho_{th} = \kappa^{-1}$)—defined under the standard condition of zero electric current—is primarily controlled by momentum-nonconserving 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. 1(a)]. 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-nonconserving 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. 1(a)]. Although the total momentum is now zero, electron-hole collisions transfer momentum between electrons and holes, giving rise to the Coulomb resistivity ρ_{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]

$$L = L_0 \left(1 + \frac{1}{\Gamma^2}\right), \tag{1}$$

where $\Gamma^2 = (3/\pi^2)(\rho_{el,dis}/\rho_{el}) \ll 1$ is the ratio of the ordinary Drude resistivity, $\rho_{el,dis}$, to the Coulomb resistivity ρ_{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, $\mu/(k_B T) \ll \Gamma$.] The resulting Lorenz ratio, $L > L_0$, is clearly seen in the experiments of Ref. [3], which we reproduce in Fig. 1(b), 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 carrier, then the requirement of 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
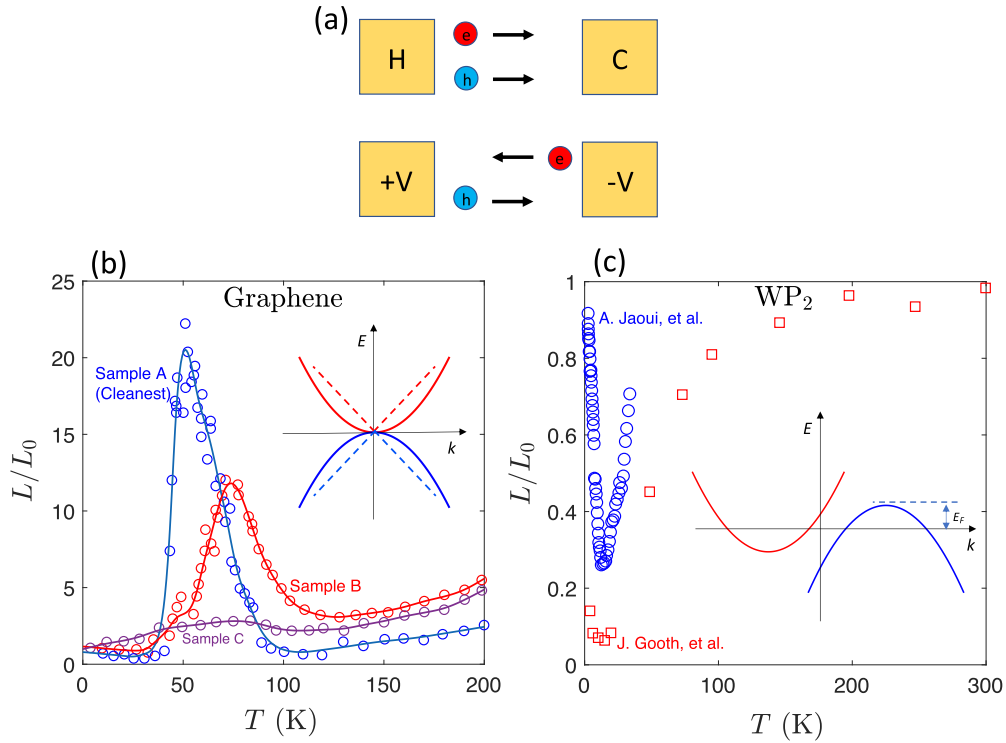
FIG. 1. (a) Schematic illustration of the difference between heat and charge current in a charge-neutral system. A thermal current can be set up in a semimetal simply by letting electrons and holes drift with equal speeds in the same direction (upper row). Electric field causes electrons and holes to 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 (c) 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.

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. 1(c). 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. 1(c), 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 types of semimetals, which we will hereafter refer to as "type I" (graphene-like) and "type II" (WP$_2$-like). Electrons and holes are degenerate Fermi liquids in type-II semimetals, where the inverse quasiparticle lifetime scales as $(k_BT)^2/\varepsilon_F$, but, in type-I semimetals, they are nondegenerate, 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): ρ_{el} is essentially independent of temperature in single- and double-layer graphene, but becomes proportional to $(k_BT/\varepsilon_F)^2 \ll 1$ in type-II semimetals. The small value of ρ_{el} suggests that the "hydrodynamicity" parameter $1/\Gamma^2$ of Eq. (1) in type-II semimetals 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 in passing 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 electric currents of electrons and holes cancel against each other.
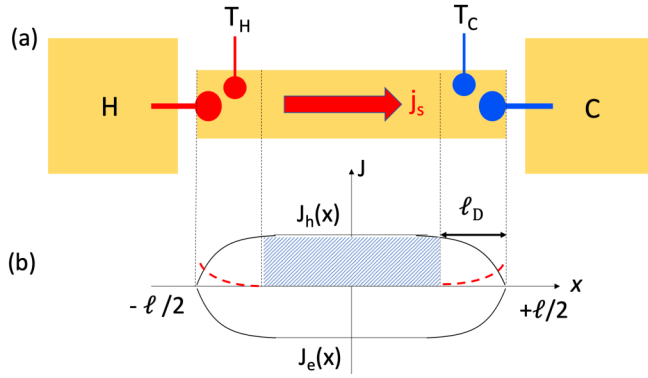
FIG. 2. (a) Schematics of the thermal conductivity measurement. (b) Spatial distribution of the electron and hole components of the electric current. ℓ_D is the diffusion length. The ambipolar transport region is shaded. The red dashed lines represent the enhanced density of electrons and holes near the contacts.

Another possible mechanism for the reduction of L in type-II semimetals is inelastic electron-phonon scattering. Indeed, below the Bloch-Grüneisen temperature inelastic electron-phonon scattering becomes nearly momentum-conserving and its contributions to the electrical and the thermal resistivity [in reduced units, as explained after Eq. (7)] scale as T^5 and T^3, respectively [20,27]: this shifts the balance in favor of the electric conductivity and thus tends to reduce the Lorenz ratio. In the same regime, however, it is possible that electron-electron scattering dominates over electron-phonon scattering, shifting again the balance in favor of the thermal conductivity and making $L > 1$. For semimetals in which this is the case, we are then left with the following problem: Why is the ambipolar channel of thermal conduction apparently disabled in these systems while it is clearly operative in graphene? In this paper we propose a resolution of this puzzle.

II. AMBIPOLAR TRANSPORT IN THE PRESENCE OF CONTACTS: QUALITATIVE DESCRIPTION

In a typical thermal conductivity measurement [see Fig. 2(a)] 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 noninteracting 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 [28] 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. 2(b). 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, $D \sim (k_B T/m)\tau$ in a nondegenerate 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 Ref. [29]) 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 $\rho_{th,ambi}$. (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 $\rho_{th,uni} = (\kappa_e + \kappa_h)^{-1}$. The thermal resistivities of the three

sections add in series, leading us to the final result

$$\rho_{\text{th}} = \alpha(\ell)\rho_{\text{th,uni}} + [1 - \alpha(\ell)]\rho_{\text{th,ambi}}, \tag{4}$$

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

$$\alpha(\ell) = \frac{2\ell_D}{\ell} \tanh\left(\frac{\ell}{2\ell_D}\right). \tag{5}$$

If, as we expect in very clean samples, $\rho_{\text{th,ambi}} \ll \rho_{\text{th,uni}}$, Eq. (4) can be further simplified to

$$\rho_{\text{th}} = \alpha(\ell)\rho_{\text{th,uni}}. \tag{6}$$

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 $\rho_{\text{th}} \simeq \rho_{\text{th,uni}}$ 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 type-II compensated semimetals and graphene arises from the difference between their electron-hole equilibration times. In type-II semimetals 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 nonequilibrium 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.

III. 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 ($-\ell/2 \leqslant x \leqslant \ell/2$). 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 $|F_{ns}\rangle = {}^t(-eE, -k_B\partial_x T)$ 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

$$\hat{\rho} = \begin{pmatrix} \rho_{\text{el}} + Q^2\rho_{\text{th}} & -Q\rho_{\text{th}} \\ -Q\rho_{\text{th}} & \rho_{\text{th}} \end{pmatrix}, \tag{7}$$

where ρ_{el} and ρ_{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 $k_B^2 T$, respectively, while Q is the Seebeck coefficient in units of k_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 downfold 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 Lorenz ratio and the Seebeck coefficient.

In the three-mode theory, the fields are related to the currents via a 3×3 resistivity matrix:

$$\begin{pmatrix} |F_{ns}\rangle \\ F_\delta \end{pmatrix} = \left(\begin{array}{c|c} \hat{\rho}_{ns} & |\rho_\delta\rangle \\ \hline \langle\rho_\delta| & \rho_{\delta\delta} \end{array} \right) \begin{pmatrix} |j_{ns}\rangle \\ j_\delta \end{pmatrix}. \tag{8}$$

Here $\hat{\rho}_{ns}$ is a 2×2 block, whereas $|\rho_\delta\rangle$ is a two-component vector. Hereafter $\langle v|$ denotes the transpose 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-nonconserving 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 downfolding 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 the Appendix: 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, a 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:

$$\partial_t n_\delta + \partial_x j_\delta = -\frac{\nu_0}{\tau_r}(\mu_e - \mu_h), \tag{9}$$

where τ_r is a phenomenological electron-hole recombination time, while ν_0 is the density of states of electrons and holes (assumed to be equal) at the Fermi level. Equation (9) encodes the fact that nonequilibrium electron and hole densities are not separately conserved and therefore the imbalance density decays over time with a typical timescale 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, which locally preserves charge neutrality, 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

$$\partial_x^2 j_\delta(x) = \frac{\nu_0}{\tau_r} F_\delta. \tag{10}$$

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

$$j_\delta(x) = -\frac{\langle \rho_\delta | j_{ns} \rangle}{\rho_{\delta\delta}}[1 - \alpha(\ell, x)]. \tag{11}$$

In this equation,

$$\alpha(\ell, x) = \frac{\cosh(x/\ell_D)}{\cosh[\ell/(2\ell_D)]}, \tag{12}$$

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-nonconserving 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-nonconserving 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 timescales of few picoseconds [30].

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

$$\tilde{\rho}(x) = \hat{\rho}_{ns} - \frac{|\rho_\delta\rangle\langle\rho_\delta|}{\rho_{\delta\delta}}[1 - \alpha(\ell, x)]. \quad (13)$$

According to the discussion above, from Eq. (13) we can define the unipolar and ambipolar resistivity matrices as

$$\hat{\rho}_{\text{uni}} \equiv \lim_{\alpha(\ell,x)\to 1} \tilde{\rho}(x) = \hat{\rho}_{ns},$$

$$\hat{\rho}_{\text{ambi}} \equiv \lim_{\alpha(\ell,x)\to 0} \tilde{\rho}(x) = \hat{\rho}_{ns} - \frac{|\rho_\delta\rangle\langle\rho_\delta|}{\rho_{\delta\delta}}. \quad (14)$$

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

$$\hat{\rho} = \hat{\rho}_{\text{uni}}\alpha(\ell) + \hat{\rho}_{\text{ambi}}[1 - \alpha(\ell)]. \quad (15)$$

We now observe that $\hat{\rho}_{\text{uni}}$ can be calculated, as shown in Ref. [17], by solving Boltzmann's kinetic equation within a two-mode approximation. As such, the electric and thermal resistivities calculated from it are always finite [17], and the resulting Lorenz ratio is *lower* than L_0. On the contrary, $\hat{\rho}_{\text{ambi}}$ corresponds to the resistivity matrix calculated by explicitly accounting for the conserved momentum mode [16,17]. Therefore, in the absence of momentum-nonconserving interactions, it exhibits a vanishing thermal resistivity ($\rho_{\text{th,ambi}} = 0$) when the semimetal is compensated [16,17]. Furthermore, as a consequence of momentum conservation, the determinant of the 3×3 matrix of Eq. (8), $\det(\hat{\rho}_{\text{ambi}})$, is always zero regardless of the doping level [17]. This in turn implies that the product $\rho_{\text{th,ambi}}\rho_{\text{el,ambi}} = 0$; i.e., either the electric or the thermal resistivity vanish, but not both at the same time since there is only one conservation law. Since $\rho_{\text{th,ambi}} = 0$ in a perfectly compensated system, we conclude that $\rho_{\text{el,ambi}}$ must be different from zero even in the absence of momentum-nonconserving interactions [16,17], and therefore the Lorenz

ratio calculated from $\hat{\rho}_{\text{ambi}}$ is *larger* than L_0. As shown in Ref. [16], momentum-nonconserving interactions such as quenched disorder allow for similar violations to be observed in a window of densities around perfect compensation. It is evident that the role of the function $\alpha(\ell)$ in Eq. (15) is to interpolate between these two opposite limits. In passing, we note that $\rho_{\text{el,ambi}}$ in a compensated system equals the electrical resistivity that could be naively calculated from $\hat{\rho}_{\text{uni}}$.

Note that we can easily add a contribution due to momentum-nonconserving processes $\hat{\rho}_D$ in series to $\hat{\rho}$ by replacing $\hat{\rho}_{\text{uni}} \to \hat{\rho}_{\text{uni}} + \hat{\rho}_D$ and $\hat{\rho}_{\text{ambi}} \to \hat{\rho}_{\text{ambi}} + \hat{\rho}_D$. Comparing Eq. (15) with the definition (7), we immediately identify the thermal resistivity

$$\rho_{\text{th}} = \alpha(\ell)\rho_{\text{th,uni}} + [1 - \alpha(\ell)]\rho_{\text{th,ambi}}. \quad (16)$$

It follows from what was said above that, in a compensated semimetal and in the absence of momentum-nonconserving interactions, $\rho_{\text{th}} = \alpha(\ell)\rho_{\text{th,uni}}$. Using Eqs. (8) and (15) we can also derive the Seebeck coefficient as

$$Q = \frac{\alpha(\ell)\rho_{\text{th,uni}}Q_{\text{uni}} + [1 - \alpha(\ell)]\rho_{\text{th,ambi}}Q_{\text{ambi}}}{\alpha(\ell)\rho_{\text{th,uni}} + [1 - \alpha(\ell)]\rho_{\text{th,ambi}}}, \quad (17)$$

where Q_{uni} and Q_{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., $\rho_{\text{el}} = \rho_{\text{el,ambi}}$. In fact, given that the boundary conditions do not 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 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 $\hat{\rho} = \hat{\rho}_{\text{ambi}}$. Therefore, the Lorenz ratio reads

$$L = \frac{\rho_{\text{el}}}{\alpha(\ell)\rho_{\text{th,uni}} + [1 - \alpha(\ell)]\rho_{\text{th,ambi}}}. \quad (18)$$

In the absence of interactions and in the unipolar regime, L tends to L_0, the value prescribed by the Wiedemann-Franz law. In a compensated semimetal and in the absence of momentum-nonconserving interactions, Eq. (18) reduces to the Lorenz ratio calculated within a two-mode approximation [17], which is always smaller than L_0, divided by $\alpha(\ell)$. The latter function becomes vanishingly small when $\ell \gg 2\ell_D$, i.e., when the ambipolar limit is approached, and therefore the true Lorenz ratio calculated from (18) can become *larger* than L_0.

IV. NUMERICAL RESULTS AND DISCUSSION

In this section we present numerical results for Eqs. (16)–(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-nonconserving interactions are completely neglected. These will be
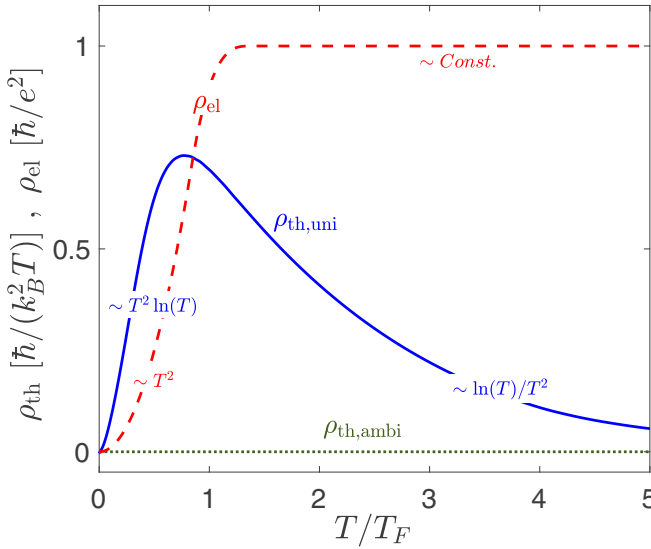
FIG. 3. Intrinsic unipolar and ambipolar thermal resistivities (as defined in the text) as well as the intrinsic electric resistivity as functions of temperature (scaled with T_F). In the presence of disorder the total electric and thermal resistivities, respectively, become $\rho_{el} + \rho_{el,dis}$ and $\rho_{th,uni/ambi} + \rho_{th,dis}$ (not shown in the figure), where $\rho_{el,dis}$ 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 noninteracting disordered system, we have assumed $\rho_{el,dis}/\rho_{th,dis} = \pi^2/3$ (in reduced units).

reintroduced 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, ρ_{el} and ρ_{th}, as functions of temperature. The thermal resistivities $\rho_{th,uni}$ and $\rho_{th,ambi}$, 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, $\rho_{th,uni} = 1/(2\kappa_e)$. We use the well-established kinetic-equation methods of Refs. [16,17] to calculate the intrinsic thermal conductivity κ_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 nonequilibrium distribution function [16] and find $\rho_{th,uni} \simeq I_{th}/(2D_{th}^2)$, where $D_{th} = \sum_{\bm{k}}(\partial f_{\bm{k}}/\partial \epsilon_{\bm{k}}) \bm{v}_{\bm{k}} \cdot \bm{v}_{\bm{k}}[(\epsilon_{\bm{k}} - \varepsilon_F)/k_B T]^2 \simeq 9\zeta(3)/(4\pi \hbar^2 \beta)$ is the thermal Drude weight, while I_{th} is the Coulomb collision integral projected onto

the thermal channel. For a system with a single parabolic band, the velocity $v_k \sim \bm{k}$ coincides with the momentum zero mode, and therefore only the thermal moment of the collision integral, I_{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_{th} to be

$$I_{th} = -\frac{1}{4\pi (k_B T)^3} \sum_{\bm{q}} \int_{-\infty}^{\infty} d\omega \frac{|V(\bm{q},\omega)|^2}{\sinh^2(\hbar\omega/2k_B T)}$$
$$\times [(\operatorname{Im} \Pi_1)^2 - \operatorname{Im} \Pi_0 \operatorname{Im} \Pi_2], \quad (19)$$

where $V(\bm{q},\omega) = v_q/|1 - v_q \Pi_0(\bm{q},\omega)|$ 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

$$\Pi_n(\mathbf{q},\omega) = 2 \sum_{\bm{k}} \frac{(\tilde{\epsilon}_k v_k - \tilde{\epsilon}_{k+q} v_{k+q})^n (f_k^0 - f_{k+q}^0)}{\epsilon_k - \epsilon_{k+q} + \hbar\omega + i0^+}, \quad (20)$$

where $\tilde{\epsilon}_k = \epsilon_k - \varepsilon_F$ is the band energy measured from the Fermi energy. In the degenerate Fermi liquid regime ($T \ll T_F$), Π_0 is the well-known zero-temperature 2D Lindhard function [31] and we find that for $\hbar\omega \ll \varepsilon_F$, $[\Pi_1^2 - \Pi_0 \Pi_2] \sim \omega^6$, resulting in $I_{th} \sim T^4$. As $D_{th}(T \ll T_F) \sim T$ we find

$$\rho_{th,uni}(T \ll T_F) \sim T^2 \ln(T). \quad (21)$$

This behavior of the thermal resistivity is consistent with previous results obtained for degenerate Fermi liquid graphene [22]. We find that $\rho_{th,uni}$ peaks around $T \simeq T_F$ and decreases as $\sim \ln(T)/T^2$ for $T \gg T_F$ (i.e., in the nondegenerate regime); see blue solid curve in Fig. 3.

Next, we study the the temperature dependence of the intrinsic electric resistivity ρ_{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 ρ_{el} (in reduced units) is given by

$$\rho_{el} \sim \frac{1}{D_{el}(T)\tau_{eh}(T)}, \quad (22)$$

where $\tau_{eh}(T)$ is the electron-hole scattering rate and $D_{el}(T)$ the Drude weight in the electric channel, which is defined as $D_{el} = \sum_{\bm{k}} \bm{v}_{\bm{k}}^2 (\partial f_{\bm{k}}/\partial \epsilon_{\bm{k}})$. In the Planckian regime the maximum scattering rate allowed by the energy-time uncertainty principle [32] is $1/\tau_{eh}(T) \sim k_B T/\hbar$ while in the Fermi liquid regime $1/\tau_{eh}(T) \sim (k_B T)^2/\hbar\varepsilon_F$. Similarly, in the Planckian regime the Drude weight $D_{el} \propto k_B T$, whereas in the Fermi liquid regime it is independent of temperature. Hence, the intrinsic resistivity ρ_{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-nonconserving interactions and discuss the calculation of the Lorenz ratio of Eq. (18).
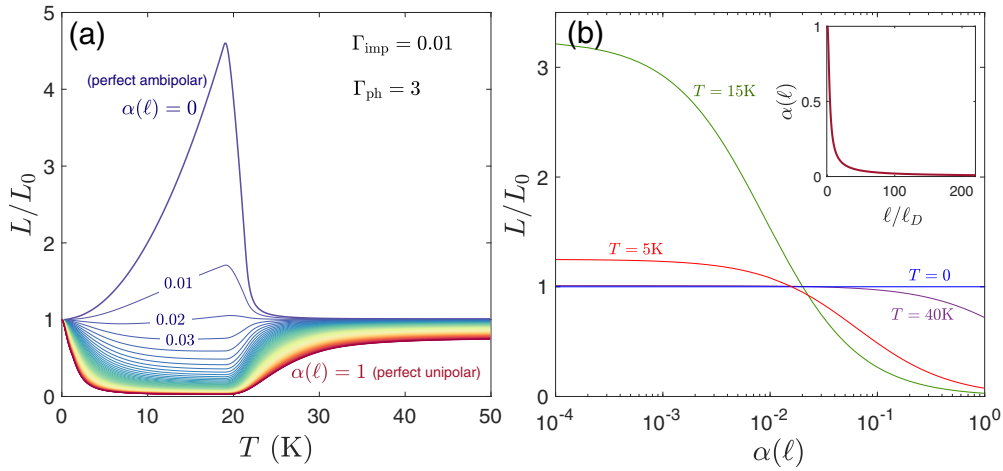
FIG. 4. Lorenz ratio as a function of (a) temperature for different values of $\alpha(\ell)$ from 0 (perfect ambipolar regime) to 1 (perfect unipolar regime) with a step of 0.01 and (b) as a function of $\alpha(\ell)$ (in logarithmic scale) for different temperatures as labeled. The strength of the charge impurity- and phonon-limited resistivities were defined through the hydrodynamicity parameters $\Gamma^2_{\mathrm{imp}} = \rho_{\mathrm{el,imp}}/\rho_{\mathrm{el}}(T \gg T_F)$ and $\Gamma^2_{\mathrm{ph}} = \rho_{\mathrm{el,ph}}/\rho_{\mathrm{el}}(T \gg T_F)$, 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 $T_{\mathrm{ph}} = 25$ 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-nonconserving scattering is necessary to regularize results in the ambipolar limit: as is clear from Fig. 3 and the definition (18), the ratio between the electrical and thermal resistivity ($\rho_{\mathrm{th,ambi}}$) 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 $\rho_{\mathrm{el,imp}} = m^*/(n\tau_{\mathrm{imp}})$, where m^* and n are the electron effective mass and density, respectively. The impurity scattering rate $1/\tau_{\mathrm{imp}}$ 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, $\rho_{\mathrm{el,imp}}$, is independent of T.

As the temperature increases, the resistivity due to collisions with phonons $\rho_{\mathrm{el,ph}}$ becomes the dominant contribution to the total electric resistivity. Above the Bloch-Grüneisen temperature, ρ_{ph} increases linearly with T [33]. We therefore posit the following model for the momentum-nonconserving scattering: $\rho_{\mathrm{el,dis}}(T \lesssim T_{\mathrm{ph}}) \simeq \rho_{\mathrm{el,imp}} \sim$ constant and $\rho_{\mathrm{el,dis}}(T \gtrsim T_{\mathrm{ph}}) \simeq \rho_{\mathrm{el,ph}} \sim T$, where T_{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-nonconserving (electron-impurity or electron-phonon) processes are taken into account, we obtain in particular that the thermal resistivity of impurities in reduced units is $\rho_{\mathrm{th,imp}} = (3/\pi^2)\rho_{\mathrm{el,imp}}$.

Figures 4(a) and 4(b) 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 $T_{\mathrm{ph}} \simeq 25$ K. We determine the strength of the charge impurity as well as phonon resistivities, respectively, through the hydrodynamicity parameters $\Gamma^2_{\mathrm{imp}} = \rho_{\mathrm{imp}}/\rho_{\mathrm{el}}(T \gg T_F)$ and $\Gamma^2_{\mathrm{ph}} = \rho_{\mathrm{ph}}/\rho_{\mathrm{el}}(T \gg T_F)$ (i.e., the ratio of the charge impurity/phonon resistivities to the intrinsic Coulomb resistivity in the nondegenerate regime, $T \gg T_F$). For the results in Figs. 4(a) and 4(b) we have taken $\Gamma_{\mathrm{imp}} = 0.01$ and $\Gamma_{\mathrm{ph}} = 3$.

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. 1(b)], while the unipolar one occurs in compensated semimetals such as WP$_2$ [see Fig. 1(c)]. 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.

Figure 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 $T \gtrsim T_{\mathrm{ph}}$, $L/L_0 \to 1$ resulting from the disorder-limited transport in these regimes, i.e., impurity-dominated at $T \to 0$ and phonon-dominated at $T \gtrsim T_{\mathrm{ph}}$. In the hydrodynamic regime ($0 \lesssim T \lesssim T_{\mathrm{ph}}$), one can tune the Wiedemann-Franz ratio from an enhancement when $\ell \gg \ell_D$ (bipolar condition) to a reduction when $\ell \ll \ell_D$ (unipolar condition).

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-nonconserving
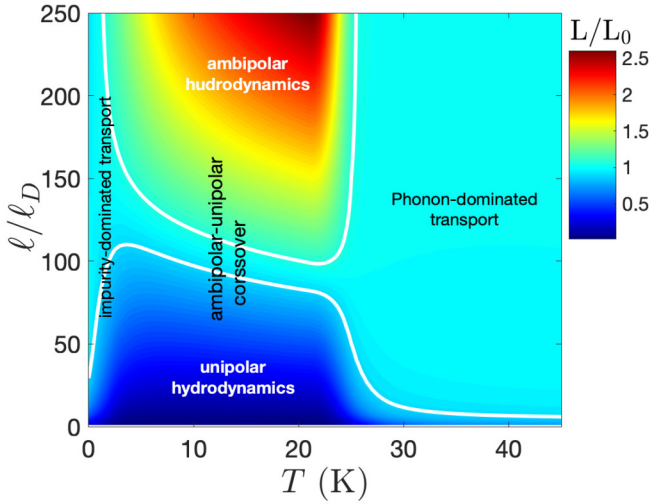
FIG. 5. 2D plot of the Lorenz ratio as a function of ℓ/ℓ_D (sample length scaled with the diffusion length) and T. Different transport regimes are indicated on the figure. The onset temperature of phonons $T_{ph} = 25$ K and the hydrodynamicity parameters $\Gamma_{imp} = 0.01$ and $\Gamma_{ph} = 3$ were 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.

processes are absent, $\rho_{\mathrm{th,ambi}} = 0$. This in turn implies that $\rho_{\mathrm{th}} = \alpha(\ell)\rho_{\mathrm{th,uni}}$ and $Q = Q_{\mathrm{uni}}$. In this case, while the thermal resistivity depends on the system size, the Seebeck coefficient is independent of it.

V. 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 liquids 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 Wiedemann-Franz 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.

APPENDIX: 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 downfolded 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 Sec. III. 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 downfolding process.

The derivation here parallels that given in Sec. III. Fields and currents are now related by the 4×4 resistivity matrix $\hat{\rho}$. Explicitly,

$$\begin{pmatrix} F_0 \\ |F_{ns\delta}\rangle \end{pmatrix} = \begin{pmatrix} \rho_{00} & \langle\rho_0| \\ |\rho_0\rangle & \hat{\rho}_{ns\delta} \end{pmatrix} \begin{pmatrix} j_0 \\ |j_{ns\delta}\rangle \end{pmatrix}. \qquad (A1)$$

Here $|j_{ns\delta}\rangle = {}^t(j_n, j_s, j_\delta)$ and $|F_{ns\delta}\rangle = {}^t(-eE, -k_B\partial_x T, F_\delta)$, $\hat{\rho}_{ns\delta}$ is a 3×3 block, and $|\rho_0\rangle$ is a three-component vector.

The determinant of the 4×4 matrix in Eq. (A1) is

$$\det(\hat{\rho}) = \det(\hat{\rho}_{ns\delta})\big[\rho_{00} - \langle\rho_0|\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle\big], \qquad (A2)$$

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. (A1) 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

$$\rho_{00} = \langle\rho_0|\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle. \qquad (A3)$$

The property (A3) plays a pivotal role in the following proof.

To downfold the four-mode theory of Eq. (A1) 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. (A1) implies that

$$j_0 = -\frac{\langle\rho_0|j_{ns\delta}\rangle}{\rho_{00}}. \qquad (A4)$$

When this relation is substituted into the last line of Eq. (A1) we get

$$|F_{ns\delta}\rangle = \left[\hat{\rho}_{ns\delta} - \frac{|\rho_0\rangle\langle\rho_0|}{\rho_{00}}\right]|j_{ns\delta}\rangle. \qquad (A5)$$

The matrix in square brackets on the right-hand side of Eq. (A5) 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 $\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle$. Indeed, using the property (A3), we have

$$|F_{ns\delta}\rangle = \left[\hat{\rho}_{ns\delta} - \frac{|\rho_0\rangle\langle\rho_0|}{\rho_{00}}\right]\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle$$
$$= |\rho_0\rangle - \frac{|\rho_0\rangle\langle\rho_0|\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle}{\langle\rho_0|\hat{\rho}_{ns\delta}^{-1}|\rho_0\rangle} = 0, \qquad (A6)$$

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.

[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, Science **351**, 1055 (2016).

[2] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Phys. Rev. Lett. **116**, 136802 (2016).

[3] 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).

[4] A. Principi, G. Vignale, M. Carrega, and M. Polini, Phys. Rev. B **93**, 125410 (2016).

[5] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schütt, and A. D. Mirlin, Phys. Rev. B **91**, 035414 (2015).

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

[7] L. Fritz, J. Schmalian, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 085416 (2008).

[8] M. Müller, L. Fritz, and S. Sachdev, Phys. Rev. B **78**, 115406 (2008).

[9] H.-Y. Xie and M. S. Foster, Phys. Rev. B **93**, 195103 (2016).

[10] D. Svintsov, V. Vyurkov, V. Ryzhii, and T. Otsuji, Phys. Rev. B **88**, 245444 (2013).

[11] D. Svintsov, Phys. Rev. B **97**, 121405(R) (2018).

[12] D. Y. H. Ho, I. Yudhistira, N. Chakraborty, and S. Adam, Phys. Rev. B **97**, 121404(R) (2018).

[13] A. Principi and G. Vignale, Phys. Rev. B **91**, 205423 (2015).

[14] A. Lucas and S. Das Sarma, Phys. Rev. B **97**, 245128 (2018).

[15] A. Lucas and S. A. Hartnoll, Phys. Rev. B **97**, 045105 (2018).

[16] M. Zarenia, A. Principi, and G. Vignale, 2D Mater. **6**, 035024 (2019).

[17] M. Zarenia, T. B. Smith, A. Principi, and G. Vignale, Phys. Rev. B **99**, 161407(R) (2019).

[18] M. Zarenia, I. Yudhistira, S. Adam, and G. Vignale, Phys. Rev. B **101**, 045421 (2020).

[19] M. Zarenia, S. Adam, and G. Vignale, Phys. Rev. Research **2**, 023391 (2020).

[20] 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 Mater. **3**, 64 (2018).

[21] J. Gooth, F. Menges, N. Kumar, V. Süß, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Nat. Commun. **9**, 4093 (2018).

[22] A. Principi and G. Vignale, Phys. Rev. Lett. **115**, 056603 (2015).

[23] A. Jaoui, B. Fauqué, and K. Behnia, arXiv:2006.01861.

[24] C. F. Gallo, B. S. Chandrasekhar, and P. H. Sutter, J. Appl. Phys. **34**, 144 (1963).

[25] C. Uher and H. J. Goldsmid, Phys. Stat. Sol. **65**, 765 (1974).

[26] S. Li and D. L. Maslov, Phys. Rev. B **98**, 245134 (2018).

[27] L. Pitaevskii and E. Lifshitz, *Physical Kinetics* (Elsevier Science, Oxford, 2012).

[28] D. Pines and P. Nozières, *The Theory of Quantum Liquids: Normal Fermi Liquids* (W. A. Benjamin, Boca Raton, 1966).

[29] B. G. Streetman, *Solid State Electronic Devices* (Prentice-Hall, Inc., 1990).

[30] Y. Chen, Y. Li, Y. Zhao, H. Zhou, and H. Zhu, Sci. Adv. **5**, eaax9958 (2019).

[31] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005).

[32] S. A. Hartnoll, Nat. Phys. **11**, 54 (2015).

[33] S. Kasap, C. Koughia, and H. E. Ruda, *Electrical Conduction in Metals and Semiconductors* (Springer International Publishing, Cham, 2017).