Quasiparticle Boltzmann Equation in Semiconductors

Václav Špička
The Blackett Laboratory, Imperial College, London, SW7 2BZ, United Kingdom

Pavel Lipavský
Institute of Physics, Academy of Sciences, Na Slovance 2, 18040 Praha 8, Czech Republic
(Received 2 June 1994)

The Boltzmann equation for electrons in semiconductors at room temperature is recovered from nonequilibrium Green's functions. Our approach eliminates gradient corrections to scattering rates, reduces "possible" collisional broadening, provides backflows, and allows one to evaluate scattering rates of two-phonon processes.

PACS numbers: 72.10.Bg, 72.20.Dp

The Boltzmann equation (BE) [1],

$$\frac{\partial f}{\partial t}+\frac{\partial \varepsilon}{\partial k}\frac{\partial f}{\partial r}-\frac{\partial \varepsilon}{\partial r}\frac{\partial f}{\partial k}=I_{\mathrm{in}}-I_{\mathrm{out}}, \tag{1}$$

is the most useful theoretical tool of nonequilibrium statistics. It links a macroscopic phenomenological picture of an interplay between diffusion and dissipation with a microscopic theory of these processes.

A theoretical challenge is to derive this equation as a quasiclassical limit of the nonequilibrium quantum statistics. Although this limit has a lot of general features, there is no general theory which would cover all (at least fermionic) systems. Instead, there are a number of special approaches to individual systems and a corresponding number of criteria for the validity of the BE. The electron gas in nondegenerate semiconductors is covered only by theories on the level of the Kadanoff and Baym (KB) ansatz [2]. These approaches are not sufficient to guarantee the validity of the BE in wide gap semiconductors, say GaAs or Si, at room temperature. Here we advance the original KB approach in two directions. First, we eliminate some frustrating weaknesses. Second, we derive a new generating functional for scattering rates which works also for two-phonon processes.

As a guide to the KB approach, we recall the theory of metals. As visualized by Landau, the structure of a transport equation for electrons in crystals is identical to the BE. However, physical quantities entering such an equation are not associated with (bare) electrons but with electronlike elementary excitations—the quasiparticles [1]. Thus, to derive the BE one has to define a function which can be interpreted as the quasiparticle distribution f, prove that f obeys the BE, and express observables in terms of f. This program has been fulfilled with the help of many-body Green's functions [1,3].

The theory of metals benefits from the high density of particles, which implies that the only excitations accessible at room temperature are fast quasielectrons and slow acoustic phonons. In this particular regime, the perturbative expansion in the electron-electron and electron-phonon interactions simplifies to self-consistent single-loop diagrams in the spirit of the random phase approximation and the Migdal theorem.

Another simplification following from the high density limit is that the quasiparticle distribution can be defined as an integral over the bare electron kinetic energy $\epsilon \sim k^2/2m$ from the single-electron Green's function [3]

$$f \equiv \int \frac{d\epsilon}{2\pi} G^{<}(\omega,k,r,t). \tag{2}$$

The definition (2) allows us to derive the BE simply by taking the integral over ϵ from the KB equation

$$[G_0^{-1}-\Sigma,G^{<}]+[G,\Sigma^{<}]=A\Sigma^{<}-\Gamma G^{<}. \tag{3}$$

Here and below, all functions depend on an energy ω, momentum k, coordinate r, and time t, i.e., $G^{<}\equiv G^{<}(\omega,k,r,t)$. The Poisson bracket $[A,B]\equiv \partial_\omega A\partial_t B-\partial_t A\partial_\omega B-\partial_k A\partial_r B+\partial_r A\partial_k B$ is a shorthand notation for the gradient approximation of commutators. The spectral functions of the single-electron propagator and self-energy are denoted by $A\equiv i(G^R-G^A)$ and $\Gamma\equiv i(\Sigma^R-\Sigma^A)$, respectively. The functions $\Sigma\equiv\frac{1}{2}(\Sigma^R+\Sigma^A)$ and $G\equiv\frac{1}{2}(G^R+G^A)$ are corresponding real parts, and $G_0^{-1}=\omega-\epsilon-U$, where U is an internal potential. After the integration over the kinetic energy ϵ, the Poisson bracket $[G_0^{-1}-\Sigma,G^{<}]$ turns into the diffusion term [the left hand side (LHS)] of the BE (1), $A\Sigma^{<}\longrightarrow I_{\mathrm{in}}$, $\Gamma G^{<}\longrightarrow I_{\mathrm{out}}$, and the Poisson bracket $[G,\Sigma^{<}]$, alien to the BE, vanishes in the high density limit. All elements of the self-energy ($\Sigma^{<}$, Γ, and Σ) are functionals of $G^{<}$, but convolutions with weakly momentum-dependent bosonic Green's functions allow us to use f instead of $G^{<}$. In this way one arrives at a closed equation for f which is the BE, see [3].

An electron gas in semiconductors, unfortunately, does not belong to this limit. Since slow quasielectrons are thermally activated, the definition (2) of the quasiparticle distribution does not work and the term $[G,\Sigma^{<}]$ does not vanish. Moreover, the Migdal theorem does not hold for optical phonons, and one has to be ready to handle multiphonon processes. In spite of these complications, we believe that Landau's vision of quasiparticles applies

0031-9007/94/73(25)/3439(4)$06.00

also to semiconductors and that a more general theory covering the two systems can be developed.

A draft of the theory which treats metals and semiconductors equally has already been proposed by Kadanoff and Baym [2]. They have introduced the ansatz [based on the equilibrium relation $G^<(\omega,k) = f_{\rm FD}(\omega)A(\omega,k)$]

$$G^<(\omega,k,r,t) \approx f(k,r,t)A(\omega,k,r,t) \tag{4}$$

$$\approx f(k,r,t)2\pi z(k,r,t) \times \delta(\omega - \varepsilon(k,r,t)), \tag{5}$$

where $z^{-1} = \partial_\omega(G_0^{-1} - \Sigma) = 1 - \partial_\omega\Sigma$. Using its form (5) in the KB equation (3) and neglecting the nonsingular term $[G,\Sigma^<]$, they have obtained the BE with no regard to the electron density. In contrast to the theory of metals, this approach does not guarantee the validity of the BE because of the following shortcomings.

(A) There is no estimate saying under which conditions $[G,\Sigma^<]$ can be neglected. This term is expected to provide gradient corrections of scattering rates.

(B) The ansatz (5) is an approximation of (4), but the form (4) does not lead to the BE without this additional approximation [3]. In particular, using (4) one violates the energy conservation of individual collisions [4].

(C) From the reduced density matrix $\rho \equiv \int(d\omega/2\pi)G^<$ one can evaluate all single-electron observables, but (4) or (5) are too crude as approximations of $G^<$. For instance, there is no backflow and form (5) fails to provide the correct number of particles.

(D) Multiphonon scattering rates are incorrect. These problems are really artifacts of the KB approach because they persist also in the high density limit while the theory of metals is free of (A) – (C). The problem (D) appears in the theory of metals too, but it is ignored in the spirit of the Migdal theorem. All the problems (A)–(D) point to the same property of the Green's function neglected by the KB ansatz: (A) the Poisson bracket $[G,\Sigma^<]$ is an off-pole term; (B) form (5) differs from (4) in the off-pole region; (C) the backflow and missing particles are hidden in the off-pole part of $G^<$; and (D) it is a missing off-pole propagation between two consecutive "single-phonon" processes which makes the two-phonon scattering rates incorrect.

Apparently, we have to fix the off-pole part of $G^<$. To this end, however, we have to change our attitude towards the KB approach. Their theory includes two kinds of approximations, the ansatz itself and the omissions in the transport equation. The two kinds are closely related. One can dream of an ansatz which directly solves the KB equation; therefore no approximation of (3) would be necessary. Such a task is unrealistic, and we have to distinguish which phenomena can be better approximated on the level of an ansatz and which phenomena have to be solved from the transport equation. The basic key for this distinction is given by two time scales inherent for systems with quasiparticles: the long (hydrodynamical) and short (collision duration, quasiparticle formation) time scales associated with the pole and off-pole parts of $G^<$, respectively. The ansatz is constructed from the quasielectron (and boson) distribution at the instant t. Therefore it can provide reliably only the behavior of $G^<$ in the vicinity of t. Thus the short time phenomena should go into the ansatz. The differential transport equation deals primarily with the long time phenomena. Accordingly, the ansatz has to have at least two parts: a pole contribution similar to (5) which solves the pole part of the KB equation provided that f solves some auxiliary Boltzmann-like transport equation and an off-pole part which solves the KB equation far from the pole. In other words,

$$G^< = f\mathcal{A} + \delta G^<, \tag{6}$$

where $\mathcal{A} \approx 2\pi z\delta(\omega - \varepsilon)$ and $\delta G^<$ is an off-pole correction. A choice of $\mathcal{A}$ can be motivated by a form of the KB equation derived by Botermans and Malfliet [5], Eq. (2.109). We find it shorter to deduce the functions $\mathcal{A}$ and $\delta G^<$ directly from the KB equation (3).

Let us identify $\delta G^<$ first. After substitution of the ansatz (6) into the KB equation (3), one finds

$$A[G_0^{-1} - \Sigma, f] + f[G_0^{-1} - \Sigma, \mathcal{A}] + [G_0^{-1} - \Sigma, \delta G^<] + [G,\Sigma^<] = A\Sigma^< - \mathcal{A}\Gamma f - \Gamma\delta G^<. \tag{7}$$

The off-pole correction $\delta G^<$ is fitted such that $[G_0^{-1} - \Sigma, \delta G^<]$ cancels the off-pole term $[G,\Sigma^<]$. The latter can be rearranged as

$$[G,\Sigma^<] = -\frac{1}{2}[G_0^{-1} - \Sigma, G^R\Sigma^<G^R + G^A\Sigma^<G^A] - \frac{i}{4}[\Gamma, G^R\Sigma^<G^R - G^A\Sigma^<G^A]. \tag{8}$$

Here, the first term is dominant. Thus the function

$$\delta G^< = \frac{1}{2}(G^R\Sigma^<G^R + G^A\Sigma^<G^A) \tag{9}$$

is the off-pole part of $G^<$. Indeed, far from the pole both propagators become real, $G^R \approx G^A \approx G$. Therefore $G^< = G^R\Sigma^<G^A \approx G\Sigma^<G \approx \delta G^<$.

Now we can specify $\mathcal{A}$. The function $G^< - \delta G^<$ is a "quasiparticle" correlation function. In equilibrium, $G^< - \delta G^< = f_{\rm FD}\mathcal{A}$, where

$$\mathcal{A} = A - \frac{1}{2}(G^R\Gamma G^R + G^A\Gamma G^A) \tag{10}$$

$$= \frac{\Gamma^3/2}{[(G_0^{-1} - \Sigma)^2 + \Gamma^2/4]^2} \approx 2\pi z\delta(\omega - \varepsilon) \tag{11}$$

is a quasiparticle spectral function. The ansatz (6) is in this sense a pole approximation of $G^< - \delta G^<$. From (11) one can see that the quasiparticle spectral function $\mathcal{A}$ is a better δ function than A, because far from the pole $\mathcal{A} \sim (\omega - \varepsilon)^{-4}$ while $A \sim (\omega - \varepsilon)^{-2}$.

The approximation of $G^<$ on the short time scale is already specified by $\delta G^<$, Eq. (9), and $\mathcal{A}$, Eq. (10). Now we turn to approximations on the long time scale. According to our expectation, a substitution of $\delta G^<$ from (9) should make the KB Eq. (7) free of all off-pole terms.

If this is true, all terms left in the transport equation will play a role on the long time scale, and we have to include them in an auxiliary transport equation, even if they do not match the structure of the BE.

Elimination of the off-pole part simplifies the transport equation remarkably:

$$\mathcal{A}[G_0^{-1} - \Sigma, f] - \frac{1}{2}GA\Gamma[\Gamma, f] = \mathcal{A}(\Sigma^< - \Gamma f). \tag{12}$$

In the LHS we have disregarded the term $[\Gamma, \frac{1}{2}GA(\Gamma f - \Sigma^<)]$ which vanishes in the quasiclassical limit being effectively quadratic in gradients. Indeed, from (12) one can see that $\Sigma^< - \Gamma f$ is proportional to gradient terms. This step has a very clear physical meaning: If there are linear gradient corrections to scattering-in and -out integrals, they mutually compensate, leaving only the term proportional to gradients of the distribution f.

To complete the derivation of the auxiliary equation, we integrate (12) over ω. The pole of $\mathcal{A}$ is at

$$\varepsilon = \varepsilon_0 - \frac{1}{4}z^2\Gamma\frac{\partial\Gamma}{\partial\omega}\bigg|_{\omega=\varepsilon_0}, \tag{13}$$

where ε_0 is a root of $G_0^{-1} - \Sigma = 0$, i.e., $\varepsilon_0 = \epsilon + U + \Sigma_{\omega=\varepsilon_0}$. In the pole approximation one finds that the first term of (12), $\int\frac{d\omega}{2\pi}\mathcal{A}[G_0^{-1} - \Sigma, f] = \partial_t f - [\varepsilon_0, f]$, gives the diffusion term known from metals, and the second term, $\int\frac{d\omega}{2\pi}\frac{1}{2}GA\Gamma[\Gamma, f] = [\varepsilon - \varepsilon_0, f]$, gives a correction (13) to the quasiparticle energy. Apparently, the second term in (12) is not alien to the transport equation, but it is a vital part, making the diffusion term consistent with the spectral function. If we approximate $\mathcal{A}$ by a δ function, see (11), the auxiliary equation we have obtained is just the BE for quasiparticles

$$\frac{\partial f}{\partial t} + \frac{\partial\varepsilon}{\partial k}\frac{\partial f}{\partial r} - \frac{\partial\varepsilon}{\partial r}\frac{\partial f}{\partial k} = z(\Sigma^< - \Gamma f)_{\omega=\varepsilon}. \tag{14}$$

The structure of the transport equation (14) is identical to the BE (1). The quantities entering this equation, however, are evaluated in a slightly different way than in previous treatments. First, there are off-pole contributions in the scattering integrals. Second, the quasiparticle energy includes the correction (13).

Now we are ready to discuss to what extent we have managed to get the four problems (A)–(D) under control.

(A) the term $[G, \Sigma^<]$ nearly cancels with other off-pole contributions. The rest of this term does not result in gradient corrections of scattering rates, but in a small correction of the quasiparticle energy. This correction can easily be understood in a complex energy plane. The pole of the propagator G^R is given by $\varepsilon_R = \epsilon + U + \Sigma^R(\varepsilon_R)$. The simplest approximation is ε_0, and a linear correction is $\varepsilon_R - \varepsilon_0 = -\frac{i}{2}\Gamma + (\varepsilon_R - \varepsilon_0)\partial_\omega\Sigma^R$. Since the derivative of the self-energy is in general complex, $\partial_\omega\Sigma^R = \partial_\omega\Sigma - \frac{i}{2}\partial_\omega\Gamma$, the real part of ε^R shifts into ε, Eq. (13). This shows that our theory makes only the first step from the real axis but does not change the interpretation of the quasiparticle energy.

(B) Let us estimate what happens if we do not approximate $\mathcal{A}$ by a δ function. Within the KB approach one obtains the so-called collisional broadening given by a convolution of two spectral functions [4]. Starting from (12), one finds identical expressions with $\mathcal{A}$ substituted for A. To estimate the difference between the A and $\mathcal{A}$ broadening, we have evaluated the time evolution of an isotropic and monoenergetic distribution of electrons scattered by impurities. The elastic scattering cannot change this distribution; therefore, a time evolution following from the broadening is clearly an artifact of the theory. The correct value of the distribution, $F_\delta(\epsilon, t) = \delta(\epsilon - E)$, we take as an initial condition at $t = 0$. For $\Gamma t \gg 1$, the distribution spreads from the initial energy E as

$$f_A(\epsilon, t) \sim \frac{2t\Gamma^2}{(E - \epsilon)^2 + t^2\Gamma^4}, \tag{15}$$

$$f_{\mathcal{A}}(\epsilon, t) \sim \sqrt{\frac{4\pi}{t\Gamma^3}}\exp\left(-\frac{(E - \epsilon)^2}{t\Gamma^3}\right). \tag{16}$$

The A broadening creates overestimated high energy tails. Moreover, even in the low energy region, the width of f_A exceeds the thermal smearing after a time given by $t\Gamma^2 \sim k_BT$, which for a typical lifetime $1/2\Gamma = 10^{-12}$ s and room temperature restricts the validity to times shorter than 4×10^{-11} s. Clearly, the ansatz (5) does not follow from (4), because the results of the two approximations are strikingly different. On the other hand, with the $\mathcal{A}$ broadening, the power-law high energy tails are absent and the time after which the width of $f_{\mathcal{A}}$ exceeds room temperature, $\sqrt{t\Gamma^3} \sim k_BT$, is 1.6×10^{-9} s. Unfortunately, our approach does not rule out the collisional broadening, but it at least reduces its extent. These estimates do not include inelastic processes which reduce the artifacts of broadening. The above characteristic times show how strong the inelastic scattering rate has to be to keep the broadening negligible. This provides the criterion, $\Gamma_{\text{inel}}(k_BT)^2 > (\Gamma_{\text{el}} + \Gamma_{\text{inel}})^3$, for when $\mathcal{A}$ can be approximated by a δ function. This criterion is sufficient but overly restrictive. Nevertheless, it is easily met in GaAs or Si at room temperature. We note that the spectral function $\mathcal{A}$ is a first term of an expansion in small Γ, i.e., it is an imaginary part of $G^R(\omega - \frac{i}{2}\Gamma) \approx G^R(\omega) + \frac{i}{2}\Gamma G_R^2(\omega) = \frac{1}{4}GA\Gamma - \frac{i}{2}\mathcal{A}$. Provided one finds a way to generate corrections to the ansatz which will give us the entire expansion, one arrives at an exact δ function. If we extrapolate experience from the first step, such corrections will contribute to subtle details of higher order scattering integrals while the δ functions will be associated with initial and final states just like in the BE.

(C) The functional $\rho[f]$ evaluated from the ansatz (6) includes the off-pole part

$$\rho \equiv \int\frac{d\omega}{2\pi}G^< \approx fz + \int\frac{d\omega}{2\pi}\delta G^<. \tag{17}$$

This off-pole part is only approximate. It is the first term of the expansion in small Γ; thus it applies only

for systems with a weak interaction, where $z \approx 1 + \partial_\omega \Sigma$. Within this accuracy, (17) reproduces the backflows known from the theory of metals.

(D) Except for single-phonon processes, the scattering integrals evaluated from the ansatz (6) differ from those obtained within the KB ansatz (5). This is because $I_{\rm in}[f]$ results from $\Sigma^<[G^<]$ as

$$I_{\rm in}[f] = z\Sigma^<[\mathcal{A}f + \delta G^<]_{\omega=\varepsilon}. \tag{18}$$

As $\delta G^<$ depends on $\Sigma^<$, this nesting generates higher orders in the interaction strength. Let us compare (18) with the Fermi golden rule for the electron-phonon scattering up to two-phonon processes.

The Green's function expansion is represented by the self-consistent diagrams (a) and (b), the Fermi golden rule by the non-self-consistent diagrams (c)–(e). The scattering rates resulting from the KB ansatz are easily reproduced if we neglect $\delta G^<$. In this approximation, diagram (a) corresponds to diagram (c) and diagram (b) to diagram (e). Keeping $\delta G^<$, we obtain a correction from the nesting of diagram (a) into itself. From $\Sigma^<_{(a)}(\omega) = \int \frac{dE}{2\pi} D^<(\omega - E) G^<(E)$, where $D^<$ is a phonon Green's function, we find that the lowest $\delta G^<$ correction to $\Sigma^<_{(a)}$ contributes to the two-phonon scattering rate as

$$\begin{aligned}\Sigma^<_{(d)}(\omega) &= \int \frac{dE}{2\pi} D^<(\omega - E) G^{(2)}(E) \Sigma^<(E) \\ &= \int \frac{dE}{2\pi} \frac{d\mu}{2\pi} D^<(\omega - E) G^{(2)}(E) \\ &\quad \times D^<(E - \mu) f \mathcal{A}(\mu). \end{aligned} \tag{19}$$

The form of this contribution corresponds to the diagram (d), which is missing in the KB approach. Note that the intermediate propagation $G^{(2)} = \frac{1}{2}(G^R G^R + G^A G^A)$ has only the off-pole part; therefore, a double count of consecutive single-phonon processes is excluded.

We have modified the original KB approach so that we separate the pole and off-pole parts of the single-electron Green's function. Although this is only the first step in the small Γ expansion, it is sufficient to prove that the Boltzmann equation holds for wide gap semiconductors at room temperature. In particular, it (A) eliminates gradient corrections of scattering rates; (B) reduces "possible" collisional broadening to a negligible extent; (C) provides the observables in terms of the quasiparticle distribution including backflows; and (D) allows one to evaluate scattering integrals of two-phonon processes.

This work was supported by the Grant Agency of Academy of Sciences of Czech Republic, the Royal Society, London, and the Norwegian Research Council.

[1] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice Hall, Englewood Cliffs, New Jersey, 1963).

[2] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).

[3] R. E. Prange and L. P. Kadanoff, Phys. Rev. **134**, A338 (1964).

[4] L. Reggiani, P. Lugli, and A. P. Jauho, Phys. Rev. B **36**, 6602 (1987).

[5] W. Botermans and R. Milfliet, Phys. Rep. **198**, 115 (1990) [$\mathcal{A} \equiv a(1 - \mathrm{Re}\,\alpha)$ in Eq. (2.109)].