Exact solution of a transport equation for hot-electron effects in semiconductors and metals

Sanjoy K. Sarker
Department of Physics, University of Alabama, Tuscaloosa, Alabama 35487

Y. -K. Hu, C. J. Stanton,* and J. W. Wilkins
Laboratory of Atomic and Solid State Physics, Cornell University, Clark Hall, Ithaca, New York 14853-2501
(Received 22 January 1987)

We present and solve a Boltzmann equation for hot-electron transport where the elastic and inelastic collision terms are characterized by separate relaxation times τ_{el} and τ_{in}, respectively. The model is solved exactly for metals and semiconductors as a function of $R=\tau_{in}/\tau_{el}$. The exact solution is compared to two standard approximation schemes: the *sp* (or diffusive) approximation in which the Legendre expansion is truncated after the second term, and the effective-temperature model. For large R, the *sp* approximation becomes exact. In metals, both the effective-temperature model and the *sp* approximation are qualitatively correct for all R. In semiconductors, a strong dimensionality dependence is seen; in one dimension the approximations are valid, but in three dimensions they in general are not. The longitudinal (parallel to the field) and transverse projections of the exact distribution functions are calculated for semiconductors, and in a regime where the *sp* approximation is poor, the longitudinal function is seen to be approximated well by the one-dimensional solution with renormalized R.

I. INTRODUCTION

In semiconductor devices, substantial electron heating occurs[1] because of the existence of large electric fields ($\sim 10^4$ V/cm). Heating has also been observed in metals at low temperatures.[2] While most device designs are based on the drift diffusion[3] or the conservation[4] equations, a proper theoretical treatment of hot-electron transport would require solving the Boltzmann-type equation for arbitrary field strength. The Monte Carlo method (equivalent to the Boltzmann equation) has been successfully employed in both nondegenerate[5] and degenerate[6] semiconductors, but the lack of simplicity of this method could make it difficult to predict trends in the distribution function of the electrons, and hence the behavior of semiconductor devices. Solving the Boltzmann equation could supply a qualitative understanding of the underlying physics; however, this is, in general, difficult to do, and one often has to resort to approximate methods.

We are therefore interested in the validity of approximate methods. In this paper we shall look at two approximation schemes that are frequently used when the electric field **E** is assumed to be constant in space: the effective temperature model and the *sp* or diffusive approximation.[7] In the effective temperature model the electrons are assumed to be thermally distributed with an effective temperature T_{eff} which is a function of both the lattice temperature and the applied field. This approximation is admittedly extremely crude and thus has limited applicability, but it has the advantage of being extremely simple. In the *sp* approximation, the electron distribution function is first expanded in terms of Legendre polynomials

$$f(\mathbf{p})=\sum_l f_l(p)P_l(\cos\theta) \tag{1}$$

where θ is the angle between **p** and **E**. The Boltzmann equation is then converted into an infinite set of coupled differential equations[8–10] for $f_l(p)$. These equations are then truncated after $l=1$, leaving only two equations involving only f_0 and f_1. This truncation procedure is expected to be valid when the main scattering mechanism is low-energy acoustic phonon scattering. Then the scattering is quasielastic; the main effect of collisions is to randomize the momentum of the electrons gained from the field so that the spherically symmetric part of the distribution function $f_0(p)$ should dominate all other Legendre components $f_l(p)$. This randomization effect would be enhanced when either impurity scattering (as in metals at low temperature) or carrier-carrier collisions (as in highly doped semiconductors) is important. The question remains, given the complexity of the Boltzmann equation, how can we verify if the effective temperature model and the *sp* approximation are valid approximations? In addition there are more general questions that deserve an answer. For example, how does nonlinear transport differ in metals as compared to semiconductors? Are there any dimensional effects? Often, in desperation, one-dimensional transport equations are used for three-dimensional systems—is this approximation justified?

In this paper we address some of these issues, utilizing a transport equation which is sufficiently simple to be solved exactly, yet which contains much of the relevant physics. We assume throughout the entire paper that the energy band of the semiconductor is parabolic and that the Fermi surface of the metal is spherical. Within the relaxation-time approximation, we treat the quasielastic and inelastic collisions separately. For a constant field E in the negative z direction the transport equation for electrons (charge $-e$) reads

$$eE\frac{\partial f(\mathbf{p})}{\partial p_z}=-\frac{1}{\tau_{el}}[f(\mathbf{p})-f_0(p)]-\frac{1}{\tau_{in}}[f(\mathbf{p})-f_{eq}(p)]\ , \tag{2}$$

where τ_{el} and τ_{in} are the relaxation times for elastic and inelastic scattering, respectively. Eq. (2) was used by Arai[10] and by Tremblay and Vidal[11] to study electron heating and noise in metals. The first term on the right-hand side reflects the fact that in a quasielastic collision electrons lose very little energy, and thus they relax to the spherically symmetric (angular averaged) part f_0 of the unknown distribution function. For hot electrons this term includes not only impurity scattering but also low-energy acoustic phonon processes scattering as well since such a process can be considered essentially elastic when the characteristic electron energy is large compared to the energy lost. The second term represents inelastic collisions in which electrons are assumed to relax to the equilibrium distribution function $f_{eq}(p)$. For hot electrons, this term includes not only single-phonon processes (acoustic and optical emission), but also multiphonon emission cascades as discussed by Mahan.[12] Here we ignore electron-electron collisions although, in principle, one could include a term describing such collisions within the relaxation-time approximation.[13]

Assuming $\tau_{el} \ll \tau_{in}$, Arai has solved Eq. (2) for metals within the *sp approximation*, and found that the angular averaged distribution function f_0 closely resembles a Fermi-Dirac distribution function at an effective electron temperature T_{eff} which is higher than the lattice temperature T and that for high fields, $T_{eff} \propto E^{2/5}$. Furthermore, Arai showed that even when a more realistic energy-dependent inelastic collision rate was introduced into the model, within the *sp* approximation both the power law and the effective temperature model remain valid. The $\frac{2}{5}$ power law has been confirmed by recent experiments.[2,14] This remarkable result suggests that despite the crudeness of the relaxation-time approximation, the apparently simple Eq. (2) retains much of the physics of real metals.

In this paper we present an *exact* solution of Eq. (2). This allows us to test the validity of the *sp* approximation and of the effective-temperature model, and to make a comparative study of hot-electron effects in metals and nondegenerate semiconductors. We do so mainly by looking at f_0, the angular averaged part of the distribution function. To start with, we introduce a few momentum scales (or equivalently energy or time scales) that appear frequently in the analysis and which characterize the distribution function (see Table I). In the absence of the field the only scale is the thermal momentum $p_{th}=(2mk_BT)^{1/2}$ (and the Fermi momentum p_F in the case of metals). The field introduces two more relevant momentum scales: the drift momentum p_D, which is the average momentum of the distribution function, and a momentum scale p_E which is associated with the average excess energy over the thermal energy that the field provides (exact definitions are given in Sec. II).

An important parameter is the ratio of the elastic to in-

TABLE I. Frequently used symbols in this paper.

Time scales
τ_{in} relaxation time for inelastic scattering
τ_{el} relaxation time for elastic scattering
$1/\tau=1/\tau_{in}+1/\tau_{el}$ (total scattering rate)
$\tau_0=(\tau\tau_{in})^{1/2}$ (energy relaxation time)
$R=\tau_{in}/\tau_{el}$ (ratio of elastic to inelastic scattering *rates*)
Momentum scales
$p_{th}=(2mk_BT)^{1/2}$ (thermal momentum)
$p_F=(2m\varepsilon_F)^{1/2}$ (Fermi momentum)
$p_0=k_BTp_F/2\varepsilon_F$ (momentum spread at Fermi surface due to temperature T)
$p_D=eE\tau$ (average drift momentum)
$p_E=eE\tau_0/\sqrt{d}$ (momentum associated with energy due to the field)
$k=p/p_{th}$
Energy scales
$u=\begin{cases}\varepsilon/k_BT_{eff}\text{ (semiconductors, }\varepsilon\text{ measured with respect to the band edge)}\\(\varepsilon-\varepsilon_F)/k_BT_{eff}\text{(metals)}\end{cases}$
Temperature scales
T lattice temperature
$k_BT_E=\begin{cases}\frac{2}{md}(eE\tau_0)^2=\frac{2p_E^2}{m}\text{ (semiconductors)}\\\left[\frac{12\varepsilon_F}{\pi^2md}\right]^{1/2}eE\tau_0=\left[\frac{6}{\pi^2}\right]^{1/2}\frac{p_Ep_F}{m}\text{(metals)}\end{cases}$
$T_{eff}=\begin{cases}T+T_E\text{ (semiconductors)}\\(T^2+T_E^2)^{1/2}\text{(metals)}\end{cases}$

elastic scattering rates $R=\tau_{in}/\tau_{el}$. Arai expected that in metals the *sp* approximation is valid for large values of R (i.e., when elastic scattering dominates). We find that for metals the spherically averaged part of the distribution function $f_0(p)$ varies rather weakly with R, and both the *sp* approximation and effective-temperature model remain qualitatively correct over a large range of R. This does not hold for semiconductors, for which a considerably stronger dependence is found. In particular as far as Eq. (2) is concerned the effective-temperature model is not a good approximation for semiconductors. In addition, the exact spherical part of the distribution function $f_0(p)$ is seen to be strongly dependent on spatial dimensionality in semiconductors, but essentially independent of spatial dimensionality in metals. We also considered the longitudinal (parallel to the field) and transverse (perpendicular to the field) projections of the *full* distribution function for three-dimensional semiconductors. In contrast to $f_0(p)$ the projected longitudinal distribution function is found to agree fairly well with the one-dimensional distribution function provided the value of R is properly renormalized. This suggests that for qualitative studies one-dimensional models are quite useful, provided that translational symmetry exists.

This paper is organized as follows. In Sec. II Eq. (2) is solved in Fourier space and relevant momentum scales are identified. In Sec. III the spherically symmetric part of the distribution function is calculated exactly and a comparative analysis of metals and semiconductors is presented. In Sec. IV we calculate the longitudinal and transverse distribution function exactly and compare the former with the one-dimensional case. Sec. V contains our conclusions.

II. SOLUTION IN FOURIER SPACE

In this section we obtain expressions for the Fourier transforms of $f(\mathbf{p})$ and its angular average $f_0(p)$ in terms of the Fourier transform of $f_{eq}(p)$, the equilibrium distribution. We also introduce some time, momentum, and temperature scales that will be useful in discussions in later sections.

Define $g(\mathbf{r})$ as the fourier transform of $f(\mathbf{p})$

$$g(\mathbf{r})=\int \frac{d^d p}{(2\pi)^d} e^{-i\mathbf{p}\cdot\mathbf{r}} f(\mathbf{p}) , \tag{3}$$

where d is the spatial dimensionality. Then, by Fourier transforming Eq. (2), we obtain

$$g(\mathbf{r})=\frac{\tau}{1+ieEz\tau}\left[\frac{1}{\tau_{el}}g_0(r)+\frac{1}{\tau_{in}}g_{eq}(r)\right] , \tag{4}$$

where $\tau^{-1}=\tau_{el}^{-1}+\tau_{in}^{-1}$ is the total scattering rate, and g_0 and g_{eq} are Fourier transforms of f_0 and f_{eq}, respectively. We first average Eq. (4) over the angles to find $g_0(r)$. This gives

$$g_0(r)=g_{eq}(r)/[(1+R)\chi(r)-R] , \tag{5}$$

where $R=\tau_{in}/\tau_{el}$, and

$$\chi(r)=\begin{cases}[1+(p_D r)^2], & d=1 \quad (6a)\\ [1+(p_D r)^2]^{1/2}, & d=2 \quad (6b)\\ (p_D r)/\tan^{-1}(p_D r), & d=3 , \quad (6c)\end{cases}$$

where p_D is the average momentum of the distribution function

$$p_D\equiv eE\tau . \tag{7}$$

Substituting $g_0(r)$ from Eq. (5) in Eq. (4) we obtain

$$g(\mathbf{r})=\frac{g_{eq}(r)}{1+ip_D z}\frac{\chi(r)}{(1+R)\chi(r)-R} . \tag{8}$$

Once $g(\mathbf{r})$ is known, $f(\mathbf{p})$ can be obtained by Fourier inverting Eq. (8). However, $g(\mathbf{r})$ itself contains considerable information. For example, the particle density is $n=2g(0)/\hbar^d$. The current density $\mathbf{j}$ is given by

$$\mathbf{j}=-i\frac{2e}{m\hbar^d}\nabla g(r)\,|_{r=0}=\frac{ne^2\tau\mathbf{E}}{m}=\frac{nep_D}{m}\hat{\mathbf{z}} , \tag{9}$$

where m is the effective mass. This shows that the average drift momentum of the distribution is in fact equal to p_D, as asserted previously. The energy per particle w is given by

$$w=\frac{-1}{g_{eq}(0)}\frac{1}{2m}\nabla^2 g(\mathbf{r})\,|_{r=0}=w_{eq}(T)+(eE\tau_0)^2/m , \tag{10}$$

where $\tau_0=(\tau\tau_{in})^{1/2}$. τ_0 is associated with the typical length with which a particle *diffuses* before undergoing an inelastic collision[10] and $w_{eq}(T)$ is the average energy per particle when the system is at thermal equilibrium at temperature T. Note that the excess energy due to the field enters additively in Eq. (10). We define a momentum scale

$$p_E\equiv eE\tau_0/\sqrt{d} \tag{11}$$

associated with the excess energy acquired from the field. The definition implies that the excess energy acquired from the field is equal to $p_E^2 d/m$, so that p_E is a measure of the broadening of the distribution in each direction in momentum space due to the application of the field. In general, p_E is quite different from the drift momentum $p_D\equiv eE\tau$, which determines the current,

$$p_E=p_D\left[\frac{1+R}{d}\right]^{1/2} . \tag{12}$$

Thus, if R is large, p_E, and hence the excess energy of the distribution over the thermal energy, can be large even if the drift momentum (and hence the current) is small. This is because a large R means that elastic scattering dominates over inelastic scattering, and elastic scattering degrades the current but not the energy of the distribution.

We now define an effective temperature T_{eff} for a nonequilibrium distribution of electrons in an electric field as follows: The energy of the nonequilibrium distribution of

electrons is equal, by definition, to the energy these same electrons would have in thermal equilibrium with temperature T_{eff},

$$w = w_{\text{eq}}(T_{\text{eff}}) \ . \tag{13}$$

In semiconductors, $w_{\text{eq}}(T) = dk_B T/2$ which leads to $T_{\text{eff}} = T + T_E$, where T_E, the excess electron temperature is given by

$$\frac{d}{2} k_B T_E = (eE\tau_0)^2/m \ . \tag{14}$$

For metals, since the Fermi energy ε_F is much greater than all other energy scales (i.e., $\varepsilon_F \gg k_B T_E, k_B T$) we use the Sommerfeld expansion[15] for an expression to obtain the total energy of an equilibrium distribution as a function of temperature $T_{\text{eff}}^2 = T^2 + T_E^2$, where

$$k_B T_E = \left[\frac{12\varepsilon_F}{\pi^2 md} \right]^{1/2} (eE\tau_0)$$

$$= (6/\pi^2)^{1/2} \left[\frac{p_E p_F}{m} \right] , \tag{15}$$

with the Fermi momentum $p_F = (2m\varepsilon_F)^{1/2}$.

The physical properties of an electron distribution can roughly be characterized by the momentum scales we have described above. In the absence of an electric field, the only momentum scales are the thermal momentum $p_{\text{th}} = (2mk_B T)^{1/2}$ and the Fermi momentum $p_F = (2m\varepsilon_F)^{1/2}$. Because there are two independent time scales in the transport equation Eq. (2), τ_{in} and τ_{el}, the introduction of an electric field introduces two independent momentum scales: p_D which is a measure of the current, and p_E which is a measure of the additional spread of the distribution function (and hence the increase in the energy) that the field has produced.

Despite the fact that Eqs. (9) and (10) are similar to the semiclassical expression for $\mathbf{j}$ and w the electron temperature does not have a simple field dependence. This is because τ_{in} depends on the field, since as the average electron energy increases the collision rate must increase. For metals at low temperature, $\tau_{\text{in}} \propto T^3$ and $\tau \simeq \tau_{\text{el}}$ is temperature independent. This means that $\tau_0 \propto T_{\text{eff}}^{-3/2}$ which gives $T_{\text{eff}} \propto E^{2/5}$. This result has been obtained independently by Anderson, Abrahams, and Ramakrishnan[16] and by Arai[10] and has been confirmed by experiments.[2,14]

Note that the field dependence of the current Eq. (9) is much weaker since current depends on τ, not τ_0. As long as $\tau_{\text{in}} \gg \tau_{\text{el}}$, the current is ohmic. However, when τ_{in} becomes comparable to τ_{el} the mobility begins to decrease. Such a decrease is a standard feature in hot-electron transport.

III. THE SPHERICALLY SYMMETRIC DISTRIBUTION FUNCTION $f_0(p)$

In obtaining the Fourier transform of the exact distribution function $g(\mathbf{r})$ we have obtained in principle the exact function. However, comparing the exact distribution function to the two approximation schemes described in the introduction is no trivial matter. We would like a simple litmus test to provide us with an indication of whether or not these approximations are adequate. A necessary condition for an approximate distribution function to be valid is that the spherically symmetric part f_0 of the approximate distribution function be nearly the same as the f_0 of the exact distribution function. If it is not, then the approximate function is clearly inadequate. Even if the f_0's do match, we are not assured that the *full* distribution function is necessarily adequately described by the approximation. Nevertheless, knowing that f_0 is a good approximation is useful. For example, Arai[10] has shown that in the case of $R \gg 1$ the magnitude of the Legendre components $f_l \sim R^{-l/2}$, and so we would only be interested in f_0 since the other components would be negligibly small, and also that in this case, the noise spectrum of the current fluctuations can be calculated with the knowledge of f_0

$$S_I(\omega) \propto \int_{-\infty}^{\infty} d\varepsilon f_0(\varepsilon)[1 - f_0(\varepsilon - \hbar\omega)] + (\omega \to -\omega) \ . \tag{16}$$

This underscores the significance of f_0.

From Eq. (5), the Fourier transform of $f_0(p)$ can be written as

$$g_0(r) \equiv g_{\text{eq}}(r)\psi(r) \ , \tag{17}$$

with

$$\psi(r) = \frac{1}{(1+R)\chi(r) - R} \ , \tag{18}$$

where $\chi(r)$ is given by Eqs. (6). Thus, $\psi(r)$ gives an indication of how the field distorts the distribution function from the equilibrium configuration.

Using the method of Arai[10] it can be shown that the *sp* approximation corresponds to expanding the denomination in (18) to order r^2 so that

$$\psi_{sp}(r) = \frac{1}{1 + (p_E r)^2} \ . \tag{19}$$

This works best when $R \gg 1$, since $\psi(r)$ will be negligibly small when the expansion of $\chi(r)$ starts to break down. Note that the *sp* approximation is essentially independent of dimension, apart from a simple dependence through p_E. Furthermore $\psi_{sp}(r)$ and hence $f_0^{(sp)}$ are functions of τ_0 alone. While this is correct for one dimension (the *sp* approximation is exact for $d = 1$ since substituting the exact expression for $\chi(r)$ [Eq. (6a)] into Eq. (18) yields Eq. (19)) for $d > 1$ the true distribution function will depend on τ_{el} and τ_{in} separately. The parameter $R = \tau_{\text{in}}/\tau_{\text{el}}$ is a convenient measure of the deviations from the *sp* approximation, which is expected to be good when R is large, i.e., when elastic scattering dominates. Note also that $\psi_{sp}(r)$ in Eq. (19) decays with a characteristic length of $1/p_E$, in all dimensions. This can be shown to be true even when the *sp* approximation is not expected to hold (i.e., when R is not very much greater than 1).

The distribution function f_0 may be expressed as a one-dimensional integral in two ways. The first exploits the fact that the inverse fourier transform of a product is a convolution, and so the $d = 3$ distribution function is

$$f_0^{(d=3)}(p) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp' \frac{p'}{p} f_{\text{eq}}(p') \hat{\Psi}(p - p') \ , \tag{20}$$

where

$$\hat{\Psi}(p-\hat{p})=\int_{-\infty}^{+\infty} e^{i(p-p')r}\psi(r)dr \ . \tag{21}$$

For $d=1$ the expression is the same as (20), except that it lacks the factor p'/p

$$f_0^{(d=1)}(p)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} dp' f_{eq}(p')\hat{\Psi}(p-p') \ . \tag{22}$$

Since the characteristic decay length of $\psi(r)$ is $1/p_E$, the characteristic decay length of its Fourier transform $\hat{\Psi}(p)$ is p_E. An inspection of Eqs. (20) and (22) indicates that $\hat{\Psi}$ serves to smear out the equilibrium distribution function over the momentum scale p_E.

In the second way, $f_0(p)$ is found by directly inverse Fourier transforming $g_0(r)$, a step for which $g_{\rm eq}(r)$ is needed. For semiconductors we have

$$g_{\rm eq}(r)=\frac{n\hbar^d}{2}e^{-mk_BTr^2/2} \ . \tag{23}$$

For metals it is difficult to obtain an exact analytic expression for $g_{\rm eq}$. However at low temperatures if we linearize ε_p about the chemical potential $\varepsilon_F=p_F^2/2m$ we have

$$(\varepsilon_p-\varepsilon_F)/k_BT\approx(p-p_F)/p_0 \ ,$$

where the momentum scale $p_0=(k_BTm/p_F)$ is the momentum over which the Fermi distribution is smeared due to the temperature T. Then in one dimension

$$g_{\rm eq}(r)=\left[\frac{\sin(p_Fr)}{\pi r}\right]\left[\frac{\pi r p_0}{\sinh(\pi r p_0)}\right] , \tag{24}$$

where the first factor on the right-hand side is the Fourier transform at zero temperature. In general, the three-dimensional form of $g_{\rm eq}$ can be obtained from the one-dimensional result from the relation

$$g_{\rm eq}(r,d=3)=-\frac{1}{2\pi r}\frac{\partial}{\partial r}g_{\rm eq}(r,d=1) \ . \tag{25}$$

A. Metals

The major difference between metals and semiconductors is that there is an extra energy scale associated with metals, namely, the Fermi energy ε_F which is much larger than k_BT or k_BT_E. To illustrate this point, we shall use Eq. (20) to calculate $f_0(p)$ for $p\simeq p_F$. Here p'/p is essentially unity in the region where the integrand is significant, since $\hat{\Psi}(p-p')$ has a decay length of $p_E \ll p_F$. Hence, the factor p'/p may be set equal to one, and the resulting integral is seen to be of the same form as the one-dimensional result [Eq. (22)]. Therefore, in metals, essentially the only difference between the one-dimensional and three-dimensional results comes from the difference in the factor $\psi(r)$. Note also that because $f_{\rm eq}(p)$ has particle-hole symmetry about p_F and because $\hat{\Psi}$ is an even function, by Eq. (22), $f_0(p)$ must also have particle-hole symmetry about p_F; i.e.,

$$f_0(p_F+|\Delta p|)=1-f_0(p_F-|\Delta p|) \ .$$

In the *sp* approximation, $\psi(r)=[1+(p_Er)^2]^{-1}$ in both one and three dimensions. Then,

$$\hat{\Psi}(p-p')=(\pi/p_E)\exp(-|p-p'|/p_E) \ .$$

If we specialize to the $T=0$ case, the integral in Eq. (22) can be done analytically

$$f_0^{(sp)}(p)=\begin{cases}\frac{1}{2}e^{-(p-p_F)/p_E}, & p>p_F\\ 1-\frac{1}{2}e^{-(p_F-p)/p_E}, & p<p_F \ .\end{cases} \tag{26}$$

Obviously $f_0(p)$ in the *sp* approximation differs from the Fermi function. For a comparison, we linearize the single-particle energy $\varepsilon_p=p^2/2m$ about the Fermi level $\varepsilon_F=p_F^2/2m$. Then using the definition of the effective temperature and defining the reduced energy variable

$$u=(\varepsilon_p-\varepsilon_F)/k_BT_{\rm eff}\approx(p-p_F)p_F/mk_BT_{\rm eff}$$

we have

$$f_0^{(sp)}(u)=\begin{cases}\frac{1}{2}e^{-\alpha u}, & u>0\\ 1-\frac{1}{2}e^{-\alpha|u|}, & u<0 \ ,\end{cases} \tag{27}$$

where $\alpha=\sqrt{6}/\pi\approx 0.78$. For $|u|\ll 1$, $f_0(u)\sim\frac{1}{2}(1-\alpha u)$. The Fermi function $f_F(u)=(e^u+1)^{-1}$, on the other hand, behaves as

$$f_F(u)\sim\begin{cases}e^{-u}, & u\gg 1 \quad (28a)\\ \frac{1}{2}(1-\frac{1}{2}u), & |u|\ll 1 \quad (28b)\\ 1-e^{-|u|}, & u\ll -1 \ . \quad (28c)\end{cases}$$

Despite these differences, the *sp* approximation for $f_0(p)$ is qualitatively similar to the Fermi function as seen from Fig. 1(b). When the lattice temperature $T\neq 0$ it is difficult to obtain an analytic formula for $f_0(p)$. However, linearizing about the Fermi energy, we obtain, in the *sp* approximation,

$$f_0^{(sp)}(p)=\frac{\pi\zeta e^{-v}}{2\sin(\pi\zeta)}-\sum_{n=1}^{\infty}\frac{(-1)^n e^{-n\nu/\zeta}}{1-(n/\zeta)^2} \quad \text{for } p>p_F \ , \tag{29}$$

where $\nu=(p-p_F)/p_E$ and $\zeta=p_0/p_E=\alpha T/T_E$. At $T=0$ Eq. (29) reduces to Eq. (26), as it should.

The exact spherically symmetric distribution function $f_0(p)$ for finite values R and for various values of T were obtained numerically. These results are shown in Figs. 1(a) and 1(b) as functions of reduced energy $u=(\varepsilon-\varepsilon_F)/k_BT_{\rm eff}$ along with the *sp* approximation results. As expected the *sp* approximation is quite good for large $R(>10)$. For smaller R, there is a deviation for the *sp* results, but the dependence on R is not dramatic. It is also interesting to note that even for smaller R the *sp* approximation becomes better as the lattice temperature T is increased.

To test the effective temperature model we have plotted

$\ln(1/f_0-1)$ versus the reduced energy [Figs. 1(c) and (d)]. If $f_0=(e^u+1)^{-1}$ we should obtain a straight line. Again the agreement is quite good for large R and finite temperature $T\approx T_E$. For lower values of R and T the effective temperature model becomes worse, but the deviations from the exact result are still not very large for $u<4$. If an approximate function is to be used to perform a noise calculation using Eq. (16), it is crucial that it must not deviate significantly from the exact result in the region $|u|<4$, since the major contribution to the integral comes from this region.

B. Semiconductors

For semiconductors it is convenient to scale the momentum by the thermal momentum by defining $\mathbf{k}=\mathbf{p}/p_{\rm th}$ where $p_{\rm th}=(2mk_BT)^{1/2}$. Define the dimensionless distribution function

$$f_0(k)=\frac{2}{n}\left[\frac{p_{\rm th}}{2\pi\hbar}\right]^d f_0(p)\ , \tag{30}$$

where the prefactor is chosen so that $F(k)$ is normalized

$$\int F_0(k)d^dk=1\ . \tag{31}$$

We obtain $f_0(k)$ by inverse Fourier transforming $g_0(r)$. Using Eq. (17) for $g_0(r)$ and Eq. (23) for $g_{\rm eq}(r)$, we have in three-dimensions

$$f_0(k)=\frac{b^2}{2k\pi^2}\int_0^\infty dx\frac{xe^{-b^2x^2/4}\sin(kbx)}{(1+R)(x/\tan^{-1}x)-R}\ , \tag{32}$$

where the parameter

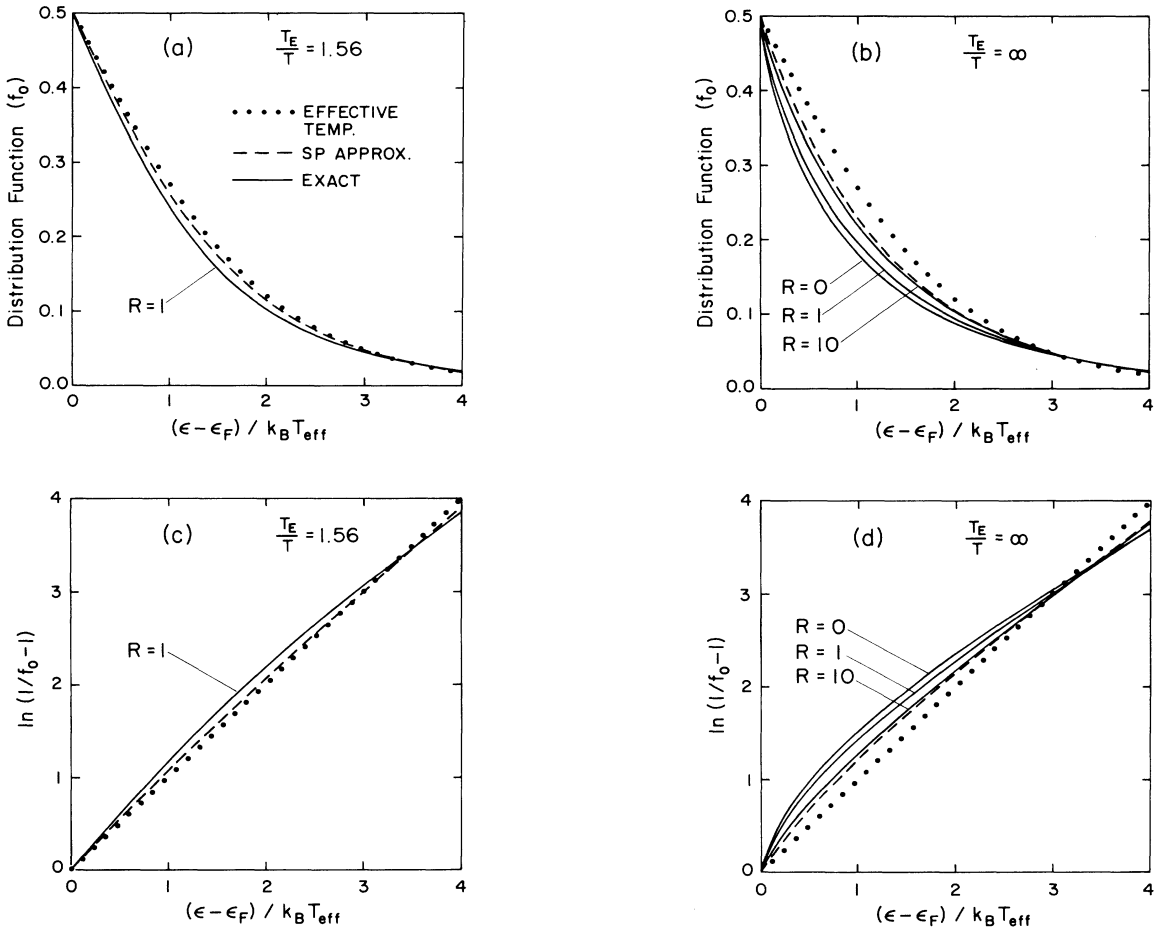

FIG. 1. The angular averaged distribution functions f_0 for metals vs the reduced energy $u=(\varepsilon-\varepsilon_F)/k_BT_{\rm eff}$ for various values of $R=\tau_{\rm in}/\tau_{\rm el}$ (solid lines), together with the distribution functions for the effective temperature model (dotted lines) and the *sp* approximation (dashed lines). (a) and (c) $T_E/T=2\sqrt{6}/\pi=1.56$; (b) and (d) $T_E/T=\infty$. In (a) and (b), f_0 vs u is plotted, while in (c) and (d), $\ln(1/f_0-1)$ is plotted vs u to facilitate comparison of the results with the effective temperature model. For the case of $T_E/T=1.56$ [(a) and (c)], the $R=0$ and $R=1$ curves are virtually identical, and the $R=10$ curve is indistinguishable from the *sp* approximation curve. At $T_E/T=\sqrt{6}/\pi=0.78$ *all* the curves are essentially indistinguishable. For $\varepsilon<\varepsilon_F$, $f_0(u)=1-f_0(-u)$.

$$b = p_{\rm th}/p_D = \left[\frac{4(1+R)}{d} \frac{T}{T_E} \right]^{1/2} . \quad (33)$$

Similarly, in one dimension one has

$$f_0(k) = \frac{b}{\pi} \int_0^\infty dx \frac{e^{-b^2x^2/4} \cos(kbx)}{(1+R)(1+x^2)-R} . \quad (34)$$

The *sp* approximation corresponds to an x expansion in the denomination of Eq. (32) to order x^2 which gives $1+\frac{1}{3}(1+R)x^2$. Such an expansion is valid for $x<1$. The exponential cuts off the integral at $x^2 \sim 4/b^2$. Therefore the small x expansion is good if $b^2/4 \gg 1$, i.e., if

$$3T(1+R)/T_E \gg 1 . \quad (35)$$

Thus, in general, the *sp* approximation is valid for low fields and large R.

For the *sp* approximation the integral (32) can be evaluated analytically with the result

$$f_0^{(sp)}(k) = \frac{1}{2k\pi} s^2 e^{-k^2} [W(s-k) - W(s+k)] , \quad (36)$$

where $s = (T/T_E)^{1/2}$ and W is related to the error function

$$W(x) = e^{x^2}[1-\mathrm{erf}(x)] . \quad (37)$$

Similarly doing the integral in Eq. (34) we have for one dimension

$$f_0(k) = \tfrac{1}{2} s e^{-k^2} [W(s-k) + W(s+k)] . \quad (38)$$

The expression for the effective temperature distribution function is

$$f_0^{(\rm eff)}(k) = \frac{1}{[\pi(1+1/s^2)]^{d/2}} \exp\left[-k^2 \Big/ \left[1+\frac{1}{s^2} \right] \right] . \quad (39)$$

One feature of Eqs. (36) and (38) is the existence of hot-electron tails. Since the error function approaches -1

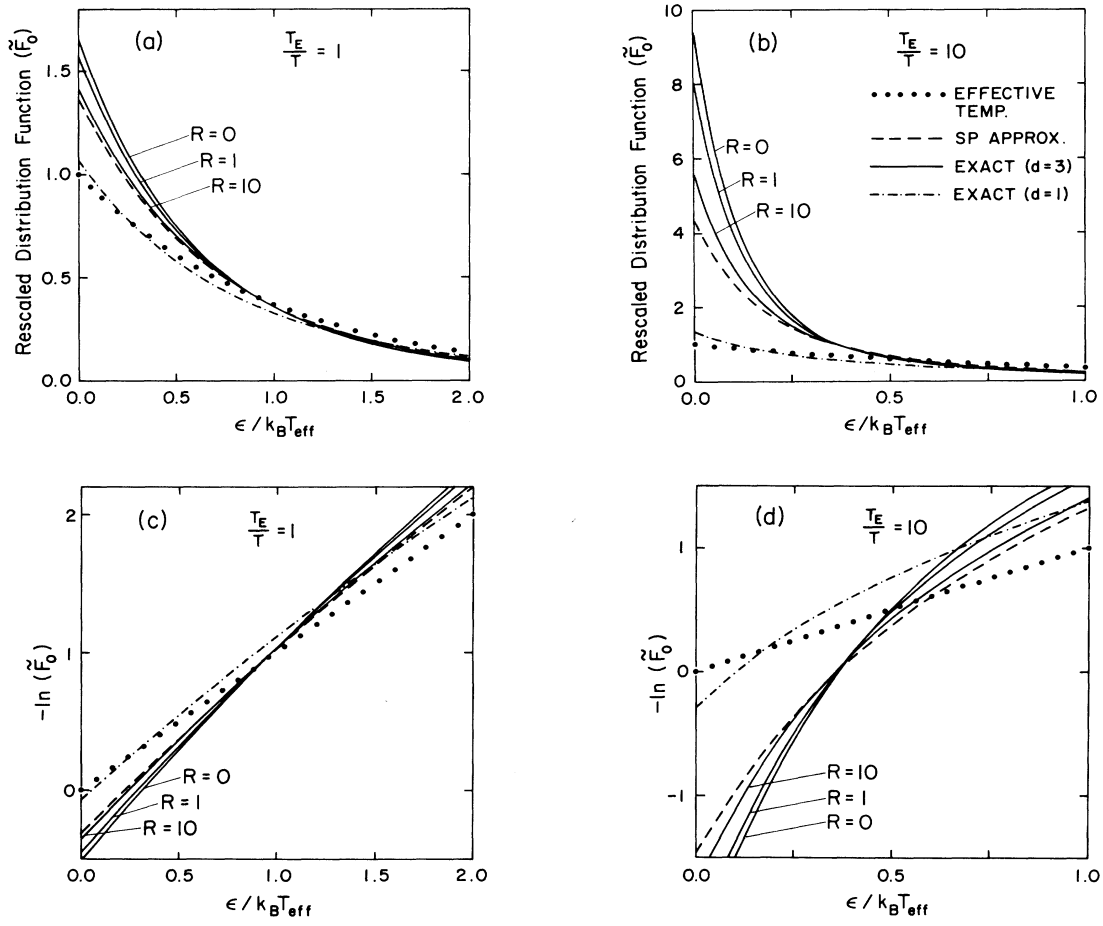

FIG. 2. The angular averaged distribution functions (F_0) for semiconductors vs reduced energy $u = \varepsilon/k_B T_{\rm eff}$ for various values of $R = \tau_{\rm in}/\tau_{\rm el}$ (solid lines for $d=3$, dashed-dotted lines for $d=1$), together with the distribution functions for the effective temperature model (dotted lines) and the *sp* approximation (dashed lines). We have plotted $\tilde{F}_0 = [\pi(1+T_E/T)]^{d/2} F_0$ so that the effective-temperature models for both one and three dimensions are e^{-u}. (a) and (c) $T_E/T = 1$; (b) and (d) $T_E/T = 10$. In (a) and (b), the $\tilde{F}_0$ is plotted vs u, while in (c) and (d), $-\ln(\tilde{F}_0)$ vs u is plotted to facilitate comparison of results with the effective-temperature model.

when its argument approaches $-\infty$, for large k we have

$$F_0(k) \sim \begin{cases} \frac{1}{k\pi^2} s^2 e^{-2sk+s^2}, & d=3 \quad (40a) \\ s e^{-2sk+s^2}, & d=1 \ . \quad (40b) \end{cases}$$

Therefore the exponent falls off rather slowly (linearly) at large k. Note also that f_0 falls off more slowly as s decreases, i.e., as the field increases, since $s \propto E^{-1}$. These tails may be suppressed if a more realistic scattering mechanism than the constant relaxation-time approximation is used.

To compare the exact results with the *sp* approximation and the effective-temperature model, Eq. (32) was evaluated numerically for various values of R. The distribution functions are plotted versus reduced energy $u = \varepsilon / k_B T_{\text{eff}}$ in Figs. 2(a) and 2(b) for $T/T_E = 1$ and $T/T_E = \frac{1}{10}$, respectively. In these figures, the distributions have been multiplied by a factor $[\pi(1+1/s^2)]^{d/2}$ so that the effective temperature model is given by e^{-u} for both one and three dimensions. To test the effective temperature model, we have plotted $-\ln(f_0)$ versus u in Figs. 2(c) and 2(d). In these graphs, the effective-temperature distribution functions are straight lines of slope one.

The first feature to note is that the effective-temperature model approximates the exact one-dimensional result rather well. However, this is certainly not the case in three-dimensions—the approximation is inadequate at low fields ($T = T_E$) and becomes worse when the electric field increases to $T_E = 10T$. The *sp* approximation is much better at the low electric fields, but deviations from the exact results are large at high fields, even at $R = 10$, when one would expect the *sp* approximation to be valid. These results differ significantly from the case of metals where (a) exact results do not deviate much from the effective temperature or the *sp* approximation, even at very large electric fields and (b) f_0 is essentially independent of the dimensionality of the system.

Part of the reason for the discrepancy between the effective temperature model and the exact results may be attributed to the use of a constant relaxation time in the inelastic term in Eq. (2). In some materials, there might be a strong increase in the rate of scattering in the high-energy region where our model predicts a significant number of electrons. Therefore the constant relaxation-time approximation underestimates scattering of these high-energy electrons, giving rise to a hot tail in the distribution. This tail contributes heavily to the second moment of the distribution, and this results in an artificially high electron temperature. An inspection of Figs. 2(c) and 2(d) indicates that a lower effective temperature would fit the exact distribution functions better.

In the region $p^2/(2mk_B T_{\text{eff}}) \lesssim 1$, the effective-temperature model and the *sp* approximation in three-dimensional semiconductors are inadequate for a fundamentally different reason—namely, because the phase space available for scattering in three-dimensional semiconductors is different from metals and one-dimensional semiconductors. In semiconductors, the phase space available for scattering increases dramatically with increasing dimension. For metals this increase is suppressed because the Fermi energy is so large that most of the scattering occurs only on the Fermi sphere. Mathematically, this can be seen by observing the factors p'/p and $\hat{\Psi}(p-p')$ in Eq. (14). At low fields $\hat{\Psi}(p-p')$ is essentially a δ function, so $p'/p \simeq 1$ and so is of no consequence. However, for large fields $\hat{\Psi}(p-p')$ decays over a longer range and so the effect of the factor p'/p in semiconductors is to increase the number of low-energy electrons. Moreover, the extent of this increase is sensitive to the exact form of $\hat{\Psi}(p-p')$. Hence in three-dimensional semiconductors, the *sp* approximation result differs significantly from the exact results because different $\hat{\Psi}(p-p')$ were used in each case. The effective-temperature model simply lacks this increase in the number of low-energy electrons, and so is the worst approximation.

For metals, $p'/p \simeq p_F/p_F = 1$ in the region of interest so this factor plays no role. In one-dimension semiconductors, (p'/p) is missing altogether. So for these cases, the effective-temperature model and the *sp* approximation are expected to be similar, and to give good approximations to the exact results (of course, in one-dimension, the *sp* approximation is the exact result).

IV. LONGITUDINAL AND TRANSVERSE PROJECTION OF THE DISTRIBUTION FUNCTIONS IN SEMICONDUCTORS

In the last section, we saw that the spherically symmetric part of the distribution function f_0 for three-dimensional semiconductors was not accurately described by either the effective temperature model or the *sp* approximation, particularly at small R. This is because the full high field, low R distribution function is strongly skewed in the field direction, and so angular averaging the function cannot be expected to give a full description. One could of course calculate functions for higher l values, but a prohibitively large number Legendre components may be needed to accurately describe the distribution function. (According to Hammar[9] 100 l components may be needed.)

This naturally leads us to attempt to find another method of simply characterizing the distribution function. To do this, we project the distribution function onto the field direction by integrating out the transverse momenta. We then compare the projected function with the full one-dimensional distribution function. In the case $R = 0$, when separation of the variables p_x and p_y from the differential equation (2) leaves the one-dimensional version of the equation, the longitudinal projected distribution function is exactly equal to the one-dimensional distribution function. Similarities between the two functions persist even when R is made nonzero, with best agreements occurring with small R. This contrasts with the *sp* approximation, which works well with large R. We can also project the distribution onto the transverse axis by integrating out p_z and, say, p_x. Together the longitudinal and transverse distribution functions provide insight into the actual shape of the full distribution.

A. Longitudinal projected distribution function

It is convenient to scale momentum by $p_{\rm th}$, defining $k_{||}=p_z/p_{\rm th}$ and $k_\perp=p_\perp/p_{\rm th}$. The longitudinal projected distribution function is then defined as

$$F_{3-d}(k_{||})=\frac{p_{\rm th}}{4n\hbar^3\pi^3}\int dp_x\int dp_y f(\mathbf{p})\ , \tag{41}$$

where, as before, the prefactor is chosen to normalize $F(k_{||})$, i.e., $\int F_{3-d}(k_{||})dk_{||}=1$.

Using the Fourier transform of $f(\mathbf{p})$ as given by Eq. (8) we can write (41) as

$$F_{3-d}(k_{||})=\frac{p_{\rm th}}{4n\hbar^3\pi^3}\int dp_x\int dp_y\int d^3r\, g(\mathbf{r})e^{i\mathbf{p}\cdot\mathbf{r}}\ . \tag{42}$$

The integrals over p_x and p_y simply give δ functions which can be used to do the x and y integrals. Then, after a change of variables $p_Dz=x$ we have

$$F_{3-d}(k_{||})=\frac{b}{2\pi}\int_{-\infty}^{\infty}dx\frac{1-ix}{1+x^2}\frac{e^{ik_{||}bx-x^2b^2/4}}{[1+R-(R\tan^{-1}x)/x]}\ . \tag{43}$$

where, as before, $b=(p_{\rm th}/p_D)=[\frac{4}{3}(1+R)T/T_E]^{1/2}$. For a comparison, in terms of the same variables, the full one-dimensional distribution function is given by

$$F_{1-d}(k_{||})=\frac{b}{2\pi}\int_{-\infty}^{\infty}dx\frac{1-ix}{1+x^2}\frac{e^{ik_{||}bx-x^2b^2/4}}{[1+R-R/(1+x^2)]}\ . \tag{44}$$

Equation (44) can be integrated, and the result is

$$F_{1-d}(k_{||})=\frac{se^{k_{||}^2}}{2}\left[\left[1+\frac{1}{\sqrt{1+R}}\right]W(s-k_{||})+\left[1-\frac{1}{\sqrt{1+R}}\right]W(s+k_{||})\right]\ . \tag{45}$$

Comparing (43) with (44) we see that the only difference between the two equations is that $(R\tan^{-1}x)/x$ appears in the denominator of Eq. (43), while in Eq. (44) the term $R/(1+x^2)$ appears in the corresponding place. Thus we see mathematically that the two are equal when $R=0$. For $b\gg1$, the exponential cuts off the integrand

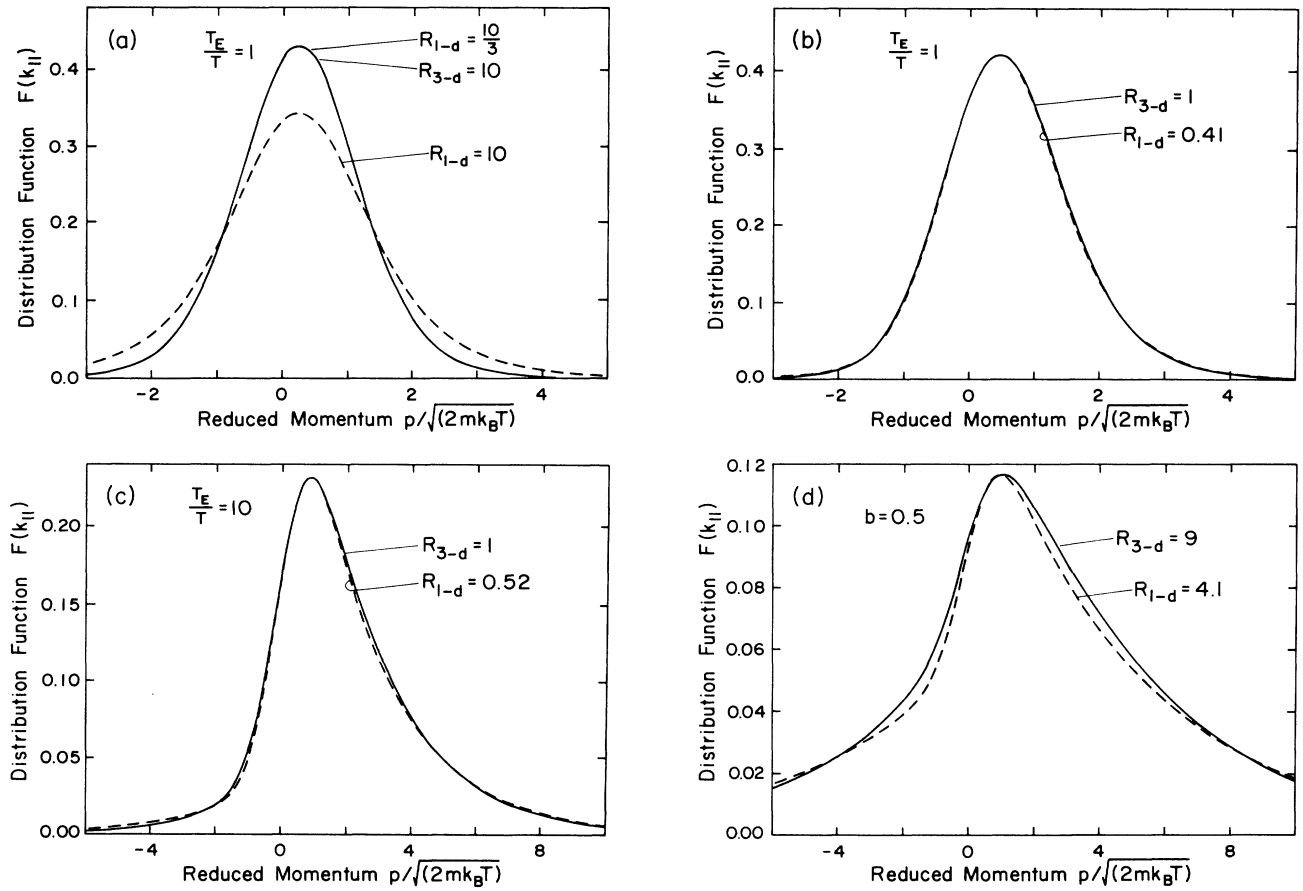

FIG. 3. Comparison of three-dimensional projected (longitudinal) distribution functions $F(k_{||})$, and one-dimensional results. These are plotted vs reduced momentum $p/(2mk_BT)^{1/2}$. The three-dimensional results are denoted by solid lines while the one-dimensional results are denoted by dashed lines. (a) $T_E/T=1$. The $R_{3-d}=10$ and $R_{1-d}=10/3$ are indistinguishable. Plotted for comparison is the $R_{1-d}=10$ result; (b) $T_E/T=1$ with $R_{3-d}=1$ and $R_{1-d}=0.41$ (again, indistinguishable); (c) $T_E/T=10$ with $R_{3-d}=1$ and $R_{1-d}=0.52$; (d) $b=0.5$ with $R_{1-d}=4.1$ and $R_{3-d}=9$. $b=p_{\rm th}/p_D$ is related to T_E/T by $b=[4(1+R)T/dT_E]^{1/2}$.

at $x \ll 1$ and so we can keep only the lowest-order terms in an expansion in x. This expansion gives factors in the denominator of $1+Rx^2$ for $d=1$, and $1+Rx^2/3$ for $d=3$; hence, we expect that for $b \gg 1$, the distributions will look very similar when $R_{3-d}=3R_{1-d}$. Physically, a larger R_{3-d} is expected because elastic scattering is much more efficient in back-scattering electrons in one dimension than in three dimensions.

For $b<1$, $R_{3-d}=3R_{1-d}$ does not give the best fit anymore. We "empirically" fit $F_{1-d}(k_{\|})$ to $F_{3-d}(k_{\|})$ by using functions with the same b (i.e., current) and by picking the value R_{1-d} so that the peak of $F_{1-d}(k_{\|})$ matched that of $F_{3-d}(k_{\|})$. We used $R_{3-d}=1$, and b was chosen so that we reproduced the conditions $T_E=T$ and $T_E=10$. The results are shown in Fig. 3(b) and 3(c). Lastly, in Fig. 3(d) we show $F_{1-d}(k_{\|})$ and $F_{3-d}(k_{\|})$ for the case of $R_{3-d}=9$ and its "empirical" fit $R_{1-d}=4.1$, for $b=0.5$, to show that while similarities between the two are less pronounced at higher R, they still persist.

B. Transverse projected distribution function

The transverse projected distribution function $F_{\perp}$ is defined by integrating out p_z and either p_x or p_y. This projected distribution function gives a picture of the motion of electrons transverse to the electric field. Should electrons gain more and more transverse kinetic energy in a semiconductor device, a possible consequence is that these electrons will spill out of potential barriers, degrading the performance of the device. For this reason, $F_{\perp}$ is worth looking at. In terms of $k_{\perp}=p_{\perp}/p_{\rm th}$, where $p_{\perp}$ is the remaining transverse momentum

$$F_{\perp}(k_{\perp})=\frac{b}{2\pi}\int_{-\infty}^{\infty}dx\frac{e^{ik_{\perp}bx-x^2b^2/4}}{1+R-(R\tan^{-1}x)/x}\ . \tag{46}$$

When $R=0$, this gives a Gaussian: $F_{\perp}(k_{\perp},R=0)=(e^{-k_{\perp}^2})/\sqrt{\pi}$. For $R>0$, $F_{\perp}(k_{\perp})$ deviates from a Gaussian as electrons are scattered elastically to the high momentum region.

For $b \gg 1$ we can expand $\tan^{-1}x/x$ in the denominator to lowest order in x

$$F_{\perp}(k_{\perp})\simeq\frac{b}{2\pi}\int_{-\infty}^{\infty}dx\frac{e^{ik_{\perp}bx-x^2b^2/4}}{1+\frac{R}{3}x^2}$$

$$=\tfrac{1}{2}s'e^{k_{\perp}^2}[W(s'-k_{\perp})+W(s'+k_{\perp})]\ , \tag{47}$$

where $s'=(b/2)\sqrt{3/R}$. Note the similarity with Eq. (38), which describes the angular averaged one-dimensional distribution function. There, $s=(b/2)\sqrt{\times 1/(1+R)}$. If we compare the projected transverse distribution function with an angular averaged one-dimensional distribution function with the same b (i.e., current), we see that in order for these to be equal [within the approximation made above to obtain Eq. (47)], we must have $R_{3-d}=3(R_{3-d}+1)$. This again shows that backscattering in the one-dimensional system is more efficient than "side-scattering" in three dimensions. In Figs. 4(a) and 4(b) we have plotted $F_{\perp}(k_{\perp})$ for the same parameters used in Figs. 3(a)—3(c).

We have seen that the projected distribution functions look very much like the exact one-dimensional distribution functions. Accordingly, either in order to simplify the problem or because of computer time and storage limitations, it may be expedient to use a one-dimensional Boltzmann equation to describe a three-dimensional system. It is not entirely clear whether this is useful, particularly since we have seen a strong dimensional dependence in f_0 in the last section. Here, the strong resemblance between the one-dimensional and the projected three-dimensional distribution functions can be used in some cases to justify this simplification.

V. CONCLUSION AND SUMMARY

In this paper we have solved a transport equation exactly to study hot-electron effects in metals and semiconductors. While the equation itself is approximate in that the collisions are treated within the relaxation-time approximation, much of the important physics is considered by

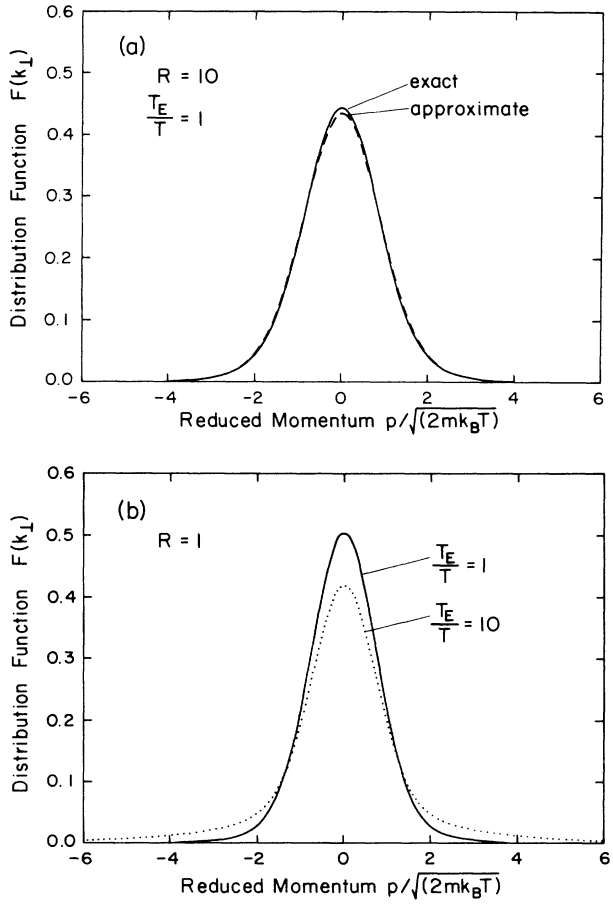

FIG. 4. Projected (transverse) distribution functions $F_{\perp}(k_{\perp})$ vs reduced momentum $p/(2mk_BT)^{1/2}$. (a) $R=10$ and $T_E/T=1$. The exact result is the solid line, while the approximate expression [Eq. (38)] is plotted with the dashed line (b) $R=1$ with $T_E/T=1$ (solid line) and $T_E/T=10$ (dotted line).

using different relaxation times for inelastic and quasi-elastic scattering. This equation was previously used and solved approximately by Arai to study hot-electron effects in metals, and his results have been confirmed by recent experiments. Our results show that both the *sp* approximation scheme used by Arai and the effective temperature model give qualitatively correct results for metals, and the results are essentially independent of spatial dimensions. In semiconductors, however, the distribution function depends strongly on spatial dimensionality, with discrepancies between the exact solution and both the effective temperature and *sp* approximation growing with increasing dimensionality. We concluded that such a dependence (which does not depend on the relaxation time approximation) is attributed to the increase in phase space available for scattering in semiconductors. We then investigated an alternative means of simply characterizing the three-dimensional semiconductor distribution function. We calculated the longitudinal and transverse projections of the distribution functions and found that both were quite well approximated by the one-dimensional semiconductor distribution function with renormalized parameters. We concluded that this might provide justification in some instances of using a one-dimensional Boltzmann equation to describe a three-dimensional system. However, there are still other interesting questions to investigate, such as a more thorough investigation of noise spectra, effects of intervalley scattering and effects of electron-hole interactions.

ACKNOWLEDGMENTS

This work was supported by the U. S. Office of Naval Research and the Semiconductor Research Corporation. In addition, the following personal support is gratefully acknowledged: From the University of Alabama a grant for S.K.S.; partial support from IBM for Y.-K. H.; and partial support from the Fannie and John Hertz Foundation for C. J. S. We also thank N. S. Wingreen and A. P. Smith for useful conversations.

---

*Present address: Coordinated Science Laboratory, University of Illinois, 1101 West Springfield Avenue, Urbana, IL 61801-3082.

[1]H. L. Grubin, in *Physics of Nonlinear Transport in Semiconductors,* edited by D. K. Ferry, J. R. Barker, and C. Jacoboni (Plenum, New York, 1980).

[2]M. L. Roukes, M. R. Freeman, R. S. Germain, R. C. Richardson, and M. B. Ketchen, Phys. Rev. Lett. **55**, 422 (1984).

[3]S. M. Sze, *Physics of Semiconductor Devices,* 2nd ed. (Wiley, New York, 1981), p. 657.

[4]K. Blotekjaer, IEEE Trans. Electron Devices **ED-17**, 38 (1970).

[5]C. Jacoboni and L. Reggiani, Rev. Mod. Phys. **55**, 645 (1983).

[6]P. Lugli, Ph.D. thesis, Colorado State University, 1985; and M. Artaki, Ph.D. thesis, University of Illinois, 1987.

[7]K. Seeger, *Semiconductor Physics,* 2nd ed. (Springer-Verlag, Berlin, 1982), p. 49.

[8]E. Conwell, *Advances in Research and Applications* (Academic, New York, 1967), Suppl. 9.

[9]C. Hammar, J. Phys. C **6**, 70 (1973).

[10]M. R. Arai, Appl. Phys. Lett. **42**, 906 (1983); and Ph.D. thesis, Cornell University, 1983.

[11]A.-M.S. Tremblay and F. Vidal, Phys. Rev. B **25**, 7562 (1982).

[12]G. Mahan, J. Appl. Phys. **58**, 2242 (1985).

[13]N. S. Wingreen, C. J. Stanton, and J. W. Wilkins, Phys. Rev. Lett. **57**, 1084 (1986).

[14]G. J. Dolan and D. D. Osheroff, Phys. Rev. Lett. **43**, 721 (1979).

[15]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), p. 49.

[16]P. W. Andersson, E. Abrahams, and T. V. Ramakrishnan, Phys. Rev. Lett. **43**, 718 (1979).