Boltzmann approach to high-order transport: The nonlinear and nonlocal responses

M. Battiato,[1,*] V. Zlatic,[2] and K. Held[1]

[1]*Institute of Solid State Physics, Vienna University of Technology, 1040 Vienna, Austria*
[2]*Institute of Physics, 10001 Zagreb, Croatia*

The phenomenological textbook equations for charge and heat transport are extensively used in a number of fields ranging from semiconductor devices to thermoelectricity. We provide a rigorous derivation of transport equations by solving the Boltzmann equation in the relaxation-time approximation and show that the currents can be rigorously represented by an expansion in terms of the "driving forces". Besides the linear and nonlinear response to the electric field, the gradient of the chemical potential and temperature, there are also terms that give the response to the higher-order derivatives of the potentials. These nonlocal responses might play an important role for some materials and/or under certain conditions, such as extreme miniaturization. Our solution provides the general solution of the Boltzmann equation in the relaxation-time approximation (or, equivalently, the particular solution for the specific boundary conditions). It differs from the Hilbert expansion, which provides only one of infinitely many solutions which may or may not satisfy the required boundary conditions.

I. INTRODUCTION

The phenomenological transport equations for the charge and heat currents are at the core of the description of electric and electronic devices of any type. These equations relate the local charge and heat current densities, $\mathbf{J}(\mathbf{r})$ and $\mathbf{J}_{\mathcal{E}}(\mathbf{r})$, to the local thermodynamic forces given by the gradients of the electrical potential ϕ, chemical potential μ, and temperature T. Often, they are written as

$$\mathbf{J}(\mathbf{r}) = \sigma\left(\mathbf{E} + \frac{\nabla\mu}{e}\right) - \sigma\alpha\nabla T + \sigma^{[\mathbf{E}^2]}\left(\mathbf{E} + \frac{\nabla\mu}{e}\right)^2, \quad (1)$$

$$\mathbf{J}_{\mathcal{E}}(\mathbf{r}) = \sigma\alpha T\left(\mathbf{E} + \frac{\nabla\mu}{e}\right) - (\kappa + \sigma\alpha^2 T)\nabla T, \quad (2)$$

where $\mathbf{E} = -\nabla\phi$ is the local electric field, and the conductivity σ, the Seebeck coefficient α, and the thermal conductivity κ are the position-dependent transport coefficients describing the linear part of the response. In some cases, the experiments indicate the presence of the nonlinear response, but, of all the possible terms, we wrote down only the one that is proportional to the square of the electrochemical force, with $\sigma^{[\mathbf{E}^2]}$ as the transport coefficient.

Within the linear response theory, Onsager [1], Kubo [2], and Luttinger [3] explained how to relate the coefficients of the driving fields to microscopic quantities, but the microscopic content of $\sigma^{[\mathbf{E}^2]}$ is less clear. Our aim, within the semiclassical Boltzmann approach, is to relate $\sigma^{[\mathbf{E}^2]}$ and similar terms in the expansion of the current densities to the energy dispersion and the scattering matrix of the electrons. However, as shown in detail below, the expansion should contain all the powers of the thermodynamic forces and their derivatives, i.e., a consistent theory should include all the driving forces up to a given order. Beyond the linear order, the current response is usually very small but, under specific conditions, the nonlinear driving forces can play a role; for instance, when the forces are large or when they exhibit large variations across the sample.

In this paper, we show how to obtain a systematic expansion of the current densities in terms of the driving forces. We find that Eqs. (1) and (2) are incomplete and that nonlocal terms proportional to the higher-order derivatives of the potentials, such as $\nabla\mathbf{E}$, $\partial^2\mu/\partial\mathbf{r}^2$, and $\partial^2 T/\partial\mathbf{r}^2$, have to be included as well. This has an obvious fundamental relevance. While some nonlocal effects have been studied so far within different approximations and approaches (as, for instance, the anomalous skin effect [4,5]), there is no systematic and broad classification of nonlocal effects, consistent with the definition of other effects such as conductivity, Seebeck, Hall, and so on. This consequently leads to neglecting effects which can be important in the description of devices, as, for, instance the charge current proportional to $\nabla\mathbf{E}$ in charged bulk regions. However, our results are also of a more direct and specific interest for the treatment of modern semiconductor devices. The miniaturization requires ever smaller components, with smaller and smaller active regions [6–9]. In these regions, the thermodynamic potentials charge enormously over small distances and their higher-order derivatives gain in importance. Thus, even if the prefactors of these terms are very small, their overall contribution to the total current could be significant and they could have an impact on the performance of the device.

Similarly, modern thermoelectric devices [10–14], designed so as to optimize the efficiency under given operating conditions [15–18], are often heterogeneous and have a nonlinear distribution of temperature and chemical potential. In that case, additional terms are needed to describe the charge and energy current densities. An advance along these line requires a quantum mechanical engineering that is hardly possible without an insight from theory [19–21].

The theoretical analysis is usually performed in several steps. First, the electronic structure is calculated by density functional theory and the transport coefficients are obtained by the linearized Boltzmann equation [22] because the Kubo approach is most often too difficult to use for real materials. Even the Boltzmann approach requires a number of simplifications, as described in the classic textbooks [23,24]. Once the transport coefficients are known, the currents given by Eqs. (1) and (2) can be substituted in the continuity equations

———————

*marco.battiato@ifp.tuwien.ac.at

for the charge and energy conservation. Given the appropriate boundary conditions, these equations provide, together with the Poisson equation for the electric field, the temperature and electrical and chemical potential at every macroscopic point of the sample. Thus, one can find the operating conditions, engineer the right composition of the material, and optimize the overall efficiency of the device.

Using the semiclassical Boltzmann theory and the relaxation-time approximation (RTA), the present paper formalizes the procedure outlined above and shows how to obtain, in a systematic and rigorous way, the transport equations and transport coefficients of inhomogeneous samples. Expanding the solution of the Boltzmann equation (BE) in terms of the driving forces, we derive the terms beyond the linear response, reproduce the terms in Eqs. (1) and (2), and show that additional terms arise. We also show that the effect of the microscopic boundary conditions can be neglected when the size of the system is large with respect to the diffusion length $l = \tau v_{\mathbf{k}}$. Here, τ is the scattering relaxation time, $v_{\mathbf{k}}$ is the group velocity of the electron wave packet, and we are referring to the microscopic boundary conditions for the BE, which define the momentum space distribution of incoming electrons at the boundary.

From the mathematical point of view, the proposed expansion has a major advantage over other expansions that are commonly used for the BE, such as the Hilbert [25] and Chapman-Enskog expansion [26–29]. These expansions disregard the boundary conditions, i.e., they expand just one of infinitely many possible solutions. In general, the solution generated by these expansions will not satisfy the boundary conditions of the real problem. The expansion proposed in this paper yields, within the RTA, the general solution of the BE. This is then used to find the particular solution satisfying the required boundary conditions. Thus, we can deal in a consistent way not only with the response to the higher-order driving forces, but we can also describe the transport properties in the vicinity of the boundary (for instance, the anomalous skin effect [4,5]). For simplicity, the present paper is restricted to the static case and it addresses in more detail only the former effects.

The paper is organized as follows. Section II introduces the Boltzmann and Poisson equations and the relaxation-time approximation, which leads to two further equations required by the energy and particle conservation. For a given boundary condition, these four coupled equations determine the electron distribution function, the temperature, the chemical potential, and the electrical field everywhere in the sample. Supplementarily, we show in Appendix A the equivalence of the equations for the charge and energy conservation to the continuity equations for the charge and heat current densities. In Sec. III, rather than solving all these equations simultaneously, we focus on the Boltzmann equation treating the temperature, chemical, and electrical potential as arbitrary known functions. The central part of the paper is the expansion of these arbitrary functions in Taylor series, which generates an expansion of the nonequilibrium electron distribution function in terms of the driving forces. The coefficients in that expansion satisfy coupled differential equations, as shown in Sec. III A. The problem of setting up the proper boundary conditions and solving these equations is discussed in Sec. III B. The

major advantage of our expansion over the Hilbert expansion is explained in Appendix B. In Sec. IV, we show that for macroscopic samples, a further approximation can be done which leads to major simplifications (this approximation is equivalent to neglecting the effects that are relevant only close to the surface). In Sec. V, we derive the generalized transport equations for the heat and charge currents in the bulk and find additional terms. Section VI provides one example which shows the influence of the thermodynamic forces on the behavior of materials: we analyze the depletion region in a metal-semiconductor junction. Section VII briefly discusses and summarizes our results.

II. TRANSPORT EQUATIONS

Bloch's quantum extension of the Boltzmann's theory derives the transport properties of a degenerate electron gas from the distribution function $g(t,\mathbf{r},\mathbf{k})$, where $\mathbf{r}$ and $\mathbf{k}$ are the coordinates of an electron at time t in the real and momentum space, respectively. For electrons moving in the presence of scalar and vector potentials $V(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, and confined to a single band, the distribution function satisfies the Boltzmann equation [22,23] (we neglect, for simplicity, the anomalous contribution to the velocity, which can easily be included in the BE [30,31]),

$$
\frac{\partial g}{\partial t} + \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar} \cdot \nabla_{\mathbf{r}} g
$$
$$
+ \frac{e}{\hbar}\left[\nabla_{\mathbf{r}} V + \frac{\partial \mathbf{A}}{\partial t} - \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar} \times (\nabla_{\mathbf{r}} \times \mathbf{A})\right] \cdot \nabla_{\mathbf{k}} g = \left(\frac{dg}{dt}\right)_{\text{col}}.
$$
(3)

Here, $\mathcal{E} = \mathcal{E}(\mathbf{k})$ is the energy dispersion provided by the band structure calculations, $e < 0$ is the electron charge, $\hbar$ is the Planck constant, and $(dg/dt)_{col}$ is the collision integral which describes the change of the distribution function due to the electron-electron (e-e) and, possibly, the electron-phonon (e-ph) scatterings. The extension to a multiband system is straightforward and it amounts to a summation over a band index [23,24]. The charge and current densities defined by $g(t,\mathbf{r},\mathbf{k})$ have to be compatible with the electrodynamic potentials $V(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, as required by the Maxwell equations. The exact solution of the Boltzmann-Maxwell system of equations satisfies all the conservation laws compatible with the invariance of the Hamiltonian with respect to the symmetry operations [32] and it determines completely the transport properties of the system. Unfortunately, except in the most simple cases, the presence of the collision integral makes the exact solution inaccessible.

In many applications, for example, when engineering an optimal material for a thermoelectric device, one tries to infer the transport properties from the available band structure data, neglecting the details of the relaxation mechanisms. In that case, the standard approach is to assume that the scattering drives the system towards equilibrium and to replace the collision integral by a simple expression,

$$
\left(\frac{dg}{dt}\right)_{\text{col}} \approx -\frac{g - g_0}{\tau},
$$
(4)

where $g_0(T,\mu,\mathcal{E})$ is the unperturbed local distribution function defined by local temperature $T(\mathbf{r})$ and local chemical potential $\mu(\mathbf{r})$. The assumption that the main effect of the scattering processes is the restoration of local thermodynamic equilibrium on the time scale given by τ defines the RTA of the Boltzmann equation [24]. Since we are interested in the transport properties of an electron fluid, we choose

$$g_0(\mathbf{r},\mathbf{k}) = f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k})) = \frac{1}{1 + e^{\frac{\mathcal{E}(\mathbf{k})-\mu(\mathbf{r})}{k_B T(\mathbf{r})}}}, \quad (5)$$

where k_B is the Boltzmann constant and f_{FD} is the Fermi-Dirac distribution.

In what follows, we consider the transport properties in a stationary state, such that $\partial g/\partial t = 0$, and solve rigorously the static Boltzmann equation using the RTA. We take into account the electric field $\mathbf{E}(\mathbf{r}) = -\nabla V(\mathbf{r})$ but neglect, for simplicity, the magnetic field. Thus, we replace the integrodifferential equation (3) by a generalized drift-reaction (convection-reaction) equation,

$$\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E} \cdot \nabla_{\mathbf{r}}g - \frac{e}{\hbar}\mathbf{E} \cdot \nabla_{\mathbf{k}}g = -\frac{g - f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k}))}{\tau}. \quad (6)$$

The relaxation time τ is treated either as a free parameter, which provides the best fit to the experimental data, or it is calculated in the perturbation theory [23].

Unlike the exact solution of Eq. (3), the solution of Eq. (6) does not automatically satisfy the fundamental conservation laws, such as the particle number and the energy conservation. To make the RTA physically acceptable, we enforce the local particle and energy conservation by constraining the functions $T = T(\mathbf{r})$ and $\mu = \mu(\mathbf{r})$. The conservation of the local particle density $n(\mathbf{r}) = \int g(\mathbf{r},\mathbf{k})d^3k$ follows from the requirement

$$\int g(\mathbf{r},\mathbf{k})d^3k = \int f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k}))d^3k, \quad (7)$$

while the conservation of the total-energy density of interacting electrons is enforced by the equation

$$\int \mathcal{E}(\mathbf{k}) \, g(\mathbf{r},\mathbf{k})d^3k = \int \mathcal{E}(\mathbf{k}) \, f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k}))d^3k. \quad (8)$$

If the electrons scatter on some additional degrees of freedom, like phonons, the scattering process changes their energy by $\Delta\epsilon_{e-ph}$ which has to be added to the right-hand side of Eq. (8). The consistency of the charge density and the electrical field is enforced by the Poisson equation,

$$\nabla \mathbf{E}(\mathbf{r},t) = \frac{e\,n(\mathbf{r}) + \rho_{\text{ion}}(\mathbf{r})}{\epsilon_0}, \quad (9)$$

where $\rho_{ion}(\mathbf{r})$ is the background charge that ensures the overall charge neutrality. The self-consistent solution of Eqs. (6)–(9) provides $g(\mathbf{r},\mathbf{k})$, $\mathbf{E}(\mathbf{r})$, $T(\mathbf{r})$, and $\mu(\mathbf{r})$ at every point in the sample.

As shown in Appendix A, the conservation of charge and energy implies the continuity equations for the charge and energy current densities. In a stationary state (and in the absence of electron-phonon scatterings), the current densities

satisfy

$$\nabla \cdot \mathbf{J}(\mathbf{r}) = 0, \quad (10)$$

$$\nabla \cdot \mathbf{J}_{\mathcal{E}}(\mathbf{r}) = W, \quad (11)$$

where the charge and energy current densities are

$$\mathbf{J}(\mathbf{r}) = \frac{e}{\hbar}\int \nabla_{\mathbf{k}}\mathcal{E} \, g(\mathbf{r},\mathbf{k})d^3k, \quad (12)$$

$$\mathbf{J}_{\mathcal{E}}(\mathbf{r}) = \frac{e}{\hbar}\int \nabla_{\mathbf{k}}\mathcal{E} \, \mathcal{E} \, g(\mathbf{r},\mathbf{k})d^3k, \quad (13)$$

and W is the work done by the applied electric field per unit time,

$$W(\mathbf{r}) = \frac{e}{\hbar}\int \nabla_{\mathbf{k}}\mathcal{E} \cdot \mathbf{E}(\mathbf{r}) \, g(\mathbf{r},\mathbf{k}) \, d^3k = \mathbf{J} \cdot \mathbf{E}. \quad (14)$$

Thus, finding the electron distribution function in the RTA implies solving the Boltzmann equation (6) for $g(\mathbf{r},\mathbf{k})$, together with the continuity equations (10) and (11) and the Poisson equation (9). Equivalently, one can solve the Boltzmann equation (6) for $g(\mathbf{r},\mathbf{k})$ and Eqs. (9)–(13) for five unknown quantities $\mathbf{J}(\mathbf{r})$, $\mathbf{J}_{\mathcal{E}}(\mathbf{r})$, $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$.

In general, the above equations have infinitely many solutions and the physically relevant one is defined by the specific boundary conditions which provide $g(\mathbf{r}_B,\mathbf{k}) = g_B(\mathbf{r}_B,\mathbf{k})$ at every point $\mathbf{r}_B$ of the boundary. Note the difference between these detailed, microscopic boundary conditions and the one which specifies just the macroscopic quantities, such as temperature and electrical and chemical potentials, at the interfaces. The microscopic boundary conditions for the Boltzmann equation specify the momentum distribution of the electrons coming from the neighboring material and also take into account the reflection and scattering of the incoming electrons at the interface. In addition to the temperature of the injection, they should provide, for example, detailed information on the band structure of the neighboring material, the $\mathbf{k}$-dependent injection probability or reflectivity. The construction of boundary conditions that determine a specific physical situation at the interface is a nontrivial problem, which is not addressed further in this work. In the following, we simply assume the boundary conditions to be known.

III. EXPANSION OF THE BOLTZMANN DISTRIBUTION FUNCTION IN TERMS OF THE GENERALIZED FORCES

Although the Boltzmann equation simplifies considerably within the RTA, solving Eq. (6) for $g(\mathbf{r},\mathbf{k})$, together with Eqs. (7)–(9) for $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$, is still a formidable task. To solve it, we integrate Eq. (6) for completely arbitrary functions $\mu(\cdot)$, $T(\cdot)$, and $E(\cdot)$, using an expansion of $g(\mathbf{r},\mathbf{k})$ in terms of the forces that arise out of equilibrium and to which the system responds by setting up the currents. Such approximations, based on physical arguments and irreversible thermodynamics, are often used but, here, we present a systematic expansion which clarifies the range of validity of the textbook solutions and identifies the different driving forces.

For a given microscopic boundary condition (BC), the solution of Eq. (6) has a unique value at every point $\{\mathbf{r},\mathbf{k}\}$ of the phase space, so that $g(\mathbf{r},\mathbf{k})$ is a functional g_F defined on

functions $\mu(\cdot)$, $T(\cdot)$, and $\mathbf{E}(\cdot)$, and the BC themselves. That is,

$$g(\mathbf{r},\mathbf{k}) = g_F[\mathbf{r},\mathbf{k},T(\cdot),\mu(\cdot),\mathbf{E}(\cdot),\mathrm{BC}]. \qquad (15)$$

Assuming $\mu(\cdot)$, $T(\cdot)$, and $\mathbf{E}(\cdot)$ are analytic functions, we expand them into Taylor series around point $\mathbf{r}$ and treat $g(\mathbf{r},\mathbf{k})$, without any loss of information [33], not as a functional but as a function of infinitely many variables,

$$g(\mathbf{r},\mathbf{k}) = \tilde{g}\bigg(\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r}),\mathbf{E}(\mathbf{r}),\nabla T(\mathbf{r}),\nabla\mu(\mathbf{r}),$$
$$\times \nabla\mathbf{E}(\mathbf{r}),\frac{\partial^2 T(\mathbf{r})}{\partial\mathbf{r}^2},\frac{\partial^2\mu(\mathbf{r})}{\partial\mathbf{r}^2},\frac{\partial^2\mathbf{E}(\mathbf{r})}{\partial\mathbf{r}^2},\dots,\mathrm{BC}\bigg). \qquad (16)$$

(Operator ∇ denotes $\nabla_\mathbf{r}$ whenever the function operated on depends solely on the position $\mathbf{r}$ and no ambiguity can arise.) Obviously, analytic functions $T(\cdot),\mu(\cdot),\mathbf{E}(\cdot)$ are completely defined by their values and the values of all their derivatives at any point of the sample. For example, $\mu(\cdot)$ is determined everywhere in its region of definition, if we provide, at point $\mathbf{r}$, the values $\mu(\mathbf{r})$, $\nabla\mu(\mathbf{r})$, $\partial^2\mu(\mathbf{r})/\partial\mathbf{r}^2$, $\partial^3\mu(\mathbf{r})/\partial\mathbf{r}^3$, etc. If one

is interested in the charge transport close to the interface, say, to model the Kapitza resistance [34], the interface defines a discontinuity and the Taylor extension cannot be used. To circumvent that problem, one can split the system into two halves, i.e., one to the left and one to the right of the boundary, and use separate Taylor expansions on each side of the interface. However, solving for such boundary conditions, including the Kapitza resistance, becomes cumbersome.

To proceed, we introduce the vector $\vec{\xi} = \{\xi_1,\xi_2,\xi_3,\dots\}$, where $\xi_1 = \mathbf{r}$, $\xi_2 = \mathbf{k}$, $\xi_3 = T$, $\xi_4 = \mu$, $\xi_5 = \mathbf{E}$, $\xi_6 = \nabla T$, etc. The components ξ_i, for $i \geqslant 5$, describe the driving forces to which the system responds. This notation indicates that even though $\tilde{g}(\vec{\xi})$ is defined on an infinite-dimensional vector space, only the values assumed by $\tilde{g}(\vec{\xi})$ on a small subspace of the whole definition space are physically relevant. In particular, we are interested in $\tilde{g}(\vec{\xi})$ on the hypersurface defined by $\xi_1 = \mathbf{r}$, $\xi_2 = \mathbf{k}$, $\xi_3 = T(\mathbf{r})$, $\xi_4 = \mu(\mathbf{r})$, $\xi_5 = \mathbf{E}(\mathbf{r})$, $\xi_6 = \nabla T(\mathbf{r})$, $\xi_7 = \nabla\mu(\mathbf{r})$, etc.

Next, we expand $\tilde{g}(\vec{\xi})$ in a Taylor series around the point $\vec{\xi}_0 = \{\xi_1,\xi_2,\xi_3,\xi_4,0,0,0,\dots\}$, i.e., we expand $\tilde{g}(\vec{\xi})$ with respect to the variables ξ_i around $\xi_i = 0$, for all $i \geqslant 5$. Thus, we write

$$\tilde{g}(\vec{\xi}) = g^{[0]}[\xi_1,\xi_2,\xi_3,\xi_4] + \delta g^{[\xi_5]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_5 + \delta g^{[\xi_6]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_6 + \delta g^{[\xi_7]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_7$$
$$+ \delta g^{[\xi_5^2]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_5^2 + \delta g^{[\xi_8]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_8 + \delta g^{[\xi_8^2]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_8^2 + \delta g^{[\xi_5\xi_6]}[\xi_1,\xi_2,\xi_3,\xi_4]\,\xi_5\xi_6 + \cdots, \qquad (17)$$

where the coefficients $\delta g^{[\alpha,\beta,\gamma,\dots]}[\xi_1,\xi_2,\xi_3,\xi_4]$ depend on the first four variables and the boundary conditions. (The explicit dependence on the boundary conditions has been omitted, for brevity.) In terms of the physically more transparent symbols, we have

$$g(\mathbf{r},\mathbf{k}) = g^{[0]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})] + \delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\mathbf{E}(\mathbf{r}) + \cdots + \delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\nabla T(\mathbf{r})$$
$$+ \delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\nabla\mu(\mathbf{r}) + \delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,(\mathbf{E}(\mathbf{r}))^2 + \cdots + \delta g^{[\partial^2 T/\partial\mathbf{r}^2]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\partial^2 T(\mathbf{r})/\partial\mathbf{r}^2 + \cdots$$
$$+ \delta g^{[(\partial^2 T/\partial\mathbf{r}^2)^2]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,(\partial^2 T(\mathbf{r})/\partial\mathbf{r}^2)^2 + \delta g^{[\mathbf{E},\nabla T]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\mathbf{E}(\mathbf{r})\nabla T(\mathbf{r}) + \cdots, \qquad (18)$$

where the (still unknown) coefficients $\delta g^{[E]}$, $\delta g^{[\nabla T]}$, $\delta g^{[\nabla\mu]}$, etc. describe the change in the distribution function due to the applied forces $\mathbf{E}$, ∇T, $\nabla\mu$, etc. These coefficients can be computed by substituting $\tilde{g}(\vec{\xi})$, given by Eq. (18), into the Boltzmann equation (6) and collecting the terms to order $\mathbf{E}(\mathbf{r})$, $\nabla T(\mathbf{r})$, $\nabla\mu(\mathbf{r})$, $\partial^2 T(\mathbf{r})/\partial\mathbf{r}^2$, $\mathbf{E}(\mathbf{r})^2$, $\mathbf{E}(\mathbf{r})\nabla T(\mathbf{r})$, $\mathbf{E}(\mathbf{r})\nabla\mathbf{E}(\mathbf{r})$, etc. The first term in Eq. (6) yields (we only show the first few terms)

$$\nabla_\mathbf{r} g(\mathbf{r},\mathbf{k}) = \nabla_\mathbf{r}\big\{g^{[0]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})] + \delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\mathbf{E}(\mathbf{r}) + \delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\nabla T(\mathbf{r}) + \delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,\nabla\mu(\mathbf{r})$$
$$+ \delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,[\mathbf{E}(\mathbf{r})]^2 + \delta g^{[\mathbf{E},\nabla\mu]}[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r})]\,[\mathbf{E}(\mathbf{r})]\nabla\mu + \cdots\big\}$$
$$= \nabla_\mathbf{r} g^{[0]}[\mathbf{r},\mathbf{k},T,\mu] + \frac{\partial g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial T}\nabla T + \frac{\partial g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial\mu}\nabla\mu$$
$$+ \nabla_\mathbf{r}\delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T,\mu]\,\mathbf{E} + \frac{\partial\delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial T}\mathbf{E}\,\nabla T + \frac{\partial\delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial\mu}\mathbf{E}\,\nabla\mu + \delta g^{[\mathbf{E}]}[\mathbf{r},\mathbf{k},T,\mu]\,\nabla\mathbf{E}$$
$$+ \nabla_\mathbf{r}\delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T,\mu]\,\nabla T + \frac{\partial\delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial T}(\nabla T)^2 + \frac{\partial\delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial\mu}\nabla T\,\nabla\mu + \delta g^{[\nabla T]}[\mathbf{r},\mathbf{k},T,\mu]\,\frac{\partial^2 T}{\partial\mathbf{r}^2}$$
$$+ \nabla_\mathbf{r}\delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T,\mu]\,\nabla\mu + \frac{\partial\delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial\mu}(\nabla\mu)^2 + \frac{\partial\delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial T}\nabla T\,\nabla\mu + \delta g^{[\nabla\mu]}[\mathbf{r},\mathbf{k},T,\mu]\,\frac{\partial^2\mu}{\partial\mathbf{r}^2}$$
$$+ \nabla_\mathbf{r}\delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T,\mu]\,\mathbf{E}^2 + \frac{\partial\delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial T}\nabla T\,\mathbf{E}^2 + \frac{\partial\delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T,\mu]}{\partial\mu}\mathbf{E}^2\,\nabla\mu + 2\delta g^{[\mathbf{E}^2]}[\mathbf{r},\mathbf{k},T,\mu]\,\mathbf{E}\,\nabla\mathbf{E}$$
$$+ \cdots, \qquad (19)$$

where the effect of $\nabla_\mathbf{r}$, operating on composite functions in the first equation, is computed using the usual rules for the derivatives of the composite function, while $\nabla_\mathbf{r} g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]$, etc. in the second equation denotes the derivative of $g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]$ with respect to the first variable only.

The terms obtained by substituting Eq. (18) in the second and the third term of the Boltzmann equation (6) are easily written down and are not reported separately. The sum of all three terms can be written as

$$
\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}g - \frac{e}{\hbar}\mathbf{E}\cdot\nabla_{\mathbf{k}}g + \frac{g - f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k}))}{\tau}
$$

$$
= \left(\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}g^{[0]} + \frac{g^{[0]} - f_{FD}(T,\mu,\mathcal{E})}{\tau}\right) + \left(\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}\delta g^{[\nabla T]} + \frac{1}{\hbar}\frac{\partial g^{[0]}}{\partial T}\nabla_{\mathbf{k}}\mathcal{E} + \frac{\delta g^{[\nabla T]}}{\tau}\right)\nabla T
$$

$$
+ \left(\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}\delta g^{[\nabla\mu]} + \frac{1}{\hbar}\frac{\partial g^{[0]}}{\partial\mu}\nabla_{\mathbf{k}}\mathcal{E} + \frac{\delta g^{[\nabla\mu]}}{\tau}\right)\nabla\mu + \left(\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}\delta g^{[\mathbf{E}]} - \frac{e}{\hbar}\mathbf{u_E}\cdot\nabla_{\mathbf{k}}g^{[0]} + \frac{\delta g^{[\mathbf{E}]}}{\tau}\right)\mathbf{E} + \cdots = 0, \quad (20)
$$

where $\mathbf{u_E}$ is the unit vector in the direction of the electric field and only the linear terms are shown because the structure of the higher-order terms is obvious. Since $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$ are arbitrary functions, the above equation can only be satisfied if all the brackets vanish. We therefore have to impose that all of the expressions within brackets in Eq. (20) have to vanish separately.

Thus, we reduced the Boltzmann equation to an infinite sequence of coupled differential equations which describe the change in the distribution function in response to the driving forces. Each equation specifies a particular response function $\delta g^{[\alpha]}[\mathbf{r},\mathbf{k},T,\mu]$ (where T and μ are treated as variables), which corresponds to a particular driving force α, and the differential operator in these equations is operating on the first variable of $\delta g^{\alpha}[\xi_1,\xi_2,\xi_3,\xi_4]$ only. The solution can be constructed sequentially, starting from the lowest order and specifying, for every equation, a particular boundary condition regarding the variable $\mathbf{r}$. The construction has to ensure that the sum of all the contributions yields $g(\mathbf{r},\mathbf{k})$ which satisfies the boundary condition imposed on the solution of Eq. (6) (more on this in Sec. III B).

A. Drift-reaction equations

We now discuss the zeroth- and the first-order distribution function defined by the expansion (20) and relate them to what is known from the literature. We also provide a few typical examples of the higher-order terms. The zeroth-order distribution function is obtained by setting to zero the first bracket in Eq. (20), which gives

$$
\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}g^{[0]} + \frac{g^{[0]} - f_{FD}(T,\mu,\mathcal{E})}{\tau} = 0, \quad (21)
$$

where $\nabla_{\mathbf{r}}g^{[0]}$ is again the derivative of $g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]$ with respect to its first variable. The spatial part of Eq. (21) is a convection-reaction equation and the solution requires the value of $g^{[0]}[\mathbf{r},\mathbf{k},T,\mu]$ on the boundary.

The change in the distribution function due to an applied electric field $\mathbf{E}$ is defined by the equation [see Eq. (20)]

$$
\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\cdot\nabla_{\mathbf{r}}\delta g^{[\mathbf{E}]} - \frac{e}{\hbar}\frac{\mathbf{E}}{|\mathbf{E}|}\cdot\nabla_{\mathbf{k}}g^{[0]} + \frac{\delta g^{[\mathbf{E}]}}{\tau} = 0. \quad (22)
$$

Similarly, the response to a thermal force is obtained by collecting all the first-order terms in ∇T, which gives

$$
\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\cdot\nabla_{\mathbf{r}}\delta g^{[\nabla T]} + \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\frac{\partial g^{[0]}}{\partial T} + \frac{\delta g^{[\nabla T]}}{\tau} = 0; \quad (23)
$$

the response to a diffusion force $\nabla\mu$ is given by

$$
\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\cdot\nabla_{\mathbf{r}}\delta g^{[\nabla\mu]} + \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\frac{\partial g^{[0]}}{\partial\mu} + \frac{\delta g^{[\nabla\mu]}}{\tau} = 0. \quad (24)
$$

Equations (21)–(24), without the first term, yield the linear corrections to the Boltzmann distribution function, which is the same as in most textbooks [23,24]. These approximate expressions also agree with the results obtained, for instance, by the Hilbert expansion of Eq. (6) (see Appendix B). At this stage, it is not obvious that $\nabla_{\mathbf{r}}\delta g^{[\alpha]}$ can be neglected but, in Sec. IV, we show that the rigorous solution of Eqs. (21)–(24) indeed assumes the textbook form sufficiently far away from the boundaries.

The higher-order response follows straightforwardly from the expansion (20) and yields the terms of two basic types. The first type describes the nonlinear response due to the higher powers of the gradients of potentials [such as $(\nabla\mu)^2$, $\mathbf{E}^2$, $(\nabla T)^2$, etc.]. Many such terms have previously been discussed in the literature [23,24]. The second-order response to the diffusion force, $(\nabla\mu)^2$, is defined by the equation

$$
\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\cdot\nabla_{\mathbf{r}}\delta g^{[(\nabla\mu)^2]} + \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\frac{\partial\delta g^{[\nabla\mu]}}{\partial\mu} + \frac{\delta g^{[(\nabla\mu)^2]}}{\tau} = 0, \quad (25)
$$

and the response to $(\nabla T)^2$ is similar.

The terms of the second type describe the response to the higher-order derivatives of the potentials [such as $\partial^2\mu(\mathbf{r})/\partial\mathbf{r}^2$ or ∇E, etc.] and all their powers. These terms can be comparable to the nonlinear terms of the same (and higher) order. For example, the response to the second derivative of the chemical potential $\partial^2\mu(\mathbf{r})/\partial\mathbf{r}^2$ reads

$$
\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\cdot\nabla_{\mathbf{r}}\delta g^{[\frac{\partial^2\mu}{\partial\mathbf{r}^2}]} + \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\delta g^{[\nabla\mu]} + \frac{\delta g^{[\frac{\partial^2\mu}{\partial\mathbf{r}^2}]}}{\tau} = 0. \quad (26)
$$

A multidimensional Taylor expansion also generates a large number of mixed terms. The equations for other higher-order responses [say, the one proportional to $(\nabla\mu)^3$ or E^3], the equations for the coefficients depending on the higher derivatives of the fields [such as $\partial^3\mu(\mathbf{r})/\partial\mathbf{r}^3$], or all the cross terms are obtained by straightforward but tedious calculations.

The distribution function is now obtained by successively solving Eqs. (21)–(23), etc. The structure of these equations is always the same and we can write, in the nth order,

$$
\nabla_{\xi_1}\delta g^{[\xi_n]}[\xi_1,\xi_2,\xi_3,\xi_4] + F_n(\delta g^{[\xi_0]},\delta g^{[\xi_6]},\ldots,\delta g^{[\xi_{n-1}]})
$$

$$
+ \frac{\delta g^{[\xi_n]}}{\tau} = 0, \quad (27)
$$

where $F_n(\delta g^{[\xi_0]}, \delta g^{[\xi_6]}, \ldots, \delta g^{[\xi_{n-1}]})$ is a known function obtained from the solution of the lower-order equations.

B. Expansion of the boundary conditions

The boundary conditions for Eq. (6), written as $g(\mathbf{r}_B, \mathbf{k}) = g_B(\mathbf{r}_B, \mathbf{k})$, define how the electrons are injected in the region under consideration. For instance, a surface subject to an electron flux from vacuum (as in the case of inverse photoemission) is subject to an electron current with an electronic distribution at the entrance that depends on the energy distribution of the injected electrons. Another interesting case is that of a current flowing from a semiconductor into a metal. The injected electrons do not have the same energy distribution as in the case when they are excited by an electric field inside the metal. In general, from a mathematical point of view, the population of electrons in the k space can be described at the boundary $\mathbf{r} = \mathbf{r}_B$ by any function of momentum, not necessarily by the Fermi-Dirac distribution or the nonequilibrium distribution which is giving rise to stationary currents.

A unique solution of Eqs. (21)–(23), etc. requires a boundary condition in every order. Since the differential operator in the drift-reaction equation for $\delta g^{[\xi_n]}$ is operating on the first argument of $\delta g^{[\xi_n]}[\mathbf{r}, \mathbf{k}, T, \mu]$, the most general form of the boundary conditions for the drift-reaction equations is

$$\delta g^{[\xi_n]}[\mathbf{r}_B, \mathbf{k}, T, \mu] = \delta g_B^{[\xi_n]}[\mathbf{r}_B, \mathbf{k}, T, \mu], \quad (28)$$

where $\delta g_B^{[\xi_n]}$ specifies the value of $\delta g^{[\xi_n]}$ at the boundary for any given $\mathbf{k}$, T, and μ.

Note that any choice of boundary conditions for the drift-reaction equations is acceptable, as long as the sum of all terms in Eq. (18) yields the correct boundary condition for the solution of the Boltzmann equation (6). That is, the boundary conditions for the drift-reaction equations have to satisfy the supplementary condition

$$g_B(\mathbf{r}_B, \mathbf{k}) = g_B^{[0]}[\mathbf{r}_B, \mathbf{k}, T(\mathbf{r}_B), \mu(\mathbf{r}_B)]$$
$$+ \sum \delta g_B^{[\xi_n]}[\mathbf{r}_B, \mathbf{k}, T(\mathbf{r}_B), \mu(\mathbf{r}_B)] \xi_n(\mathbf{r}_B). \quad (29)$$

As long as Eq. (29) holds, the sum of all terms in Eq. (18) gives the particular solution of Boltzmann equation (6) which satisfies the required boundary condition (assuming the series expansion converges).

This concludes the construction of the expansion of the solution of the BE in the RTA. We emphasize that our expansion of the distribution function takes into account the boundary conditions. This implies that our solution yields not just the response to the higher driving forces or the nonlocal effects in the bulk, but it can also describe the effects caused by the specific choice of the boundary conditions (for instance, our solution can be used to discuss the anomalous skin effect). In Appendix B, we show why this cannot be achieved by other, often used expansions, as for instance the Hilbert expansion.

IV. SOLUTION OF THE DRIFT-REACTION EQUATION IN A MACROSCOPIC SAMPLE

In the rest of this paper, we focus on the high-order responses and nonlocal effects in the bulk, leaving the description of surface effects to future work (by surface effects,

we mean the features of the solution that depend on the specific form of the boundary conditions). We will proceed in two steps. First we show that due to the dissipative nature of the scattering term in the RTA, a specific shape of the BC modifies the solution only close to the boundary. This implies that for a given functional form of $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$, any BC imposed on the BE leads to the same solution in the bulk.

For a very small device, the full set of the drift-reaction equations (21)–(23), etc. has to be solved for a given choice of the boundary conditions, subject to the supplementary condition (29). This raises the problem of the optimal distribution of the boundary conditions among various drift-reaction equations. Similar issues also arise in larger systems close to the physical boundary. However, these problems are not addressed here.

To understand the extent in which the details of the BC affect the solution far away from the boundary, we study the lowest-order drift-reaction equation (21) in one dimension. The solution can be written as

$$g^{[0]}[x, \mathbf{k}, T, \mu] = f_{FD}(T, \mu, \mathcal{E}(\mathbf{k})) + C(\mathbf{k}, T, \mu) e^{-\frac{\hbar x}{\tau \partial_{k_x} \mathcal{E}}}, \quad (30)$$

where $\mathbf{k}, T, \mu$ are arbitrary and the coefficient $C(\mathbf{k}, T, \mu)$ is defined by the BC at $x = 0$. Obviously, the effect of the boundary on the solution $g^{[0]}[x, \mathbf{k}, T, \mu]$ decreases exponentially with the distance from the boundary. The characteristic decay length is given by the mean free path, $l = \tau v_k$. For $x \gg l$, the effect of the boundary is obliterated by the scattering, so that the particular value of $g^{[0]}[x, \mathbf{k}, T, \mu]$ at $x = 0$ becomes irrelevant at distances which are much larger than the mean free path. This feature also holds in higher dimensions and for every term in the expansion.

Therefore, if we are interested only in the bulk solution, which holds far from the boundary, we are free to choose the BC as we like; the difference between the true particular solution and the one obtained for a different BC vanishes far away from the boundary. This implies that the solution of the Boltzmann equation in the bulk of the sample is completely determined by the local temperature, chemical potential, and electric field, regardless of the microscopic BC.

As regards the response of the entire sample, the larger the system, the less important the region close to the boundary. In a macroscopic device, an accurate treatment of microscopic boundary conditions gives only a very small correction to the response functions, but it dramatically increases the computational complexity. Hence, we choose the boundary condition so as to minimize the computational efforts; for large enough systems, the error of using such a solution is insignificant. Note the difference between the microscopic boundary conditions for the BE and the macroscopic boundary conditions for the continuity equations providing the thermodynamic variables T and μ. The conservation laws given by Eqs. (7) and (8) require T and μ to assume the boundary values specified by the reservoirs. The microscopic boundary conditions, imposed on the Boltzmann equation, provide information on the state of the electrons at the boundary.

If we decide to disregard the surface effects, we can choose the BC for Eq. (21) as $g^{[0]}[0, \mathbf{k}, T, \mu] = f_{FD}(T, \mu, \mathcal{E}(\mathbf{k}))$, which yields $C(\mathbf{k}, T, \mu) = 0$ and makes the function $g^{[0]}$ independent of the variable x. The differential operator in Eq. (21) can now

be dropped and the solution of the zeroth-order drift-reaction equation becomes

$$g^{[0]}[\mathbf{r},\mathbf{k},T,\mu] = f_{FD}(T,\mu,\mathcal{E}(\mathbf{k})) \qquad (31)$$

everywhere in the sample. This solution deviates from the exact one only very close to the boundary, where the exponential term cannot be neglected.

Similarly, if we choose the BC for Eq. (27) as

$$\delta g^{[\xi_n]}[\mathbf{r}_B,\mathbf{k},T,\mu] = -\tau\, F_n(\delta g^{[\xi_0]},\delta g^{[\xi_6]},\ldots,\delta g^{[\xi_{n-1}]}), \quad (32)$$

the nth-order solution can be computed analytically and it will be independent of ξ_1. The set of approximate equations obtained in such a way coincides with the equations generated by the Hilbert expansion in the static approximation (see Appendix B).

The approximate solution of the BE which works very well for bulk materials is obtained by summing up all the solutions of the drift-reaction equations. The ensuing distribution function reads

$$\begin{aligned}
g(\mathbf{r},\mathbf{k}) = &\, f_{FD}(T,\mu,\mathcal{E}(\mathbf{k})) - \frac{\tau\, e}{\hbar} \cdot \nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k})) \cdot E \\
&+ \frac{\tau}{\hbar}\frac{\mathcal{E}-\mu}{T}\nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k})) \cdot \nabla T \\
&+ \frac{\tau}{\hbar}\nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k})) \cdot \nabla \mu \\
&- \frac{\tau^2}{\hbar^2}\nabla_{\mathbf{k}}\mathcal{E}\,\frac{\partial}{\partial\mu}\nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k}))(\nabla\mu)^2 \\
&- \frac{\tau^2}{\hbar^2}\nabla_{\mathbf{k}}\mathcal{E}\,\nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k}))\frac{\partial^2\mu}{\partial\mathbf{r}^2} \\
&- \frac{\tau^2}{\hbar^2}\frac{\mathcal{E}-\mu}{T}\nabla_{\mathbf{k}}\mathcal{E}\,\nabla_{\mathbf{k}} f_{FD}(T,\mu,\mathcal{E}(\mathbf{k}))\frac{\partial^2 T}{\partial\mathbf{r}^2} + \cdots,
\end{aligned} \tag{33}$$

and it satisfies Eq. (6) but does not satisfy the original BC. However, the difference with respect to the full solution is exponentially small as soon as we move away from the boundary. Note that even though the terms $\delta g^{[\alpha]}[\mathbf{r},\mathbf{k},T,\mu]$ do not have an explicit $\mathbf{r}$ dependence far away from the boundary, they become position dependent once we substitute, for $T = T(\mathbf{r})$, $\mu = \mu(\mathbf{r})$, and $\mathbf{E} = \mathbf{E}(\mathbf{r})$, the functions obtained by self-consistently solving the continuity and Poisson equations (see Sec. V).

The first few terms in Eq. (33) coincide with the expressions for the Boltzmann distribution function in the presence of the well-known driving forces, i.e., we have reproduced, in a mathematically consistent way, the known textbook expressions [23,24]. However, we also have additional terms, such as the last two. These terms, together with similar, higher-order ones, are easy to overlook in the heuristic derivations that are often used to justify the first five terms. They only appear from the second order onwards and, therefore, usually give small corrections. However, the second-order effects are not always negligible [for instance, the fifth term in Eq. (33) is sometimes very important]. The last three terms can be important in inhomogeneous materials (say, multilayers) where the concentration and temperature vary rapidly across the sample. In that case, all the terms of the same order should be treated on the same footing, i.e., one should not neglect the

nonlocal forces proportional to the higher-order derivatives of the temperature, chemical potential, and electric field.

As shown in Appendix B, the expansion obtained by neglecting the differential operator in Eq. (27) is the same as the one generated from the time-independent Boltzmann equation by the Hilbert expansion [25,35], with the Knudsen number as the expansion parameter. [The Knudsen number is given by the ratio of the mean free path, or the mean free time between collision, to some characteristic length (or time) of the system.] However, unlike the Hilbert expansion, our method retains its validity close to the boundary, provided we calculate the distribution function in each order from the differential drift-reaction equations (27). In that case, the transport coefficients are not simply defined by local thermodynamic variables, but have an explicit position dependence.

V. MACROSCOPIC TRANSPORT EQUATIONS

In the previous section, we derived an approximate solution of the Boltzmann equation (6) for arbitrary functions $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $E(\mathbf{r})$. Substituting that solution in equations for the charge and energy conservation, Eqs. (7) and (8), and using the Poisson equation (9), we can find the physical functions $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $E(\mathbf{r})$.

Substituting the power series for $g(\mathbf{r},\mathbf{k})$ in Eqs. (12) and (13) yields the transport equations

$$\mathbf{J} = \mathbf{J}^{[\nabla T]} + \mathbf{J}^{[\nabla\mu]} + \mathbf{J}^{[E]} + \mathbf{J}^{[E^2]} + \mathbf{J}^{[\nabla E]} + \cdots, \quad (34)$$

$$\mathbf{J}_{\mathcal{E}} = \mathbf{J}_\epsilon^{[\nabla T]} + \mathbf{J}_\epsilon^{[\nabla\mu]} + \mathbf{J}_\epsilon^{[E]} + \mathbf{J}_\epsilon^{[E^2]} + \mathbf{J}_\epsilon^{[\nabla E]} + \cdots, \quad (35)$$

where $\mathbf{J}^{[\alpha]}$ is the charge current density due to the driving force ξ_α,

$$\mathbf{J}^{[\alpha]} = e\left(\int \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\delta g^{[\alpha]}(\mathbf{r},\mathbf{k})d^3k\right)\xi_\alpha = N_\alpha^J \xi_\alpha, \qquad (36)$$

$\mathbf{J}_{\mathcal{E}}^{[\alpha]}$ is the corresponding energy current density,

$$\mathbf{J}_{\mathcal{E}}^{[\alpha]} = \left(\int \mathcal{E}\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}\delta g^{[\alpha]}(\mathbf{r},\mathbf{k})d^3k\right)\xi_\alpha = N_\alpha^{\mathcal{E}} \xi_\alpha, \qquad (37)$$

and N_α^J and $N_\alpha^{\mathcal{E}}$ are the transport coefficients associated with the force ξ_α. Since the charge and energy conservation imply the continuity equations for the charge and energy currents, we can equivalently obtain $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $E(\mathbf{r})$ by solving the continuity equations $\nabla\cdot\mathbf{J}(\mathbf{r}) = 0$ and $\nabla\cdot\mathbf{J}_{\mathcal{E}}(\mathbf{r}) = \mathbf{J}(\mathbf{r})\cdot\mathbf{E}(\mathbf{r})$, together with the Poisson equation and the transport equations, Eqs. (34) and (35).

The above current densities and transport coefficients reproduce all the standard results for the response due to the known driving forces. For instance, the current which is first order in the electric field $\mathbf{J}^{[E]} = \sigma\mathbf{E}$ has the conductivity coefficient

$$\sigma = -\frac{\tau\, e^2}{\hbar^2}\int \nabla_{\mathbf{k}}\mathcal{E}\,\nabla_{\mathbf{k}} f_{FD}\, d^3k, \qquad (38)$$

which is the textbook result. The same agreement is found for the Seebeck coefficient S and the thermal conductivity κ. We now report a few higher-order terms generated by the expansion of the distribution function given by Eq. (18). The current due to the second power of the electric field is $\mathbf{J}^{[E^2]} =$

$\sigma^{[\mathbf{E}^2]}\mathbf{E}^2$, where the conductivity coefficient,

$$\sigma^{[\mathbf{E}^2]} = \frac{e^3\tau^2}{\hbar^3} \int \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \frac{\partial^2 f_{FD}}{\partial\mu^2} d^3k, \qquad (39)$$

describes the Boltzmann expression for the nonlinear response to an electric field. The current driven by $\nabla\mathbf{E}(\mathbf{r})$ is $\mathbf{J}^{[\nabla\mathbf{E}]} = \sigma^{[\nabla\mathbf{E}]}\nabla\mathbf{E}$, with the conductivity coefficient

$$\sigma^{[\nabla\mathbf{E}]} = \frac{e^2\tau^2}{\hbar^3} \int \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \frac{\partial f_{FD}}{\partial\mu} d^3k, \qquad (40)$$

while the current driven by $\partial^2 T/\partial r^2$ is $\mathbf{J}^{[\frac{\partial^2 T}{\partial r^2}]} = \alpha^{[\frac{\partial^2 T}{\partial r^2}]}\partial^2 T/\partial r^2$, with the thermal conductivity coefficient

$$\alpha^{[\frac{\partial^2 T}{\partial r^2}]} = \frac{e\tau^2}{\hbar^3} \int \frac{\mathcal{E}-\mu}{T} \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \nabla_\mathbf{k}\mathcal{E} \ \frac{\partial f_{FD}}{\partial\mu} d^3k. \qquad (41)$$

The other contributions to the currents, with the corresponding transport coefficients, can be calculated in the same way.

The first few terms for the bulk current density read

$$\mathbf{J} = \sigma\left(\mathbf{E} + \frac{1}{e}\nabla\mu\right) + \alpha\nabla T + \sigma^{[\mathbf{E}^2]}\mathbf{E}^2 + \frac{\sigma^{[\mathbf{E}^2]}}{e^2}(\nabla\mu)^2$$
$$+ \sigma^{[\nabla\mathbf{E}]}\nabla\mathbf{E} + \frac{\sigma^{[\nabla\mathbf{E}]}}{e}\frac{\partial^2\mu}{\partial\mathbf{r}^2} + \cdots, \qquad (42)$$

where the first term describes the response to an electrochemical force, the second term describes the Seebeck effect due to the thermal gradient, and the third and fourth terms give the second-order response to the gradients of the electrical and chemical potentials, while the remaining terms describe the nonlocal response. A similar expression can be written for the heat current. Thus, our expansion supplements the well-known steady-state macroscopic transport equations with additional terms, which are due to the higher powers of the thermodynamic forces and their spatial derivatives. The difference with respect to the usual textbook equations is the appearance of additional, higher-order terms.

The current in Eq. (42) does not have the zeroth-order term since it can be proven to vanish because of the periodicity of the band structure and the fact that the integrand for the zeroth-order current is an exact differential. Let us also note that even-order contributions to the current densities are obtained by integrating the odd powers of the velocity. Such terms can be finite if inversion and time-reversal symmetry are broken, e.g., for a ferromagnet and a lattice without inversion symmetry. Otherwise, $\mathcal{E}_\uparrow(-\mathbf{k}) = \mathcal{E}_\downarrow(\mathbf{k})$ so that there is no charge current, but possibly a spin current if inversion symmetry is broken.

VI. EXAMPLE: DEPLETION REGION IN SEMICONDUCTORS

To show the impact of the additional terms on the behavior of real devices, we now consider an example involving the terms proportional to $\partial^2\mu/\partial\mathbf{r}^2$ and $\nabla\mathbf{E}$. These terms are proportional to τ^2, so $\sigma^{[\nabla\mathbf{E}]}$ is usually small, but if the scattering lifetime is long or the derivatives of the potential are large, their contribution can be important. Taking the case that is familiar to most readers, we examine the width and the shape of the depletion region in a metal-semiconductor (M-S) junction, shown in Fig. 1(a). We are only interested in the qualitative features due to the additional terms, revealed by our treatment, so we neglect several effects that are relevant for real junctions, such as finite jumps in temperature and chemical potential or the formation of defects.

Let us, first, recapitulate the textbook approach to isothermal M-S junctions, ignoring the heat transport [36]. The junction is described by two equations:

$$J = \sigma(\mu(x))\left[\mathbf{E}(x) - \frac{\mu'(x)}{e}\right] = \text{const.}, \qquad (43)$$

$$\epsilon_0\mathbf{E}'(x) = \rho(\mu(x)\mathbf{0}), \qquad (44)$$

where the first one is the transport equation (42) and the second one is the Poisson equation [Eq. (9)]. As shown below, the conductivity σ and the charge ρ depend explicitly on the chemical potential, so that the solution of these coupled equations yields, for a constant current density J, the spatial profile of the electric field and the chemical potential. The conductivity is calculated from Eq. (38) and to obtain an explicit expression for $\sigma(x)$ and $\rho(x)$, we consider an n-type semiconductor with a parabolic band, $\mathcal{E}(k) = \hbar^2 k^2/2m$. We also assume that the chemical potential is below the bottom of the conduction band and approximate the Fermi-Dirac
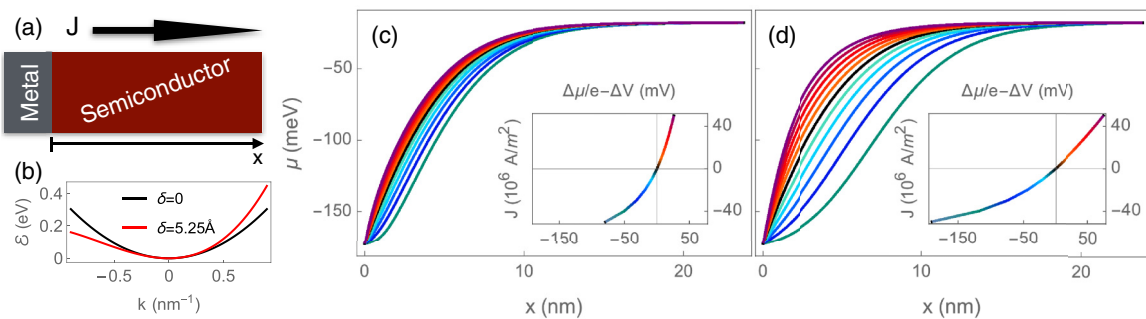

FIG. 1. (a) Geometry of the metal-semiconductor junction and the direction of positive current. The interface is at $x = 0$. (b) Modification of the bottom of the band structure as in Eq. (51). (c) Position-dependent chemical potential calculated for different currents at $T = 300$ K for the band structure with a finite δ, but using the textbook transport equations, i.e., setting $\sigma^{[\nabla\mathbf{E}]} = 0$ in the complete set of equations. The results depend on the value of the current running through the device (see the inset, which also provides the color code for the currents). (d) Same as (c), but calculated with the correct transport equations, including the additional term $\sigma^{[\nabla\mathbf{E}]}$. The insets in (c) and (d) provide the current vs voltage characteristic of the semiconductor layer.

distribution by $f_{FD} \approx \exp[(\mu - \mathcal{E})/k_B T]$. The conductivity and the charge ρ [see Eq. (A1)] assume the simple form

$$\sigma = \sigma_0 e^{\frac{\mu}{k_B T}}, \qquad (45)$$

$$\rho = \rho_D - \rho_0 e^{\frac{\mu}{k_B T}}, \qquad (46)$$

where σ_0 and ρ_0 are given by the expressions

$$\rho_0 = \frac{\sqrt{2\pi k_B T m}}{\hbar}, \qquad (47)$$

$$\sigma_0 = \frac{\tau e^2}{\hbar} \sqrt{\frac{2\pi k_B T}{m}}, \qquad (48)$$

and ρ_D is the dopant charge density. Thus, we obtain the textbook equations for the charge transport in semiconductors,

$$J = \sigma_0 e^{\frac{\mu(x)}{k_B T}} \left[\mathbf{E}(x) - \frac{\mu'(x)}{e} \right], \qquad (49)$$

$$\epsilon_0 \mathbf{E}'(x) = \rho_D - \rho_0 e^{\frac{\mu(x)}{k_B T}}, \qquad (50)$$

which have to be solved for the appropriate boundary conditions. For a complete description, we also need similar equations for the metal part of the junction (with the appropriate expressions for the conductivity and the charge) and, then, we have to link the two regions. The requirement is that the chemical potential and the electric field (or the electric displacement field, when the semiconductor and the metal have different dielectric constants) are continuous across the interface (again, we neglect the surface discontinuities in the potentials). However, if the metal is highly conductive and has a high density of states at the Fermi energy, it is sufficient to solve Eqs. (49) and (50) with the boundary conditions $\mu(0) = \mathcal{E}_{F,m}$ and $\mu(+\infty) = k_B T \ln(\rho_D/\rho_0)$. That is, we require that the chemical potential at the interface is set by the Fermi level $\mathcal{E}_{F,m}$ of the metal and that there is no net charge far away from the interface.

We consider a low-symmetry material with the dispersion (close to the band minimum) given by

$$\mathcal{E}(\mathbf{k}) = \frac{\hbar^2 (k^2 + \delta\, k_x^3)}{2m}, \qquad (51)$$

where k is the magnitude of the crystal momentum $\mathbf{k}$, k_x its the x component, and δ measures the asymmetry. The modification of band structure is depicted in Fig. 1(b). The difference with respect to the parabolic dispersion is that $\mathcal{E}(\mathbf{k})$ now has a finite third derivative at the minimum. As an example, we take a semiconductor similar to silicon ($m = 0.1 m_e$, where m_e is the electron mass, and $\tau = 10$ fs) but with an asymmetry $\delta = 5.25$ Å, doped with 3.6×10^{22} carriers/m^3, and attach it to a metal with Fermi level at -155 meV below the semiconductor conduction band.

The solution of Eqs. (49) and (50) provides the spatial profile of the electric field and the chemical potential. The results are shown in Fig. 1(c) for various operating conditions. The black line represents an open circuit (no current flowing) and shows the formation of the depletion region. As the current increases, the depletion region either expands, for negative currents, or shrinks, for positive ones. The inset shows a typical current-voltage characteristic of a metal-

semiconductor junction, i.e., it shows the dependence of the electrochemical potential in the semiconductor, $\mu(+\infty)/e - V(+\infty) - [\mu(0)/e - V(0)]$, on the current running through the junction.

We now compare this textbook solution with the case when the response of a semiconductor in the presence of additional driving forces $\partial^2\mu/\partial \mathbf{r}^2$ and $\nabla \mathbf{E}$. It is easy to prove that

$$\sigma^{[\nabla \mathbf{E}]} = \sigma_0^{[\nabla \mathbf{E}]} e^{\frac{\mu}{k_B T}}, \qquad (52)$$

where $\sigma_0^{[\nabla \mathbf{E}]}$ is a constant, and that $\sigma^{[\nabla \mathbf{E}]}$ vanishes for centrosymmetric materials ($\delta = 0$). That is, the shape of the conductivity and charge is still given by Eqs. (45) and (46), but the proportionality constants are now more complex. The macroscopic transport equation, including the additional terms, reads

$$J = \sigma_0 e^{\frac{\mu}{k_B T}} \left(\mathbf{E} - \frac{\mu'}{e} \right) + \sigma_0^{[\nabla \mathbf{E}]} e^{\frac{\mu}{k_B T}} \left(\mathbf{E}' - \frac{\mu''}{e} \right), \qquad (53)$$

while the Poisson equation is unchanged. As before, we also have to consider the transport equations in the metal and ensure that the electric field (or electric displacement field if the semiconductor and the metal have different dielectric constants), the chemical potential, and its first derivative are continuous across the interface.

In the absence of the gradient of the electrochemical potential, we can set $\sigma_0^{[\nabla \mathbf{E}]} = 0$ and reduce Eq. (53) to Eq. (49). When $\sigma_0^{[\nabla \mathbf{E}]} \neq 0$, the order of the differential equation rises and the boundary condition for the derivative is needed. We require that the first derivative has the same value as in the case $\sigma_0^{[\nabla \mathbf{E}]} = 0$ and report the solution of Eqs. (53) and (50) in Fig. 1(c). The comparison of Figs. 1(c) and 1(d) shows that in the absence of the current, the textbook equations and the complete transport equations lead to the same result [the black curves in Figs. 1(c) and 1(d) are the same], as can be deduced from Eq. (53) for $J = 0$. However, for $J \neq 0$, the dimension and shape of the depletion region are modified by the additional terms, i.e., the linear response theory differs considerably from the full description of the device. For negative currents, the depletion region given by the nonlinear description is wider than the textbook one, while for positive currents, the depletion region is reduced.

The above concepts can have a straightforward experimental verification in the case of a perfectly symmetrical metal-semiconductor-metal device, when the semiconductor has both inversion and time-reversal symmetries broken. The standard transport equations predict that when the current is flowing in the positive direction, the voltage drop is exactly the opposite to the one when the direction of the current is reversed. Thus, the standard equations predict that the two running conditions are perfectly symmetric. However, if the terms proportional to the derivative of the electric field and the second derivative of the chemical potential are taken into account, we expect the asymmetric behavior, as can be verified by applying the transformation $x \to -x$ to Eqs. (49) and (53).

We have provided just one example of the possible effects due to the additional terms obtained by the expansion of the Boltzmann distribution function. But the range of applicability and the relevance of that expansion is much wider and we expect that other terms will also become important.

VII. DISCUSSION AND CONCLUSIONS

In summary, our starting point is the stationary Boltzmann equation, in its most general form, subject to specific microscopic boundary conditions. To obtain the solution, we carefully define the relaxation-time approximation (RTA) and relate, for a given spatial profiles of temperature $T(\mathbf{r})$, chemical potential $\mu(\mathbf{r})$, and electric field $\mathbf{E}(\mathbf{r})$, the Boltzmann distribution function to the relaxation time. We then show that to obtain the physically meaningful results, $\mathbf{E}(\mathbf{r})$ has to satisfy the Poisson equation, while $T(\mathbf{r})$ and $\mu(\mathbf{r})$ have to satisfy the charge and energy conservations. Thus, together with the Boltzmann equation, we now have to self-consistently solve three additional equations. The initial problem for $g(\mathbf{r},\mathbf{k})$ has, apparently, been turned into a more complicated one, where $g(\mathbf{r},\mathbf{k})$ is a functional defined on the functions $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$.

We now focus on the strategy for solving the Boltzmann equation within the RTA, by assuming that $T(\mathbf{r})$, $\mu(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$, are analytic functions. Representing these functions by their respective Taylor series [33], we can write $g(\mathbf{r},\mathbf{k})$ as a function of infinitely many variables,

$$g(\mathbf{r},\mathbf{k}) = \tilde{g}\bigg[\mathbf{r},\mathbf{k},T(\mathbf{r}),\mu(\mathbf{r}),\mathbf{E}(\mathbf{r}),\nabla T(\mathbf{r}),\nabla \mu(\mathbf{r}),$$
$$\times \nabla \mathbf{E}(\mathbf{r}), \frac{\partial^2 T(\mathbf{r})}{\partial \mathbf{r}^2}, \frac{\partial^2 \mu(\mathbf{r})}{\partial \mathbf{r}^2}, \frac{\partial^2 \mathbf{E}(\mathbf{r})}{\partial \mathbf{r}^2}, \ldots, \text{BC}].$$

The general solution of the Boltzmann equation is obtained by expanding the distribution function with respect to all its variables, except the first five, i.e., $\mathbf{r},\mathbf{k},V(\mathbf{r}),T(\mathbf{r})$, and $\mu(\mathbf{r})$, and writing $g(\mathbf{r},\mathbf{k})$ as a multivariable power series. Since the expansion variables are completely arbitrary, substituting $g(\mathbf{r},\mathbf{k})$ in the Boltzmann equation yields an infinite number of coupled differential equations for the expansion coefficients. Integrating these equations for a particular set of microscopic boundary conditions yields $g(\mathbf{r},\mathbf{k})$ as a power series in terms of the expansion variables $\mathbf{E}(\mathbf{r}),\nabla T(\mathbf{r}),\nabla \mu(\mathbf{r}),\nabla^2 V(\mathbf{r}),\nabla^2 T(\mathbf{r}),\nabla^2 \mu(\mathbf{r}), \ldots$. Substituting $g(\mathbf{r},\mathbf{k})$ in the expressions for the charge and energy current densities, we obtain an expansion of $\mathbf{J}(\mathbf{r})$ and $\mathbf{J}_{\mathcal{E}}(\mathbf{r})$ in terms of their respective driving forces. The coefficients of the driving forces define the generalized transport coefficients. The physically relevant functions $\mathbf{J}(\mathbf{r})$, $\mathbf{J}_{\mathcal{E}}(\mathbf{r})$, $\mu(\mathbf{r})$, $T(\mathbf{r})$, and $\mathbf{E}(\mathbf{r})$ are obtained at every macroscopic point of the sample by self-consistently solving the transport equations, the charge and energy continuity equations, and the Poisson equation. We also show under which conditions the surface effects can be neglected and the distribution function of a macroscopic sample assumes a simple, textbook form.

The above procedure elucidates the commonly used derivation of transport equations and exposes various approximations employed in that derivation. In addition, it reveals contributions to the response functions which are proportional to the higher powers of the forces and their higher-order derivatives. The ensuing corrections to the charge and energy currents are usually small, which explains the success of the phenomenological transport theory in Eqs. (1) and (2). However, in certain situations, the additional terms lead to qualitatively different phenomena. For example, they can become important for heterogeneous devices, for materials in which the transport properties are strongly temperature and

potential dependent, for systems driven out of equilibrium by large thermodynamic forces (e.g., large temperature differences), or when the thermodynamic potentials vary strongly over small distances. Unlike the second-order response to the thermodynamic forces, such as the one due to $\mathbf{E}^2$, $(\nabla \mu)^2$, or $(\nabla T)^2$, here we discuss the response to the spatial derivatives of these forces, such as $\nabla \mathbf{E}$, $\nabla^2 \mu$, or $\nabla^2 T$. Since the magnitude of the additional higher-order terms is comparable to the already known ones, all the terms of the same order should be treated on the same footing. In other words, a consistent semiclassical description of transport phenomena should not just consider the higher powers of the thermodynamic forces, but should also take into account the driving forces which are proportional to the higher-order derivatives of temperature, chemical potential, and electric field.

The expansion of the distribution function described in this paper respects the microscopic boundary conditions of the Boltzmann equation. Thus, it can be used to treat, on the same footing, not just the higher-order and nonlocal effects in the bulk, but also the multitude of surface effects. Our expansion in terms of the driving forces provides a substantial improvement over the Hilbert expansion or similar expansions of the solution of the Boltzmann equation, which do not take into account the microscopic boundary conditions and therefore yield the solution which is valid only in the bulk.

APPENDIX A: CONTINUITY EQUATIONS

The particle continuity equation is obtained by integrating Eq. (3) over the whole k space. The first term of Eq. (3) becomes the time evolution of the local total number of particles, defined as

$$n(t,\mathbf{r}) = \int g(t,\mathbf{r},\mathbf{k})d^3k. \tag{A1}$$

In the second term of Eq. (3), the divergence with respect to spacial coordinates can be brought out of the k-space integral, leading to the spatial divergence of the particle current written as

$$\mathbf{J}(t,\mathbf{r}) = e \int \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar} g(t,\mathbf{r},\mathbf{k})d^3k. \tag{A2}$$

The third term of Eq. (3) can be proven to integrate to zero. The part multiplying the electric field $\mathbf{E} = \nabla V + \partial \mathbf{A}/\partial t$ vanishes due to a corollary of the divergence theorem in the k space and the periodicity of all the involved functions in k space. The part multiplying the magnetic field, $\mathbf{B} = \nabla \times \mathbf{A}$, requires the use of the identity $\nabla_{\mathbf{k}} \cdot (\nabla_{\mathbf{k}}\mathcal{E} \times \mathbf{B}) = 0$ and then the same considerations above.

The integral of the scattering term on the right-hand side of Eq. (3) depends on its precise expression. We assume the RTA in Eq. (4). As already mentioned in Sec. II, the value of the integral will depend on the local temperature $T(\mathbf{r})$ and chemical potential $\mu(\mathbf{r})$, as well as the local electron distribution $g(t,\mathbf{r},\mathbf{k})$ and, in general, can be different from zero. This would imply that some particles are either destroyed or created during the scattering. It is indeed to prevent this unphysical effect that we imposed the constraint in Eq. (7), for the RTA to make sense. Using Eq. (7), the conservation equation reduces to

$$\frac{\partial}{\partial t}\int g(t,\mathbf{r},\mathbf{k})d^3k + \nabla \cdot \int \frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}g(t,\mathbf{r},\mathbf{k})d^3k = 0, \quad (A3)$$

which states that the variation in the local number of particles n has to be equal to the divergence of the particle current density $\mathbf{J}(\mathbf{r})$.

Similarly the energy continuity equation can be obtained by multiplying Eq. (3) by the particle energy $\mathcal{E}(\mathbf{k})$ and integrating over the whole k space. The first term will give the change in the energy density, $\epsilon(t,\mathbf{r}) = \int \mathcal{E}(\mathbf{k})\, g(t,\mathbf{r},\mathbf{k})d^3k$. The second yields the divergence of the energy current density, $\mathbf{J}_\epsilon(t,\mathbf{r}) = \int \mathcal{E}\, g\, \nabla_{\mathbf{k}}\mathcal{E}/\hbar\, d^3k$. The third term in this case does not vanish, but is the work made by the electrical field on the system $W(t,\mathbf{r}) = e/\hbar \int \mathcal{E}\, \mathbf{E}\cdot\nabla_{\mathbf{k}}g\, d^3k$ leading to the Joule heating. Again we use the RTA for the scattering term. Its integral is now constrained by Eq. (8) [eventually with the effect of phonons included, as explained in the text below Eq. (8)]. We therefore obtain

$$\frac{\partial}{\partial t}\int \mathcal{E}\, g(t,\mathbf{r},\mathbf{k})d^3k + \nabla \cdot \int \mathcal{E}\frac{\nabla_{\mathbf{k}}\mathcal{E}}{\hbar}g(t,\mathbf{r},\mathbf{k})d^3k$$
$$+ \frac{e}{\hbar}\int \nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla V g\, d^3k = \frac{\Delta\epsilon_{e-ph}}{\tau}, \quad (A4)$$

where we have integrated by parts the Joule heating term and used again the periodicity of all the involved functions in k space. The equation implies that any change in the local total energy ϵ is caused by the divergence of the energy current $\mathbf{J}_\epsilon$ and the work W done on the charged particles by the electric field and the energy dissipated into phonons $\Delta\epsilon_{e-ph}$.

APPENDIX B: COMPARISON WITH HILBERT EXPANSION

Some of the results derived in this paper can also be obtained from the Hilbert expansion for the Boltzmann equation in the relaxation-time approximation. However, as shown below, our approach overcomes one fundamental limitation of the Hilbert expansion, which critically limits its range of applicability (as well as its use in the longstanding mathematical problem of the proof of the existence of the solution of the Boltzmann equation in the presence of the boundary conditions).

By adapting the Hilbert expansion to the present case, we look for the solution of Eq. (6) in the form

$$g(\mathbf{r},\mathbf{k}) = \sum_{i=0}^{\infty}\tau^i g_H^{[i]}(\mathbf{r},\mathbf{k}), \quad (B1)$$

where here the relaxation time τ plays the role of the Knudsen parameter. Substituting the above series in the Boltzmann equation and collecting all the terms of the same order in τ gives the result

$$g_H^{[0]}(\mathbf{r},\mathbf{k}) = f_{FD}(T(\mathbf{r}),\mu(\mathbf{r}),\mathcal{E}(\mathbf{k})),$$
$$g_H^{[1]}(\mathbf{r},\mathbf{k}) = -\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}g_H^{[0]} + \frac{e}{\hbar}\mathbf{E}\cdot\nabla_{\mathbf{k}}g_H^{[0]},$$
$$g_H^{[2]}(\mathbf{r},\mathbf{k}) = -\frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}\cdot\nabla_{\mathbf{r}}g_H^{[1]} + \frac{e}{\hbar}\mathbf{E}\cdot\nabla_{\mathbf{k}}g_H^{[1]},$$
$$\cdots = \cdots. \quad (B2)$$

Successively substituting the lower-order corrections into the higher-order ones, we find that the distribution function defined by Eq. (B2) is equivalent to the one given by Eq. (33). Note that the expansion in Eq. (B1) does not take into account the boundary conditions, so that it provides just one of the infinitely many solutions of the Boltzmann equation. In fact, the result given by Eqs. (B2) or (33) corresponds to a specific choice of the boundary conditions, as discussed in Sec. IV. Hence, the solution obtained by the Hilbert expansion does *not*, in general, satisfy the imposed boundary conditions.

This shows the most important difference between the approach taken in this work and the one taken by Hilbert expansion. Each term in the expansion defined by Eq. (18) of this work satisfies one of the differential equations given by Eqs. (21)–(26), and the sum of all these terms provides a particular solution of the Boltzmann equation that complies with the imposed boundary conditions. Thus, the expansion in terms of the driving forces presented in this work is much more powerful than the Hilbert expansion.

[1] L. Onsager, Phys. Rev. **37**, 405 (1931).
[2] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
[3] J. M. Luttinger, Phys. Rev. **135**, A1505 (1964).
[4] G. E. H. Reuter and E. H. Sondheimer, Proc. R. Soc. A **195**, 336 (1948).
[5] R. G. Chambers, Proc. R. Soc. A **215**, 481 (1952).
[6] R. Chau, B. Doyle, S. Datta, J. Kavalieros, and K. Zhang, Nat. Mater. **6**, 810 (2007).
[7] H. Iwai, Microelectron. Eng. **86**, 1520 (2009).
[8] I. Ferain, C. A. Colinge, and J.-P. Colinge, Nature (London) **479**, 310 (2011).
[9] J. A. d. Alamo, Nature (London) **479**, 317 (2011).
[10] J. He, Y. Liu, and R. Funahashi, J. Mater. Res. **26**, 1762 (2011).
[11] G. Rogl, A. Grytsiv, P. Rogl, E. Royanian, E. Bauer, J. Horky, D. Setman, E. Schafler, and M. Zehetbauer, Acta Mater. **61**, 6778 (2013).
[12] K. Koumoto, R. Funahashi, E. Guilmeau, Y. Miyazaki, A. Weidenkaff, Y. Wang, and C. Wan, J. Am. Ceram. Soc. **96**, 1 (2013).
[13] J. P. Heremans, M. S. Dresselhaus, L. E. Bell, and D. T. Morelli, Nat. Nanotechnol. **8**, 471 (2013).

[14] L.-D. Zhao, S.-H. Lo, Y. Zhang, H. Sun, G. Tan, C. Uher, C. Wolverton, V. P. Dravid, and M. G. Kanatzidis, Nature (London) **508**, 373 (2014).

[15] G. Mahan, J. Appl. Phys. **70**, 4551 (1991).

[16] G. Mahan, J. Appl. Phys. **87**, 7326 (2000).

[17] W. Seifert, M. Ueltzen, and E. Müller, Phys. Stat. Sol. (a) **194**, 277 (2002).

[18] H. S. Kim, W. Liu, G. Chen, C.-W. Chu, and Z. Ren, Proc. Natl. Acad. Sci. USA **112**, 8205 (2015).

[19] G. S. Nolas, J. Sharp, and H. J. Goldsmid, *Thermoelectrics Basic Principles and New Materials Developments* (Springer-Verlag, Berlin, 2001).

[20] V. Zlatić and R. Monnier, *Modern Theory of Thermoelectricity* (Oxford University Press, Oxford, 2014).

[21] M. Battiato, J. M. Tomczak, Z. Zhong, and K. Held, Phys. Rev. Lett. **114**, 236603 (2015).

[22] L. Boltzmann, Further studies of thermal equilibrium between gas molecules, in *Kinetic Theory II: Irreversible Processes, Introduction and Original Papers* (Vieweg+Teubner Verlag, Wiesbaden, 1970), p. 115.

[23] J. M. Ziman, *Electrons and Phonons* (Oxford University Press, Oxford, 1972).

[24] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976).

[25] D. Hilbert, Math. Ann. **72**, 562 (1912).

[26] S. Chapman, Philos. Trans. R. Soc. A **216**, 279 (1916).

[27] D. Enskog, Kinetic theory of the processes in not-too-diluted gases, Ph.D. Dissertation, Uppsala University, 1917.

[28] C. Cercignani, *The Boltzmann Equation and Its Applications* (Springer-Verlag, New York, 1988).

[29] B. Chopard and M. Droz, *Cellular Automata Modeling of Physical Systems* (Cambridge University Press, Cambridge, 1998).

[30] N. A. Sinitsyn, J. Phys.: Condens. Matter **20**, 023201 (2008).

[31] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010).

[32] E. M. Lifschitz and L. P. Pitaevskii, *Statistical Physics: Theory of the Condensed State* (Butterworth-Heinemann, Oxford, 1980).

[33] An analytic function $f(\mathbf{r})$ can be expanded into Taylor series $f(\mathbf{r}_0) = \sum_0^\infty (\mathbf{r} - \mathbf{r}_0)^n \nabla^n f(\mathbf{r})/n!$ anywhere in its domain of definition. Hence, the functional form of $f(\mathbf{r})$ is completely determined by the value of the derivatives $\nabla^n f(\mathbf{r})$ for $n = 0, 1, 2, \ldots$ at any point $\mathbf{r}$.

[34] G. L. Pollack, Rev. Mod. Phys. **41**, 48 (1969).

[35] S. Ukai and T. Yang, *Mathematical Theory of Boltzmann Equation*, Lecture Notes Series-No. 8 (Liu Bie Ju Center for Mathematical Sciences, City University of Hong Kong, Hong Kong, 2006), p. 116.

[36] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices* (Wiley, Hoboken, NJ, 2007).