Quantum corrections to the Boltzmann equation for transport in semiconductors in high electric fields

Sanjoy K. Sarker
Department of Physics and Astronomy, University of Alabama, University, Alabama 35486-1921

J. H. Davies, F. S. Khan,* and J. W. Wilkins
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501
(Received 24 January 1986)

We have investigated quantum corrections to the Boltzmann equation for transport in high, spatially uniform electric fields, with weak scattering. Parabolic bands show no effects for constant fields up to 10^7 V/cm; nonparabolic bands or time-varying fields give quantum corrections by distorting the energy-conserving δ function in semiclassical transition rates. We have found that in practical cases the scale of this is less than that from collisional broadening, almost invariably neglected. These results extend greatly the known range of validity of the Boltzmann equation in semiconductors.

Recent experiments[1,2] have demonstrated directly the existence of ballistic carriers in semiconducting devices. Important features of these experiments are the following: high electric fields (exceeding 10^4 V/cm), that accelerate the ballistic electrons and drive the distribution far from equilibrium; contact effects, which are used to inject the electrons over a barrier[1] or through a tunnel junction,[2] and which rely on rapid spatial variation of the electric field to launch the ballistic carriers; and very high scattering rates, which necessitate an extremely narrow active region (300–600 Å) to preserve the ballistic features.

The theory of transport in semiconductor structures is traditionally based on the semiclassical Boltzmann equation, the solution of which for a submicrometer device[3] has demonstrated the importance of ballistic carriers. However, these experiments emphasize important questions concerning the validity of the Boltzmann equation in such structures.[4–6] (a) Can transition rates calculated using the golden rule be used (as in Ref. 1), even in the presence of a high (uniform) electric field? (b) At what field strengths does the Boltzmann transport equation itself fail? (c) Do the electric field and density vary too rapidly in space for the semiclassical description to be applicable?

In this paper we address the first two questions, paving the way for an approach to the third. We identify causes of possible deviations from the Boltzmann equation and present, for the first time, quantitative estimates of the associated energy scales. Our analysis is based on an integral equation derived recently by one of us[7] from the quantum transport theory.[8] This equation is valid for nondegenerate electrons moving in a spatially homogeneous, time-varying field and coupled weakly to phonons and impurities.

Our main results are as follows. Contrary to previous suggestions[5] we find that there is no intracollisional field effect, and the effect of collision duration is negligible since the collision-duration time t_c is smaller than the time between collisions τ_{sc} by a factor of 10^{-3}. For a constant field and a parabolic band the Boltzmann equation is valid for fields as strong as 10^7 V/cm. For nonparabolic bands or time-varying fields deviations from the Boltzmann equation do appear, solely through a distortion of the energy-conserving δ function in the semiclassical transition rates. This distortion can be characterized by a suitably defined energy scale λ. We find for a nonparabolic band $\lambda_\alpha \propto \alpha^{1/3}F^{2/3}$, where α is a measure of the nonparabolicity and F is the field strength. For a time-varying field of frequency ω, $\lambda_\omega \propto \omega^{1/3}F^{1/3}$, while for a field switched on at $t=0$, $\lambda_{\rm trans} \propto F^{1/2}$. Numerically, λ_α, λ_ω, and $\lambda_{\rm trans}$ are all comparable to $\lambda_{sc} \approx 1$–6 meV, which characterizes the width due to scattering (collisional broadening) and which is almost always ignored.[9]

For a spatially homogeneous system it is convenient to introduce the field $\mathbf{F}(t)$ by a vector potential $\mathbf{A}(t) = -\int_0^t \mathbf{F}(t)dt$. The electrons belong to a single band with dispersion relation $E(\mathbf{p})$ which, for a parabolic band, reduces to $p^2/2m^*$, where m^* is the effective mass. Here we consider only electron-phonon scattering, which is most relevant to hot electron physics. For weak scattering the Born approximation for the self-energies will be used. In addition, the phonon subsystem is considered to be in equilibrium. We use units such that $\hbar = c = 1$.

The physical information is contained in the distribution function $f(\mathbf{p},T)$. In the weak-scattering limit $f(\mathbf{p},T)$ satisfies the integral equation[7]

$$f(\mathbf{p},T) = \int_0^\infty dT' \exp\left[-\int_0^{T'} dT'' \int_{-2T''}^{2T''} d\tau\, \gamma(\mathbf{k}(\mathbf{p},T'',T),\tau,T-T'')\right] \int_{-2T'}^{2T'} d\tau\, \varepsilon(\mathbf{k}(\mathbf{p},T',T),\tau,T-T') , \quad (1a)$$

where

$$\mathbf{k}(\mathbf{p},t_1,t_2)=\mathbf{p}+e\mathbf{A}(t_2)-e\mathbf{A}(t_2-t_1)\ . \tag{1b}$$

The self-energy ε, which depends on $f(\mathbf{p},t)$, and the scattering rate γ, are given by

$$\varepsilon(\mathbf{p},\tau,t)=\sum_{\mathbf{q},s=\pm1}|M_{\mathbf{q}}|^2[N_{\mathbf{q}}+\tfrac{1}{2}(1+s)]f(\mathbf{p}+\mathbf{q},t) \times\phi_s(\mathbf{p},\mathbf{q},\tau,t)\ , \tag{2a}$$

$$\gamma(\mathbf{p},\tau,t)=\sum_{\mathbf{q},s=\pm1}|M_{\mathbf{q}}|^2[N_{\mathbf{q}}+\tfrac{1}{2}(1-s)]\phi_s(\mathbf{p},\mathbf{q},\tau,t)\ , \tag{2b}$$

where $N_{\mathbf{q}}$ is the equilibrium distribution function of a phonon of momentum $\mathbf{q}$ and energy $\omega_{\mathbf{q}}$, $M_{\mathbf{q}}$ is the electron-phonon matrix element, and

$$\phi_s(\mathbf{p},\mathbf{q},\tau,t)=\exp\left[-i\int_{-\tau/2}^{\tau/2}dt_1[E(\mathbf{k}(\mathbf{p},t_1,t)+\mathbf{q}) -E(\mathbf{k}(\mathbf{p},t_1,t))-s\omega_{\mathbf{q}}]\right]\ . \tag{2c}$$

It must be stressed that in deriving Eqs. (1) and (2) from the quantum theory, all Green's functions and self-energies have been rescaled by the *exact* field-dependent scattering-free spectral function.[7] Thus ε and γ are fully field dependent though the time evolution of the momenta in ϕ_s [Eqs. (2c) and (1b)] and are not just the equilibrium self-energy and scattering rate.

We now examine the differences between Eq. (1a) and the Boltzmann equation. Since the only τ dependence of ε and γ is through the function ϕ_s, the effect of the two τ integrals in Eq. (1a) can be taken into account by defining

$$\psi_s(\mathbf{p},\mathbf{q},t',t)\equiv\int_{-2t'}^{2t'}d\tau\,\phi_s(\mathbf{p},\mathbf{q},\tau,t-t')\ . \tag{3}$$

Equation (1a) reduces to an integral form of the Boltzmann equation if ψ_s is replaced by $2\pi\delta(E(\mathbf{p}+\mathbf{q})-E(\mathbf{p})-s\omega_{\mathbf{q}})$, which is just the energy-conserving δ function of the golden-rule transition rates in semiclassical theory. Thus the quantum effects considered here are all contained in Eq. (3). The δ function is obtained if the field dependence of ϕ_s is ignored and the limits ($\pm2t'$) on the τ integral go to $\pm\infty$.

It is interesting to note that for a constant field and a parabolic band ϕ_s is independent of the field (and of its second time argument):

$$\phi_s=\exp[-i(E(\mathbf{p}+\mathbf{q})-E(\mathbf{p})-s\omega_{\mathbf{q}})\tau]\ .$$

We emphasize that this field independence of ϕ_s is also true for the exact quantum theory:[7] it is not merely a result of the weak-scattering approximation. Since there is no field dependence there can be no intracollisional field effect.[5]

The right-hand side of Eq. (3) differs from a δ function either because of the finite limits on the integral or because of the field dependence of ϕ_s arising from nonparabolic bands or time-dependent fields. Before discussing these effects we note that in Eq. (1a), as in the Boltzmann equation, the effect of scattering has been neglected *within* the self-energy and the scattering rate, consistent with the weak-coupling approximation. This effect could be included, at least phenomenologically, by multiplying ϕ_s [Eq. (2c)] by a decay term: $\phi_s\to\phi_s\exp(-\lambda_{sc}\tau/2)$, where λ_{sc} is some average scattering rate. In the Boltzmann equation this decay term would convert the δ function into a Lorentzian of width λ_{sc}. In practical calculations such broadening is almost always ignored with the assumption that as long as $\lambda_{sc}<k_BT$ its effect will be small. For semiconductors a typical collision time τ_{sc} is 10^{12}–10^{13} s, which means $\lambda_{sc}=1/\tau_{sc}\approx0.6$–$6$ meV. Room temperature corresponds to $k_BT\approx25$ meV, while for hot electrons the effective temperature can be much larger. We are only interested here in the energy scales associated with quantum corrections, so the inclusion of the decay term will make no difference. Nonetheless, λ_{sc} provides a convenient small quantity with which the quantum corrections considered below can be compared.[9]

Finite collision duration. To examine this effect we specialize to the case of a constant field $\mathbf{F}$ and a parabolic band. Then ε, γ, and ψ are independent of their second time arguments. Equation (3) gives

$$\psi_s(\mathbf{p},\mathbf{q},t)=\frac{2\sin[2t\Delta_s(\mathbf{p},\mathbf{q})]}{\Delta_s(\mathbf{p},\mathbf{q})}\ , \tag{4}$$

where

$$\Delta_s(\mathbf{p},\mathbf{q})=E(\mathbf{p}+\mathbf{q})-E(\mathbf{p})-s\omega_{\mathbf{q}}\ .$$

Note that this particular form of ψ_s is also present in equilibrium ($F=0$) and in golden-rule calculations of transition rates.[10] As long as the collision-duration time t_c (defined below) is much smaller than the time between collisions τ_{sc} the δ function is a valid approximation[11] for ψ_s. In the exponential factor in Eq. (1a) the effect of letting the limits on τ go to $\pm\infty$ is equivalent to replacing the τ integral by $\gamma(\mathbf{p}-e\mathbf{F}T'',\omega=0)$, where $\gamma(\mathbf{p},\omega)$ is the Fourier transform of $\gamma(\mathbf{p},\tau)$. Then we can thus define $\tau_{sc}(\mathbf{p},\mathbf{F})$ by the timescale on which the exponential in Eq. (1a) decays:

$$\int_0^{\tau_{sc}}dt\,\gamma(\mathbf{p}-e\mathbf{F}t,\omega=0)=1\ . \tag{5}$$

To see whether the δ-function approximation for ψ_s is valid we expand $\gamma(p,\omega)$ in powers of ω, which leads to a characteristic timescale t_c, the collision-duration time, defined as[12]

$$t_c\equiv\left[\frac{\partial}{\partial\omega}\ln\gamma(\mathbf{p}-e\mathbf{F}t,\omega)\right]_{\omega=0}\ ; \tag{6}$$

here t is the argument of γ is of order τ_{sc}. Numerical calculations based on optic-phonon scattering in Ge show that[13] $t_c/\tau_{sc}\approx10^{-3}$ over a wide range of p and F. We expect similar results to hold for other types of phonons and other semiconductors. Thus the collision-duration effect can be neglected.[14]

Now consider the factor involving ε in Eq. (1a). We can replace ψ_s [Eq. (4)] by a δ function only if $\Delta_s\gg1/T'$, i.e., for $\Delta_s\gg1/\tau_{sc}$ because of the exponential factor in Eq. (1a). Since ε contains the distribution function f this

means that the approximation is valid only if f varies slowly on the scale of $1/\tau_{sc}$, that is, if $k_BT > 1/\tau_{sc}(\mathbf{p},\mathbf{F})$. Thus $k_BT = 1/\tau_{sc}(\mathbf{p},\mathbf{F}_c)$ defines a critical field F_c above which the Boltzmann equation may not be valid. Numerical calculations show[13] that F_c is as high as 10^7 V/cm.

Next we consider the quantum effects due to field dependence of ϕ_s in Eq. (3). Since the effect of the finite limits on τ has been shown to be negligible we let them go to $\pm\infty$. Then $\psi_s(\mathbf{p},\mathbf{q},t',t)$ depends only on $t-t'$, so we drop the redundant argument to obtain

$$\psi_s(\mathbf{p},\mathbf{q},t) = \int_{-\infty}^{\infty} d\tau\, \phi_s(\mathbf{p},\mathbf{q},\tau,t) \ . \tag{7}$$

This form will be used to estimate the effects of band structure and of time-varying electric fields.

Band-structure effects. For many semiconductors the conduction band at the Γ point is not parabolic. The nonparabolicity is usually expressed by writing[15] $p^2/2m^* = E(\mathbf{p})[1+\alpha E(\mathbf{p})]$, where $\alpha = [1-(m^*/m)^2]/E_\Gamma$ and the band gap $E_\Gamma \approx 1$ eV is much larger than a typical phonon energy $\omega_{ph} \approx 40$–50 meV. To estimate the effect of a nonparabolic band we specialize to a constant field and take α to be small. Then Eq. (7) becomes

$$\psi_s(\mathbf{p},\mathbf{q},t) \simeq \int_{-\infty}^{\infty} d\tau \exp\{-i[\tau\Delta_s(\mathbf{p},\mathbf{q}) + \tfrac{1}{3}\tau^3\lambda_\alpha^3]\} = \frac{2\pi}{\lambda_\alpha}\mathrm{Ai}[\Delta_s(\mathbf{p},\mathbf{q})/\lambda_\alpha] \ , \tag{8}$$

where Ai is an Airy function, and

$$\lambda_\alpha = \left\{ \frac{\alpha(eF)^2}{4m^*} \{E(\mathbf{p}')[1+2(\hat{\mathbf{p}}'\cdot\hat{\mathbf{F}})^2] - E(\mathbf{p})[1+2(\hat{\mathbf{p}}\cdot\hat{\mathbf{F}})^2]\} \right\}^{1/3} . \tag{9}$$

Here $\mathbf{p}' = \mathbf{p}+\mathbf{q}$ and $\hat{\mathbf{p}}, \hat{\mathbf{p}}', \hat{\mathbf{F}}$ are unit vectors. Thus for $\alpha \neq 0$ the δ function is distorted with energy scale $\lambda_\alpha \propto \alpha^{1/3}F^{2/3}$. To obtain a numerical estimate for λ_α we replace $(\hat{\mathbf{p}}\cdot\hat{\mathbf{F}})^2$ and $(\hat{\mathbf{p}}'\cdot\hat{\mathbf{F}})^2$ by their angular averages, and the energy difference between the initial and the final states by the phonon energy. This gives

$$\lambda_\alpha = \left[\frac{5}{12m^*}(eF)^2 \left[\frac{\omega_q}{E_\Gamma}\right] \left[1-\frac{m^*}{m}\right]^2 \right]^{1/3} .$$

Because of the $\frac{1}{3}$ power, λ_α is insensitive to the parameter values for different semiconductors. Using $F = 10^4$ V/cm, $\omega_{ph} = 50$ meV, and $m^*/m = 0.5$ we obtain $\lambda_\alpha \approx 1$ meV. Thus λ_α is rather small and comparable to the broadening due to scattering, λ_{sc}.

Time-varying fields. Deviations from the Boltzmann equation also occur when the field varies in time. To estimate this effect we consider first a parabolic band and a slowly varying field $\mathbf{F}(t) = \mathbf{F}_0\cos\omega t$, where ω is small. Expanding the field within the exponent in Eq. (2c) to lowest nonvanishing power in ω, one obtains for ϕ exactly the same Eq. (8), but with λ_α replaced by

$$\lambda_\omega = [(eqF_0/8m^*)\omega\sin\omega t]^{1/3} \ ;$$

λ_ω increases slowly with F_0 and the frequency. To obtain a numerical estimate we compute the frequency ω that gives rise to a width of 1 meV. Using $q = (10\ \text{Å})^{-1}$ for a typical phonon wave vector and $F_0 = 10^4$ V/cm, we obtain

$$\omega = \lambda_\omega^3 \Big/ \left[\frac{eqF_0}{8m^*}\right] \approx 13 \text{ GHz} \ .$$

For $\lambda_\omega = 5$ meV, $\omega \approx 1.6\times10^3$ GHz. These are very high frequencies, so for a frequency-dependent field the quantum corrections are rather small.

Another case of interest is when a field is suddenly switched on at time $t=0$, $\mathbf{F}(t) = \mathbf{F}_0\Theta(t)$, which gives rise to transient effects. In this case

$$\phi_s(\mathbf{p},\mathbf{q},\tau,t) = \exp\left[-i\tau\Delta_s(\mathbf{p},\mathbf{q}) + \frac{ie\mathbf{q}\cdot\mathbf{F}_0}{m^*}[\Theta(t)t\tau - \tfrac{1}{2}\Theta(t+\tau/2)(t+\tau/2)^2 + \tfrac{1}{2}\Theta(t-\tau/2)(t-\tau/2)^2] \right] .$$

Deviations from the Boltzmann equation arise due to the term in square brackets, whose effect vanishes as $|t| \to \infty$. The analysis of ψ_s is somewhat complicated, but we can obtain a rough estimate for the size of the quantum effect by setting $t=0$, since the effect is largest just after the field is turned on. Then

$$\phi_s = \exp\left[-i\tau\Delta_s - i\frac{e\mathbf{q}\cdot\mathbf{F}_0}{m^*}\frac{\tau^2}{8}[\Theta(\tau)-\Theta(-\tau)] \right]$$

so that the energy of the distortion of the δ function is given by

$$\lambda_{trans} = \left| \frac{e\mathbf{q}\cdot\mathbf{F}_0}{8m^*} \right|^{1/2} \approx 8-9 \text{ meV} \ ,$$

which is comparable to λ_{sc}.

To summarize, we have shown that, for a spatially homogeneous time-dependent electric field, quantum corrections to the Boltzmann equation arise from a nonparabolic band structure or time variation of the field, and that these corrections appear through a distortion of the energy-conserving δ function in the scattering process. Our results show that the associated energy scales λ are comparable to the broadening due to scattering which is usually considered to be small at room temperature. On the other hand, at lower temperatures when $\lambda \approx k_BT$ the quantum corrections may become important.

In this paper we have treated carriers in a single unbounded energy band. Relaxing this restriction would introduce Stark ladders within a bounded band and interband processes such as Zener tunneling and impact ionization, all of which are quantum mechanical in nature. A strong magnetic field would also produce nonclassical behavior. Our result, that electrons in a simple band can be adequately described by the semiclassical Boltzmann equation even in very high electric fields, means that fu-

ture work can concentrate on the influence of intrinsically quantum effects on transport in semiconductors.

An important advantage of the integral transport equation used here, in contrast to the usual integrodifferential equation,[4–6] is that the quantum corrections appear through Eq. (3), which does not involve the distribution function. A numerical solution of Eq. (1a) with full quantum effects is therefore feasible. We have considered only electron-phonon interactions in semiconductors here, but the extension to metals and the inclusion of impurity scattering is straightforward within the weak-scattering approximation. Finally, since the field-induced quantum corrections are found to be small, it is conceivable that a careful application of the gradient-expansion technique can be used to study quantum effects in small devices where there are strong spatial variations of the electric field in the contact region, or when the scattering is strong.[9,16]

ACKNOWLEDGMENTS

We thank H. U. Baranger and C. J. Stanton for helpful discussions. This research was supported by a U.S. Office of Naval Research Grant, No. N00014-82-K-0576. One of us (J.H.D.) thanks IBM for partial support. Another of us (S.K.S.) is grateful for the hospitality of the Laboratory of Atomic and Solid State Physics at Cornell University and for the support from the University Research Grant Committee of the University of Alabama.

---

*Present address: Department of Electrical Engineering, Ohio State University, Columbus, OH 43210.

[1]A. F. J. Levi, J. R. Hayes, P. M. Platzman, and W. Weigman, Phys. Rev. Lett. **55**, 2071 (1985).

[2]M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, Phys. Rev. Lett. **55**, 2200 (1985).

[3]H. U. Baranger and J. W. Wilkins, Phys. Rev. B **30**, 7349 (1984).

[4]I. B. Levinson, Zh. Eksp. Teor. Fiz. **57**, 660 (1969) [Sov. Phys.—JETP **30**, 362 (1970)].

[5]J. R. Barker and D. K. Ferry, Phys. Rev. Lett. **42**, 1779 (1979); J. R. Barker, J. Phys. (Paris) Colloq. **42**, C7-245 (1981).

[6]V. P. Seminozhenko, Phys. Rep. **91**, 103 (1982); D. Calecki and N. Pottier, J. Phys. (Paris) Colloq. **42**, C7-271 (1981).

[7]S. K. Sarker, Phys. Rev. B **32**, 743 (1985). Our Eq. (1) is Eq. (32) of this paper with the substitutions given by Eqs. (29) and (33)–(35). Note that Eq. (32) contains a misprint: $\lambda(\mathbf{p}(T''), T-T'')$ should be $\lambda(\mathbf{p}(T''), T'', T-T'')$. To make the notation for time-dependent momenta in this reference conform with the present paper, replace $\mathbf{p}(T'')$ by $\mathbf{k}(\mathbf{p},T'',T)$ and $\mathbf{p}(T')$ by $\mathbf{k}(\mathbf{p},T',T)$ in (32), and $\mathbf{p}(t)$ by $\mathbf{k}(\mathbf{p},t,T_2)$ in (34).

[8]L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962); L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys.—JETP **20**, 1018 (1965)].

[9]The range of validity of the low-field Boltzmann equation in nondegenerate semiconductors has not yet been established. The case for metals is much clearer. It was recognized quite early by Landau that the condition $1/\tau_{sc} \ll k_B T$ should really be $1/\tau_{sc} \ll k_B T_F$, where T_F is the Fermi temperature [see R. E. Peierls, *Quantum Theory of Solids* (Clarendon, Oxford, 1955), pp. 139–142]. No such analysis exists for semiconductors, and in its absence we assume that the condition $\lambda_{sc} \approx 1/\tau_{sc} < k_B T$ limits the validity of the Boltzmann equation.

[10]See, e.g., L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), p. 283.

[11]B. K. Ridley, *Quantum Processes in Semiconductors* (Clarendon, Oxford, 1982), p. 82.

[12]Our definition of t_c is more general than that given in Ref. 8.

[13]Furrukh S. Khan and J. W. Wilkins (private communication).

[14]The condition $t_c \ll \tau_{sc}$ is violated, for example, for electrons whose energies lie near the optic-phonon frequency. It is interesting, however, to note that the violations are most severe for zero field (equilibrium). A high electric field sweeps the electrons out of this range of energies and reduces the time interval over which the condition $t_c \ll \tau_{sc}$ is not valid.

[15]H. Ehrenreich, Phys. Rev. **120**, 1951 (1960).

[16]R. E. Prange and L. P. Kadanoff, Phys. Rev. **134**, A566 (1964).