P. J. Price

The Theory of Hot Electrons

Abstract: This is a survey of methods of analysis of the hot-electron phenomenon in semiconductors. The earliest method depended on three basic assumptions: smallness of the deviation of $f(\mathbf{p})$, the carrier distribution function, from $f_0(E(\mathbf{p}))$, the distribution in energy E; the conventional relaxation-time relation between $f - f_0$ and df_0/dE; and smallness of df_0/dE. More general methods are associated with giving up, successively in the reverse order, these assumptions. Procedures for obtaining and solving equations involving f_0 only, based on a new approach due to Levinson, are developed. An inherently precise method for calculating f and related treatment of differential mobility, which requires computer implementation and which has recently come into use, is expounded. Test calculations for the case of n-germanium are reported.

1. Introduction

This is a survey of methods of theoretical analysis of the hot-electron phenomenon in semiconductors.[1] That is, of the effects of substantial displacement of the electron system from thermal equilibrium with the lattice, by an electric field ($\mathcal{E}$) driving a current (or by light or other energy-bearing disturbances). Interest and progress in the subject have been considerably augmented by the recent work on dynamical instabilities in these systems.

In the conditions which normally apply,[2] one may use a quasiclassical specification of the system of carriers (itinerant electrons or holes), with nondegenerate statistics. That is, assign each carrier, independently, time-dependent values of the Bloch state quantum numbers: crystal momentum $\mathbf{p}$ ($= \hbar$ times wavevector), spin component and band. We may explicitly ignore the band and spin variables, however, since the way formulas are generalized to include them (to take into account forces and scattering processes which change them) is obvious. We also ignore carrier-carrier scattering, although in some practical conditions this may not be justifiable and it may have a substantial effect.[3] Then the state of the hot-electron system is determined by the band energy function $E(\mathbf{p})$ and the scattering-probability function $W(\mathbf{p}_1, \mathbf{p}_2)$ for transitions $\mathbf{p}_1 \rightarrow \mathbf{p}_2$. The energy changes, $E(\mathbf{p}_2) - E(\mathbf{p}_1)$, of course have an essential role. They are normally due to absorption and emission of the lattice-vibration phonons. The associated displacement of the phonon system itself from thermal equilibrium is usually assumed negligible, though its neglect may not be justified in some circumstances.[4] The foregoing customary assumptions are not seriously restrictive for present purposes.

The author is located at the IBM Watson Research Laboratory, 612 W. 115 St., New York, N. Y.

Accordingly, the state of the system is given apart from fluctuations by the distribution function $f(\mathbf{p})$, such that the drift velocity is

$$\mathbf{u} = I\, \mathbf{v}(\mathbf{p}) f(\mathbf{p}) \equiv \langle \mathbf{v} \rangle \tag{1.1}$$

and similarly for other macroscopic quantities of interest, and f is governed by the Boltzmann equation[5]

$$\partial f/\partial t = -\mathbf{F}\cdot(\partial/\partial \mathbf{p}) f + \mathfrak{D} f, \tag{1.2}$$

$$\mathfrak{D} f \equiv I(\mathbf{p}')\, (f(\mathbf{p}') W(\mathbf{p}', \mathbf{p}) - f(\mathbf{p}) W(\mathbf{p}, \mathbf{p}')) \tag{1.3}$$

where $\mathbf{F}$ is the force on a carrier, which we may take as the "Lorentz force"

$$\mathbf{F} = (\pm e)(\mathcal{E} + (1/c)\mathbf{v} \times \mathcal{H}). \tag{1.4}$$

In these equations I or $I(\mathbf{p})$ stands for integration $\int d^3\mathbf{p} \cdots$ over the Brillouin zone, the carrier velocity is

$$\mathbf{v}(\mathbf{p}) = \partial E/\partial \mathbf{p} \tag{1.5}$$

and the normalization, in (1.1) and hereafter, is such that

$$I f = 1. \tag{1.6}$$

The second term of (1.3) may be written as $-f/\tau$, where $\tau(\mathbf{p})$ is the scattering time defined by

$$1/\tau \equiv I(\mathbf{p}')\, W(\mathbf{p}, \mathbf{p}'). \tag{1.7}$$

Theoretical analysis has been concerned primarily with solution of the Boltzmann equation (1.2), (1.3) for a steady state, corresponding to a constant electric field with spatial homogeneity, and with the resulting drift velocity (1.1); secondarily with the response of this steady state to small perturbations, such as the static and "a.c." differential mobilities; and to a limited extent with diffusion, due to a

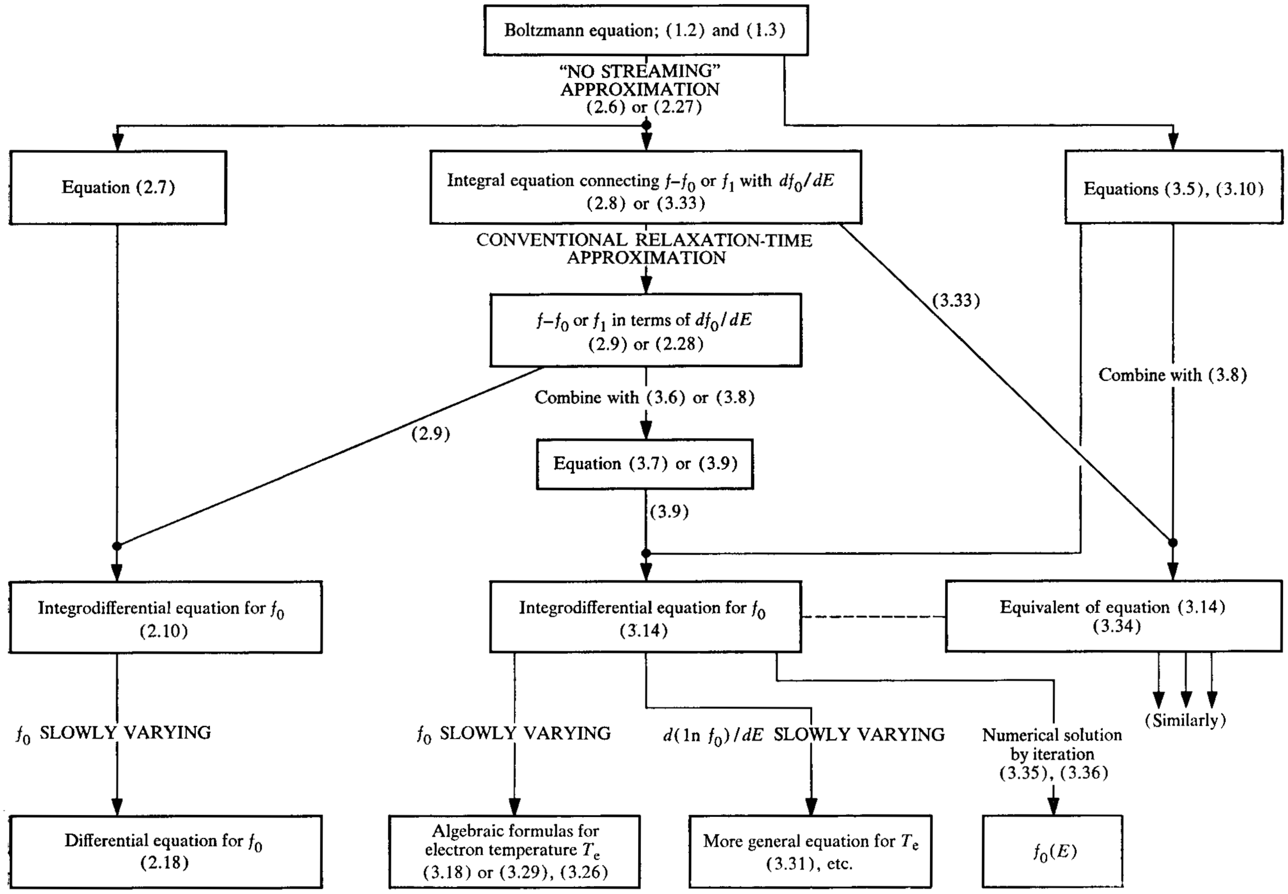


Figure 1 Synopsis of mathematical content of Sections 2 and 3.

gradient of carrier concentration or of field. The present work has the same emphasis. (Phenomena such as the Gunn domains, however, involve *large* time-varying excursions of the field; also diffusion must be assumed to have significant effect.) In general the hot-electron Boltzmann equation cannot be solved except by introducing approximations, such as are discussed in Sections 2 and 3, or by massive numerical processing, as discussed in Section 4.

Methods of solving (1.2), (1.3) by analytical, or limited numerical, means entail reducing it to an equation in a single variable, which is necessarily the energy, E.[6] The unknown of this equation will be the "energy distribution function"

$$f_0(E) \equiv \bar{f} \tag{1.8}$$

where the overhead bar means average over a surface of constant energy:

$$g(E_1)\bar{\psi}(E_1) \equiv I(\mathbf{p})\ \psi(\mathbf{p})\ \delta(E(\mathbf{p}) - E_1) \tag{1.9}$$

for any function $\psi(\mathbf{p})$; the density of states is

$$g(E_1) \equiv I(\mathbf{p})\ \delta(E(\mathbf{p}) - E_1). \tag{1.10}$$

Solution of an equation for $f_0(E)$ is, of course, of central interest in hot-electron theory. Except for the special case of Ref. 9, both derivation and (except as discussed in Section 3) solution of this energy-variable equation require approximations. These are discussed in Sections 2 and 3.

Anisotropy, in the dependence of $E(\mathbf{p})$ on the direction of $\mathbf{p}$ or of $W(\mathbf{p}_1, \mathbf{p}_2)$ on the directions of $\mathbf{p}_1$ and $\mathbf{p}_2$ separately, is a complication which in general requires additional approximations. Even for Ohmic conduction this is an incompletely resolved problem. It may be passed over here, however, without essentially affecting the analysis. Therefore for most of Sections 2 and 3 we assume isotropy in E and W;[7] and similarly the second (magnetic field) term of (1.4) is dropped:

$$\mathbf{F} = (\pm e)\boldsymbol{\mathcal{E}}. \tag{1.11}$$

The theory is developed first, however, as far as possible without assuming isotropy: in Section 2, Eqs. (2.3)—(2.18); in Section 3, Eqs. (3.6), (3.7) and the following equations with $\bar{W}$ in place of W_0.

Figure 1 shows results (in the boxes), and reasoning connecting them, for these two Sections. The left side, corresponding to Section 2, represents the procedure due to Landau and coworkers in the 30's; the remainder, corresponding to Section 3, develops the more general and serviceable formulation introduced by Levinson in 1964. The improvement is from left to right.

The essential condition of validity for Sections 2 and 3 is that f and f_0 differ little. That is, surfaces of constant f should not deviate much from the surfaces of constant E in $\mathbf{p}$ space. Section 3 attempts to develop the best available means for calculating $f_0(E)$ subject to this condition. Methods such as described in Section 4 are evidently required, in general, when one has the substantial displacement of surfaces of constant f (characteristically an elongation in the field direction, which has been termed "streaming") that occurs in some conditions. Section 4 discusses methods which provide essentially exact calculation of the carrier distribution, in principle for all circumstances, but which require computer implementation. They have been used successfully for cases which allow a two-dimensional specification of $f(\mathbf{p})$, as well as a one-dimension case. Extension to three-dimension cases does not appear to involve any essential difficulties, but only the necessary machine capacity. Sections 3 and 4 represent two limits in treatment, and there evidently is no available theory providing for orders of approximation, and calculation procedure, intermediate between these.

2. The conventional approximations

In this Section the traditional procedure for obtaining a solution of the steady-state Boltzmann equation is expounded, with the emphasis on the approximations entailed, and no discussion of any particular application. A more general version allowing anisotropy is given first; then (Eq. (2.19) onward) the conventional, more restricted but more amenable, version is introduced.

Equation (1.2) may be separated into parts which are even and odd in $\mathbf{p}$ (or its equivalent for a many-valley band), in terms of the even and odd parts of $f(\mathbf{p})$. Normally $\mathfrak{D}f$ preserves this parity; and (so long as we have (1.11) rather than (1.4)) $\mathbf{F}\cdot(\partial/\partial\mathbf{p})$ interchanges even and odd parities. Then the two equations are

$$\mathbf{F}\cdot(\partial/\partial\mathbf{p})f_{\text{odd}} = (\mathfrak{D} - \partial/\partial t)f_{\text{even}} \tag{2.1}$$

$$\mathbf{F}\cdot(\partial/\partial\mathbf{p})f_{\text{even}} = (\mathfrak{D} - \partial/\partial t)f_{\text{odd}}. \tag{2.2}$$

Now if we write

$$f \equiv f_0(E(\mathbf{p})) + f'(\mathbf{p}) \tag{2.3}$$

so that

$$f' \equiv f - \bar{f}, \qquad \overline{f'} = 0 \tag{2.4}$$

then f_0 is analogous, here, to the thermal distribution function

$$f_B(E) = a\,e^{-E/kT} \tag{2.5}$$

in the theory of Ohmic conduction. Although $f_0(E)$ is in general far from f_B, it can still be reasonable to assume (with the expectation that $f' \ll f_0$):

Approximation (a)

$$f' \text{ is odd in } \mathbf{p}. \tag{2.6}$$

(Eq. (2.27) is the equivalent Approximation (a), below.) Then (2.1) and (2.2) become (when we also drop the time dependence)

$$\mathfrak{D}f_0 = \mathbf{F}\cdot(\partial/\partial\mathbf{p})f' \tag{2.7}$$

$$\mathfrak{D}f' = \mathbf{F}\cdot(\partial/\partial\mathbf{p})f_0 = \mathbf{F}\cdot\mathbf{v}\,df_0/dE. \tag{2.8}$$

To solve (2.8) for f' we may reasonably use the standard "Ohmic" approximation

Approximation (b)

$$f' = -\tau'(\mathbf{p})\mathbf{v}\cdot\mathbf{F}\,df_0/dE \tag{2.9}$$

where τ' is not necessarily equal to the scattering time τ defined in (1.7). The extent of validity of this customary Ansatz can be taken as similar to what it is for the Ohmic case, and likewise (see Footnote 17) for its equivalent, (2.28). (For any particular model, or simplifying situation, one can obtain a "best" approximate τ' by means of Kohler's minimum principle (see, for example, P. J. Price, *IBM J. Res. Develop.* 1, 239 (1957)); but for the hot-electron situation it should be recognized that, for appreciable inelastic scattering, τ'/τ may be a functional of $f_0(E)$.) For the spherical symmetry case of (2.28), it can in any case be dispensed with by the means described in connection with Eqs. (3.33), (3.34).

By substituting (2.9) in (2.7) we eliminate f' and obtain an equation for f_0. It is expedient to average this equation over a surface of constant energy. Then

$$(\bar{\mathfrak{D}} + \mathbf{FF}\colon(\mathbf{A}\,d/dE + \mathbf{B}\,d^2/dE^2))\,f_0 = 0 \tag{2.10}$$

where the tensors $\mathbf{A}(E)$ and $\mathbf{B}(E)$ are given by

$$\mathbf{A} \equiv \overline{\partial(\tau'\mathbf{v})/\partial\mathbf{p}}, \qquad \mathbf{B} \equiv \overline{\tau'\mathbf{vv}} \tag{2.11}$$

and $\bar{\mathfrak{D}}$ is given by substituting for W in (1.3) the kernel $\overline{W}(E, E')$ obtained from $W(\mathbf{p}, \mathbf{p}')$ by double application of the averaging (1.9), (1.10).

The integrodifferential equation (2.10), which is not very amenable to solution, may be converted into a differential equation by substituting for its first term a truncation of the formal expansion

$$\bar{\mathfrak{D}}f_0 = \sum_{n=0}^{\infty} K_n(E)\,d^nf_0/dE^n \tag{2.12}$$

where

$$K_n(E) \equiv (1/n!) \int dE'\, g(E')(E' - E)^n (\overline{W}(E', E) - \delta_{n,0} \overline{W}(E, E')). \tag{2.13}$$

Replacement of the series (2.12) by only its leading terms amounts to

Approximation (*c*)

$$f_0(E) \text{ is slowly varying.} \tag{2.14}$$

On account of the detailed-balance relation

$$e^{-E(\mathbf{p}_1)/kT} W(\mathbf{p}_1, \mathbf{p}_2) = e^{-E(\mathbf{p}_2)/kT} W(\mathbf{p}_2, \mathbf{p}_1) \tag{2.15}$$

the coefficients (2.13) satisfy

$$\sum_{n=0}^{\infty} (-1/kT)^n K_n = 0, \tag{2.16}$$

in agreement with

$$\mathfrak{D} f_B = 0. \tag{2.17}$$

In the *truncated* series replacing (2.12), however, the coefficients K_n of the retained terms must be adjusted if the sum over these alone on the left of (2.16) is to be zero.[8] In particular, the *second order* differential equation

$$\begin{aligned} Qf_0 \equiv{}& K_0 f_0 + (K_1 + \mathbf{FF} : \mathbf{A})\, df_0/dE \\ &+ (-(kT)^2 K_0 + kTK_1 + \mathbf{FF} : \mathbf{B})\, d^2 f_0/dE^2 \\ ={}& 0 \end{aligned} \tag{2.18}$$

results from the first three terms of (2.12), when K_2 is replaced by $kTK_1 - (kT)^2 K_0$ to satisfy this condition. Alternative approximations will be more conveniently discussed with the formulation in Section 3.

We now restate the foregoing, a little differently, for the case of spherical symmetry:

$$E = E(p), \qquad g = 4\pi p^2/v, \quad \text{where } v = dE/dp \tag{2.19}$$

and

$$W(\mathbf{p}, \mathbf{p}') \text{ depends only on } E(p), E(p') \text{ and the angle } \widehat{\mathbf{p},\mathbf{p}'}. \tag{2.20}$$

For this case one may expand

$$f(\mathbf{p}) = f(E, x) \equiv \sum_{l=0}^{\infty} f_l(E) P_l(x) \tag{2.21}$$

(still assuming (1.11) rather than (1.4)) and

$$\begin{aligned} W(\mathbf{p}, \mathbf{p}') &= W(E, E', y) \\ &\equiv \sum_{l=0}^{\infty} W_l(E, E') P_l(y) \end{aligned} \tag{2.22}$$

where

$$x \equiv \mathbf{p} \cdot \boldsymbol{\mathcal{E}}/p\mathcal{E}, \qquad y \equiv \mathbf{p} \cdot \mathbf{p}'/pp' \tag{2.23}$$

and the P_l are the Legendre polynomials, which will be normalized to $P_l(1) = 1$. (The $\overline{W}$ used in (2.13) corresponds to W_0 here.) The Boltzmann equation may correspondingly be separated into spherical-harmonic components. The l'th equation (coefficient of P_l) is

$$\frac{1}{2l+1} \int dE'\, g(E') f_l(E') W_l(E', E) - \frac{1}{\tau} f_l(E) = Fv\, \phi_l(E) \tag{2.24}$$

where $1/\tau(E)$ is given by (1.7) in terms of W_0 only and the first two of the ϕ_l are

$$\phi_0 = \frac{1}{3}\left(\frac{2}{pv} + \frac{d}{dE}\right) f_1 \tag{2.25}$$

$$\phi_1 = \frac{d}{dE} f_0 + \frac{2}{5}\left(\frac{3}{pv} + \frac{d}{dE}\right) f_2. \tag{2.26}$$

For Approximation (a), the equivalent of Eq. (2.6) is to truncate the sum (2.21) to

$$f = f_0 + x f_1 \tag{2.27}$$

accordingly dropping the f_2 term of (2.26).[9] Then, from (2.24) with $l = 0$ and $l = 1$, we have two simultaneous equations for f_0 and f_1. For entirely isotropic scattering, (2.9) is exact with $\tau' = \tau$; and then (2.6) is a less restrictive (even for spherical symmetry) equivalent of (2.27). With spherical symmetry, if the inelastic (e.g., optical-mode phonon) component of the scattering is isotropic[10] one has from (2.24) for $l = 1$ a virtually exact version of Eq. (2.9):

$$f_1 = -\tau' v F\, df_0/dE. \tag{2.28}$$

In any case, by means of the formulation developed in Section 3, one may dispense with Approximation (b) and still obtain equations of the desired form (see the discussion in connection with Eq. (3.34)).

Substitution of (2.28) in (2.24) for $l = 0$ gives the spherical symmetry version of (2.10). Then the equivalent of (2.18) is obtained similarly. For the case of only acoustic-mode phonon, deformation-potential coupled, scattering, a spherical-symmetry version of Eq. (2.18) was obtained in the mid-30's by Landau and co-workers;[9] and for this case, in normal circumstances,[11] it is virtually exact. Essentially (2.18) has been extensively applied since then to other cases, including valence semiconductors with scattering by optical-mode (as well as acoustic-mode) phonons,[12] and polar semiconductors[13] with their different law of electron-lattice coupling. However, it has come to be recognized[14,15] that Approximation (a)—Eq. (2.6) or (2.27)—then has no general validity. The surfaces of constant f in $\mathbf{p}$-space can be far from coincident with surfaces of constant $E(\mathbf{p})$, and many terms of the sum (2.21) be appreciable. It appears that, except in special circumstances, one then has to resort to the methods discussed in Section 4. When one does have the conditions for a valid "one-dimensional" analysis of the carrier distribution—an equation for $f_0(E)$ alone—there is still a question of the accuracy of Approximation (c), or equivalents, and more generally of the most suitable form of equation and its solution. This is discussed in Section 3.

3. Levinson's formulation

An equation of the type of (2.10) or of (2.18) does not represent the most general, or necessarily the best, treatment of situations in which $f(\mathbf{p})$ may be analyzed in terms of $f_0(E)$ alone. What appears to be an effective overall approach, for such situations, is described in the present Section. It will be shown that when all three basic approximations of Section 2 apply one may obtain equivalent but more tractable equations for f_0; and that Approximation (c), and even Approximation (b), may be given up and one still has useful procedures for obtaining f_0. Approximation (a)—the absence of appreciable "streaming"—remains the basic condition.

For the particular case of Ref. 9, the second order differential operator of (2.18) factorizes into two first order operators:

$$Q = L_1 L_2 \tag{3.1}$$

and so (2.18) is satisfied by the solution of

$$L_2 f_0 = 0, \tag{3.2}$$

which is given by a single integration. It was shown by Levinson[16] that (3.2) applies generally: that is, for conditions essentially those leading to (2.18), one has

$$f_0 + U(E)\, df_0/dE = 0 \tag{3.3}$$

where $U(E)$ is a prescribed functional of $W_0(E, E')$ or of $\bar{W}$. Then (3.3) gives directly

$$f_0(E) = \exp -\int^E dE'/kT_e(E'), \qquad kT_e \equiv U. \tag{3.4}$$

Levinson's basic equation, for the steady state, is

$$0 = J(E) = J_F + J_S \tag{3.5}$$

where $g(E_1)J(E_1)$ is the net rate per unit time at which electrons pass across the surface $E(\mathbf{p}) = E_1$ in the direction of increasing energy, and the two terms on the right of (3.5) are the contributions to J due to the force $\mathbf{F}$ and the scattering respectively. One may arrive at (3.5) by multiplying the right-hand side of (1.2) by a step function of the energy and integrating over p-space. For the first term this gives

$$J_F = \mathbf{F}\cdot\overline{\mathbf{v}f}. \tag{3.6}$$

By (2.3), f may be replaced by f' on the right of (3.6). Equation (2.9) then gives

$$J_F = -\mathbf{FF}{:}\mathbf{B}\, df_0/dE \tag{3.7}$$

where $\mathbf{B}(E)$ is defined in (2.11).

If we assume (as in Ref. 16) the spherical-symmetry conditions (2.19) and (2.20), then by (2.19) and (2.21)

$$J_F = \tfrac{1}{3} F v f_1 \tag{3.8}$$

exactly. Then (2.28) gives

$$J_F = -\tfrac{1}{3} F^2 \tau' v^2\, df_0/dE \tag{3.9}$$

in place of (3.7). (The truncation (2.27) is involved in (3.9), but not in (3.8).) The second term of (3.5) is given by an integral operation on $f_0(E)$:

$$\begin{aligned} g(E)\, J_S(E) &= g\, J_S\{f_0\} \\ &= I(E_1;\, E;\, E_2)\, f_0(E_1)\, W_0(E_1, E_2) \end{aligned} \tag{3.10}$$

where

$$\begin{aligned} &I(E_1;\, E;\, E_2)\, \phi(E_1, E_2) \\ &\quad\equiv \int_{-\infty}^{E} dE_1\, g(E_1) \\ &\qquad \cdot \int_{E}^{+\infty} dE_2\, g(E_2)\, (\phi(E_1, E_2) - \phi(E_2, E_1)) \\ &\quad= 2\int_{-\infty}^{\infty} dE_-\, \mathrm{Sg}(E_-) \\ &\qquad \cdot \int_{E-|E_-|}^{E+|E_-|} dE_+\, g(E_1) g(E_2) \phi(E_1, E_2) \end{aligned} \tag{3.11}$$

with

$$E_\pm \equiv \tfrac{1}{2}(E_2 \pm E_1) \tag{3.12}$$

and

$$\mathrm{Sg}\,(x) \equiv \begin{cases} +1, & x > 0 \\ -1, & x < 0. \end{cases} \tag{3.13}$$

The equation obtained by substituting (3.9) and (3.10) into (3.5)

$$df_0/dE = (3/\tau' v^2 F^2)\, J_S\{f_0\} \tag{3.14}$$

is the equivalent of (the spherical symmetry version of) Eq. (2.10). An equivalent of (2.18) is obtained by using instead of the integral expression for $J_S\{f_0\}$ the first *two* terms of the expansion analogous to (2.12):

$$J_S\{f_0\} = -\sum_{n=0}^{\infty} B_n(E)\, d^n f_0/dE^n. \tag{3.15}$$

We then obtain an equation of the form (3.3). If, as in (2.18), we force consistency with the fundamental relation

$$J_S\{f_B\} = 0 \tag{3.16}$$

by setting $B_1 = kTB_0 \equiv D$, then

$$J_S = -D(E)\left(\frac{f_0}{kT} + \frac{df_0}{dE}\right) \tag{3.17}$$

and hence for the "electron temperature" T_e of (3.4)

$$T_e/T = 1 + (\tau' v^2/3D)F^2. \tag{3.18}$$

Levinson[16] gives for the coefficient in (3.17)

$$D(E) = \tfrac{1}{2} \int dE' \, g(E')(E' - E)^2 W_0(E, E'). \tag{3.19}$$

The exact formulas for the first two coefficients in (3.15) are

$$-gB_0 = I(E_1; E; E_2) \; W_-(E_1, E_2) \tag{3.20}$$

$$-gB_1 = I(E_1; E; E_2) \; ((E_+ - E)W_-(E_1, E_2) - E_- W_+(E_1, E_2)) \tag{3.21}$$

in the notation (3.11), where

$$W_\pm(E_1, E_2) \equiv \tfrac{1}{2}((W_0(E_1, E_2) \pm W_0(E_2, E_1)). \tag{3.22}$$

Now, so long as the maximum energy change in scattering $(E_2 - E_1)$ is small compared to E, $W_0(E_1, E_2)$ can be expected to vary little, with $E_2 - E_1$ held fixed, over the range of the second integral on the right of (3.11); and $g(E_1)g(E_2)$ should also vary little. Accordingly we may drop the W_- term of (3.21). For the other term of (3.21), similarly the factor $g(E_1)g(E_2)W_+(E_1, E_2)$ can be taken outside the second integral of (3.11), with E_+ replaced by E in W_+ and $g(E_1)g(E_2)$ set equal to $g(E)^2$. Then we obtain (with $\epsilon = 2E_-$)

$$B_1 = \tfrac{1}{2} g \int_{-\infty}^{\infty} W_+(E - \tfrac{1}{2}\epsilon, \, E + \tfrac{1}{2}\epsilon)\epsilon^2 \, d\epsilon. \tag{3.23}$$

The equivalent result for B_0 from (3.20) is

$$B_0 = g \int_{-\infty}^{\infty} W_+(E - \tfrac{1}{2}\epsilon, \, E + \tfrac{1}{2}\epsilon)\epsilon \tanh(\epsilon/2kT) \, d\epsilon \tag{3.24}$$

since on account of the "detailed balance" relation, (2.15),

$$W_+ \sinh(E_-/kT) + W_- \cosh(E_-/kT) = 0. \tag{3.25}$$

The actual result of substituting the first two terms of (3.15) in (3.14)

$$T_e/T = (B_1/B_0kT) + (\tau' v^2/3B_0kT)F^2 \tag{3.26}$$

in general differs from (3.18) in that $B_1 \neq B_0kT$, except in the limit $kT \gg$ (the predominant energy change in scattering). For simple dispersionless optical-mode-phonon plus elastic scattering, in particular,

$$B_1/B_0kT = (T_o/2T) \coth(T_o/2T) \tag{3.27}$$

where $kT_o \equiv \hbar\omega_o$ is the optical-mode quantum. In fact the truncation of (3.15) is admissible for the high fields, and appropriate energies E, where $kT_e(E)$ is large compared to the energy changes in scattering. For weak fields such that f_0 is close to f_B, however, one may better approximate J_S by truncating the series

$$J_S = -f_B \sum_{n=1}^{\infty} D_n(E) \, d^n(f_0/f_B)/dE^n. \tag{3.28}$$

Retaining the first term alone of the sum in (3.28) gives, in place of (3.26),

$$T_e/T = 1 + (\tau' v^2/3D_1)F^2. \tag{3.29}$$

Assumptions and reasoning like those leading to (3.23) and (3.24) give, to an equivalent approximation,

$$D_1 = kTB_0. \tag{3.30}$$

Thus D_1 differs numerically from the D given by (3.19), again unless kT is large compared to energy changes in scattering.

We cannot have a single equation like (3.26) or (3.29) applying over the whole range of field strengths. But by assuming that $T_e(E)$ varies inappreciably over a range equal to the predominant energy change in scattering, so that $f_0(E_1)$ may be replaced by $f_0(E) \exp(E - E_1)/kT_e(E)$ in (3.10), one can obtain from (3.14) an algebraic equation for T_e. With the same assumption as leads to (3.23), (3.24) and (3.30) (but not that required to justify truncating (3.15), or (3.28)) this becomes

$$\tfrac{1}{3}\tau' v^2 F^2 = 2(kT_e)^2 g \cdot \int_{-\infty}^{\infty} W_+(E - \tfrac{1}{2}\epsilon, \, E + \tfrac{1}{2}\epsilon) G(\epsilon, T, T_e) \, d\epsilon \tag{3.31}$$

$$G(\epsilon, T, T_e) \equiv \sinh\left(\frac{\epsilon}{2kT} - \frac{\epsilon}{2kT_e}\right) \cdot \sinh\left(\frac{\epsilon}{2kT_e}\right) \Big/ \cosh\left(\frac{\epsilon}{2kT}\right).$$

For simple dispersionless optical-mode plus elastic scattering, (3.31) gives

$$\frac{\tau' v^2}{3D_1} F^2 = \left(\frac{2T_e^2}{TT_o}\right) \sinh\left(\frac{T_o}{2T} - \frac{T_o}{2T_e}\right) \cdot \sinh\left(\frac{T_o}{2T_e}\right) \Big/ \sinh\left(\frac{T_o}{2T}\right) \tag{3.32}$$

with D_1 given by (3.30) and (3.24) for this case. Eq. (3.32) reduces to (3.26), with (3.27) and (3.30), for $2T_e \gg T_o$; and to (3.29) for $T_e \simeq T$. It should be recalled, however, that where optical-mode scattering is important we have still in effect assumed (in (3.23) etc.) that $E \gg \hbar\omega_o$; and results for the limit $T_e \simeq T$ (and more generally, for small T/T_o) may accordingly have a limited applicability. Figure 2 displays the relation (3.32), as T_e/T versus its left-hand side for some values of T/T_o.

So far we have been concerned with the integrodifferential equation (3.14), which was obtained subject to any approximations entailed in the use of (2.28). After dropping the f_2 term of (2.26), the resulting integral equation (2.24) for f_1 has a solution of form (2.28) only for some cases.[17]

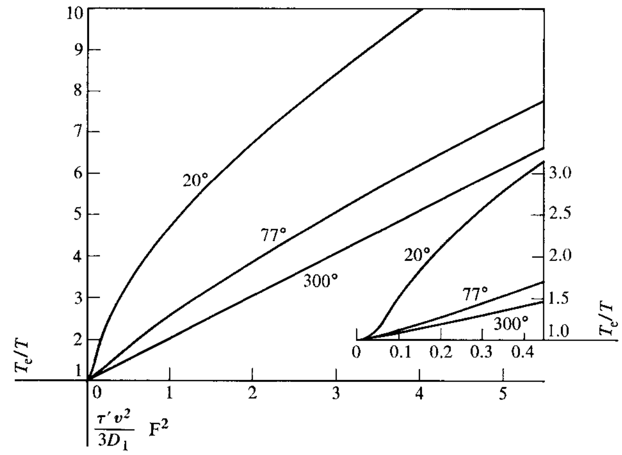


Figure 2 T_e/T (vertical scales) versus the left-hand side of Eq. (3.32), for the values of T indicated on the curves with $T_o = 440°$.

Given the assumed spherical symmetry (in addition to the necessary dropping of the f_2 term), however, there is in this respect no essential difficulty with the formulation of the present Section. The equation for f_1 is of form

$$-Fv\tau\, df_0/dE = M\{f_1\}, \tag{3.33}$$

where M is the integral operator (on f_1) appearing on the left of (2.24) for $l = 1$, after multiplying by $-\tau$. But (3.5) and (3.8) give an exact equation $f_1 = -(3/Fv)J_S\{f_0\}$ which can be substituted in (3.33) to give the integro-differential equation

$$\tfrac{1}{3}\tau v F^2 df_0/dE = M\{(1/v)J_S\{f_0\}\} \tag{3.34}$$

with kernel obtained from a convolution of W_0 and W_1. The procedures for estimating T_e, and that described below in terms of Eq. (3.35), should apply to (3.34) as well as to (3.14). The corresponding generalizations of (3.27), (3.32), etc. will involve additional terms in T_o like those in $T_o/2$.

Although it has been useful to develop in some detail the approximate analytical solutions of (3.14), the entailed restriction to E fairly large compared to energy changes in scattering implies a limit on their applicability. However, Eq. (3.14)—and, similarly, Eq. (3.34)—is suitable for numerical solution, without the approximations made in these analytical solutions, by iteration. An appropriate iteration sequence is

$$f_0^{(n+1)}(E) = \exp \int^E \frac{1}{F^2}\left(\frac{3}{\tau' v^2}\frac{1}{f_0^{(n)}} J_S\{f_0^{(n)}\}\right)_{E'} dE'$$
$$\equiv \Lambda\{f_0^{(n)}\}. \tag{3.35}$$

The integrand of the integral on the right of (3.35), at convergence of the sequence, is $-1/kT_e(E) \equiv d(\ln f_0)/dE$. Not only does this iteration scheme not have the inherent stability of that described in Section 4; it has an inherent instability. By (2.15), for T_e constant over the energy range of integration in J_S the integrand of (3.35) has the sign opposite to that of $(1/kT) + d(\ln f_0^{(n)})/dE'$, and its dependence on the latter is monotonic. Therefore, when $f_0^{(n)}$ is close to the solution of (3.14), if it has $T_e(E)$ values which are too large then (3.35) should give an $f_0^{(n+1)}$ with T_e values which are too small, and conversely. For low fields especially, a resulting alternation of errors can be unstable. This difficulty was met, in the calculation described below, by using instead of (3.35) the sequence

$$f_0^{(n+1)} = (1 - a)f_0^{(n)} + a\Lambda\{f_0^{(n)}\} \tag{3.36}$$

and controlling the parameter a so as to approach the fastest convergence compatible with stability. (Still, it is evident that this iteration scheme is a "high field" procedure which cannot be expected to extend down to $\mathcal{E} = 0$; the same reservation applies to the method expounded in Section 4.)

Since the iteration method described above is new, and may be considered to give the best possible "one-dimensional" calculation of the distribution function on the basis of the Landau truncation of (2.21) to its first two terms (2.27), a trial application was made. The case was n-germanium, with lattice scattering only, and intervalley scattering taken to be so weak that it may be neglected in calculating the distribution in a single valley. The $E(\mathbf{p})$ function of a valley was taken to be the "tensor-mass parabolic" $\mathbf{p}\cdot(1/2\mathbf{m})\cdot\mathbf{p}$ with constant $1/\mathbf{m}$ tensor, and the in-valley scattering to be isotropic, elastic acoustic-mode-phonon plus single-energy ($\hbar\omega_o \equiv kT_o$) optical-mode. Then $\tau' = \tau$, and the scattering rate may be written[18]

$$\frac{1}{\tau} = R_1\,(T/T_o)\sqrt{X}$$
$$+ R_2\,(N_o\sqrt{(X+1)} + (N_o+1)\sqrt{(X-1)}) \tag{3.37}$$

where

$$X \equiv E/kT_o, \qquad N_o \equiv 1/(e^{T_o/T} - 1) \tag{3.38}$$

and $\sqrt{(\ )}$ is to be taken as zero for negative argument. The parameters R_1, R_2 are acoustic-mode and optical-mode coupling constants. This system with axial symmetry is reduced to spherical symmetry by a linear transformation of $\mathbf{p}$ space, equivalent to having a "crystal momentum" equal to $p^*(E) \equiv \sqrt{(2Em^*)}$, where[19]

$$\frac{1}{m^*} = \frac{1}{m_\perp}\left(\frac{\mathcal{E}_\perp}{\mathcal{E}}\right)^2 + \frac{1}{m_\parallel}\left(\frac{\mathcal{E}_\parallel}{\mathcal{E}}\right)^2 \tag{3.39}$$

With the actual scattering function (represented by (3.37), (3.38)) and actual field strength $\mathcal{E}$, this gives the correct energy distribution $f_0(E)$; and the correct drift velocity is the scalar product of $1/\mathbf{m}$ and the expectation of $\mathbf{p}^*$ (which will be parallel to $\boldsymbol{\mathcal{E}}$) in this transformed system.

It will be of interest to compare the computed results with those given by the algebraic formulas. In the limit represented by (3.26), the density-of-states factors $\sqrt{(X \pm 1)}$ occurring in the evaluation of B_0 or B_1 may be replaced by $\sqrt{X}$. Then (3.26) would give

$$\frac{T_e}{T} = \frac{T_o}{2T} \coth\left(\frac{T_o}{2T}\right) + \frac{2}{3m^*kTR_2R} F^2 \qquad (3.40)$$

where

$$R \equiv R_1(T/T_o) + R_2(2N_o + 1). \qquad (3.41)$$

As expected, T_e is in this approximation independent of E. To the same approximation the drift velocity would be $\mu_0 \mathcal{E}(T/T_e)^{\frac{1}{2}}$, where

$$\mu_0 = (4/3\sqrt{\pi})(e/m^*R)(T_o/T)^{\frac{1}{2}} \qquad (3.42)$$

is the value the Ohmic mobility, in the $\mathcal{E}$ direction, would have if τ were equal to $1/(R\sqrt{X})$.

In these calculations the field $\mathcal{E}$ was taken to be in the (1, 0, 0) direction; so that m^* was the free electron mass divided by 8.39.[19] The value 440° was used for T_o. The results reported below are for $T = 300°$, with $R_1 = 0.37 \times 10^{13}$ sec^{-1} and $R_2 = 0.022 \times 10^{13}$ sec^{-1}.[20] The iteration scheme (3.35), (3.36) was implemented on the APL terminal-based interactive computation facility,[21] which allowed its operation to be monitored and freely interrupted and controlled, and the numbers generated to be freely and selectively accessed. A "grid" of equally spaced X values (varying in number, up to 271 for $\mathcal{E} = 5000$ V/cm, and in spacing) was used to represent f_0.

A feature of interest in the results is the energy dependence of T_e and f_1/f_0, which are constant in the approximation represented by (3.40). Figure 3 shows T_o/T_e and f_1/f_0 versus $X \equiv E/kT_o$, for $\mathcal{E} = 1000$ V/cm. ($T_o/T_e(X)$ is more informative, displayed, than $f_0 = \exp - \int^X (T_o/T_e)\, dX'$.) At higher fields the sharp shoulders of these curves soften to smooth bends connecting the rising and flat parts, and the location of these bends shifts slightly to the right; but the curves remain otherwise similar.[22]

Figure 4 compares the limiting value of f_1/f_0 for large X (indicated by the points) with the constant value given by (3.40) (lower curve) and with that given by (3.40) with the first term on the right replaced by 1, i.e., by (3.29) instead of (3.26) (upper curve). The points cross over from the "low field" to the "high field" curve with increasing $\mathcal{E}$, as they should. The analogous behavior was found for the limiting value of T_o/T_e at large X, and for $\langle X \rangle$, but the displacements entailed in the crossovers were less. For each of these two quantities, a single curve may be constructed from the two limiting curves and the computed points; these are shown in Fig. 5. Figure 6 shows drift velocity versus field.[23]

It would be of interest to extend these computations to systems of many valleys with scattering between them,

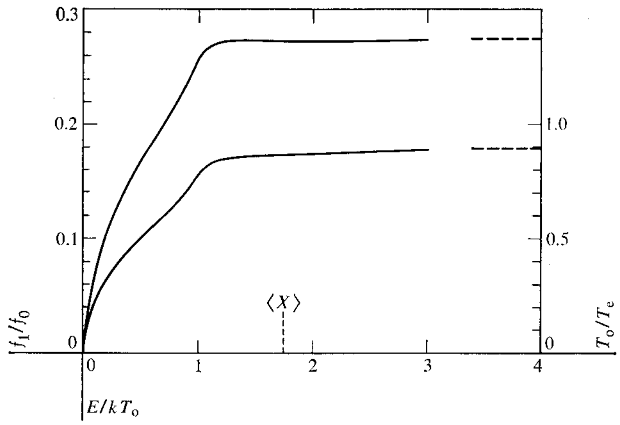


Figure 3 T_o/T_e (lower curve, right scale) and f_1/f_0 (upper curve, left scale) versus E/kT_o for a field of 1000 V/cm.

Figure 4 Limit of f_1/f_0, for large E, versus field. Curves and points obtained as described in the text.

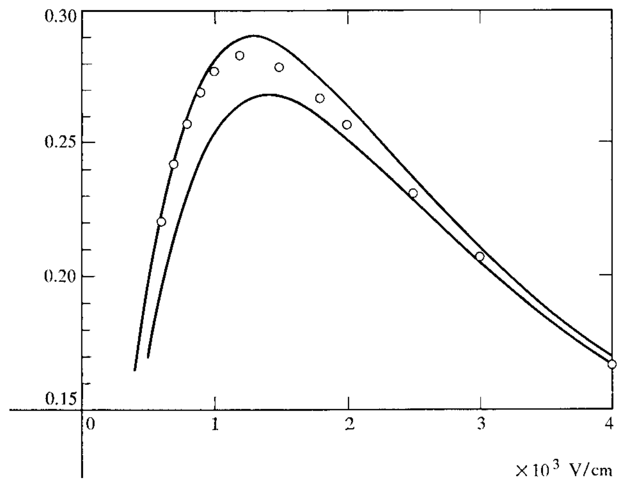
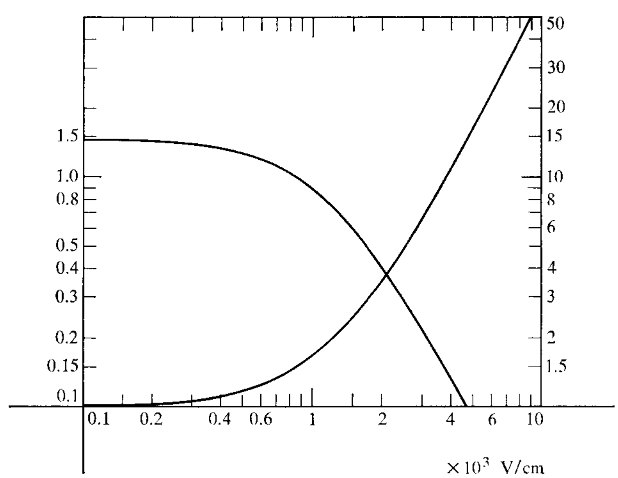


Figure 5 Limit of T_o/T_e, for large E (descending curve, left scale) and expectation of E/kT_o (ascending curve, right scale) versus field.



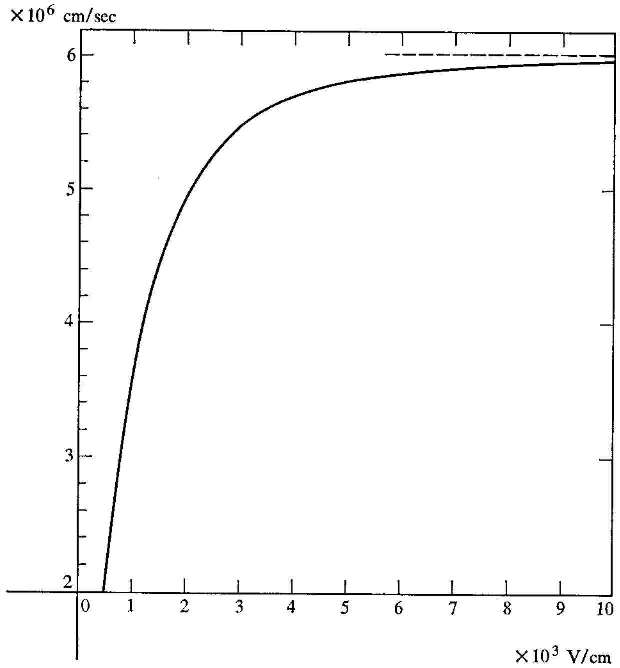


Figure 6 Drift velocity versus field.

especially since (see the discussion at the end of Section 4) the fundamental condition for validity of Eq. (3.14) evidently is satisfied in the present case. A change in the form of the equations is entailed, however. The left-hand side of Eq. (3.5), for each valley, is no longer zero; it is equal to a net "upward" flow of carriers, given by an integral over valley f_0 functions times intervalley W functions. Since this integral will extend over at least the energy range to the band edge, it could not be subsumed in J_S for approximations such as lead to (3.31).

The main deficiencies of the formulation developed in this Section, and hence of the iterative solution of Eq. (3.14), appear to be (a) that appreciable (band-energy and scattering) anisotropy can be adequately taken account of only in special cases; (b) that this formulation is not the lowest order in a systematically generated sequence of approximations, which could deal with the extent of deviation of $f(\mathbf{p})$ from $f_0(E)$.[24] This deviation does not affect the validity of the methods discussed in the following Section, which are quite general, but also their implementation is more demanding. Obviously it would still be desirable to have available the fullest possible extent of applicable "one dimensional" formulations.

4. Precise calculations

In general the complete solution of the Boltzmann equation or its equivalent, for the hot-electron distribution, is not available in closed form. But there are methods for systematic numerical evaluation. The cases which these have dealt with, so far, have been reducible to two dimensions in $\mathbf{p}$ space (i.e., systems having spherical symmetry, for which $f(\mathbf{p})$ has axial symmetry and depends on p and the angle $\widehat{\mathbf{p}\mathcal{E}}$; and the n-Ge case of the preceding Section, which can be transformed to spherical symmetry as indicated there), in addition to a one-dimension case. The number of arithmetical operations required (for a given precision) is proportional to, and many times, the number of points in $\mathbf{p}$ space used in representing the distribution function; and this figure determines the needed computer capacity.

One of these procedures is the Monte Carlo method,[25] which has been applied to the present problem[26] by Kurosawa[27] and by Boardman, Fawcett and Rees.[28] The other, to be discussed here, could be considered as a systemized, and arguably preferable, version of Monte Carlo[29] or [30] as based on a description of the electron system in terms of ensembles differing from the conventional ensembles of statistical mechanics. The distribution is obtained as the result of repeated application (iteration) of a linear integral operator (in the form of two such operators applied successively). There should normally be stable convergence (but see Footnote 39) to a result which can be expected to be independent of the starting function. The method, in various versions, has been applied to several hot-electron systems.[31–33] In addition to these calculations of the hot-electron steady state, the linear response to small perturbations[34,35] may be obtained in terms of sums over the results of successive iterations,[33] as well as by a perturbation of the iterative computation itself.[36,37]

The traditional "μ-space" ensemble for a steady state may be considered as referring to a set of single-particle states, i.e., here a set of $\mathbf{p}$ values $\{\mathbf{p}\} = \mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \cdots$, obtained by sampling $\mathbf{p}(t)$ for an individual particle (undergoing the accelerations and scatterings represented by $\mathbf{F}$ and W in (1.2) and (1.3)) at *arbitrary* times t_1, t_2, $\cdots$. Then $f(\mathbf{p})$ describes $\{\mathbf{p}\}$. By sampling instead at times correlated with the scattering events we may obtain equally representative, but different, ensembles. In particular, the t_n may be *the successive scattering times themselves.* If the nth scattering is the transition $\mathbf{p}_b^{(n)} \rightarrow \mathbf{p}_a^{(n)}$ then we have "before" and "after" ensembles $\{\mathbf{p}_b\} \equiv \mathbf{p}_b^{(1)}, \mathbf{p}_b^{(2)}, \cdots \mathbf{p}_b^{(n)}, \cdots$ and $\{\mathbf{p}_a\} \equiv \mathbf{p}_a^{(1)}, \mathbf{p}_a^{(2)}, \cdots$, with distribution functions $f_b(\mathbf{p})$ and $f_a(\mathbf{p})$, respectively. Since $\{\mathbf{p}_b\}$ is the set of initial states of all the scattering events, its distribution function must be f times the scattering rate:

$$f_b(\mathbf{p}) = (1/\tau(\mathbf{p}))\, f(\mathbf{p})/I\,(1/\tau)f \tag{4.1}$$

where the denominator on the right conserves normalization ($I\,f_b = I\,f$).

The two distributions are connected by the relations

$$f_b = A f_a \tag{4.2a}$$

$$f_a = C f_b \tag{4.2b}$$

where A and C are linear integral operators, given by (4.10) and by (4.6). Then

$$f_b = ACf_b. \tag{4.3}$$

The iteration scheme referred to above is expressed by

$$f_b^{(n)} = Af_a^{(n-1)} \tag{4.4}$$

$$f_a^{(n)} = Cf_b^{(n)} \tag{4.5}$$

starting from $f_a^{(0)}$, say, and continuing until the sequence of resulting functions is stationary. One may consider these successive functions as successive distribution functions of a set of particles, with each particle "observed" at its own scattering times t_n. Thus $f_a^{(0)}$ gives the distribution of the initial $\mathbf{p}$ values of the set of particles; then $f_b^{(1)}$ gives the distribution of the $\mathbf{p}_b^{(1)}$, the $\mathbf{p}$ values at the instants preceding the first scattering of each of the particles; and so on. Although these scatterings are not simultaneous (the set of t_n for a given n are not all equal), nevertheless the sequence[38] $\cdots f_a^{(n-1)}, f_b^{(n)}, f_a^{(n)}, f_b^{(n+1)}, \cdots$ from this point of view represents the thermalization of the system, from an initial distribution $f_a^{(0)}$ and presumably to an ultimate steady state.[39] That is, one may expect the iteration scheme (4.4), (4.5) to converge for those hot-electron systems which do thermalize to a steady state.

Because of the relation between the "a" and "b" states, it is evident that

$$C\psi = I(\mathbf{p}')\, \psi(\mathbf{p}')\tau(\mathbf{p}')W(\mathbf{p}', \mathbf{p}) \tag{4.6}$$

for any function $\psi(\mathbf{p})$. To obtain A it is convenient to introduce the operator Z such that, if ψ is a function of the state of a particle, $Z\psi$ is the expectation of ψ at the instant preceding the first subsequent scattering of a particle after it is in a given initial state (with $Z\psi$ considered as a function of the latter). Since $Z\psi$ is the expectation of ψ at the next "b" state following a given state, $I\, f_a(Z\psi) = I\, f_b\psi$. Therefore $I\, f_b\psi = I\, (Z^*f_a)\psi$, where Z^* means the adjoint of Z.[40] Hence $f_b = Z^*f_a$, or

$$A = Z^*. \tag{4.7}$$

Now, the probability that the elapsed time to the next collision, after a given initial state, will exceed t is

$$P(\mathbf{p}, t) = \exp -\int_0^t dt'/\tau(\mathbf{p} \mid t') \tag{4.8}$$

where the argument $(\mathbf{p} \mid t)$ means the state ($\mathbf{p}$ value) reached by an electron after a time t, starting from $\mathbf{p}$, along the trajectory $d\mathbf{p}/dt = \mathbf{F}$.[41] Then

$$Z\psi = \int_0^\infty \psi(\mathbf{p} \mid t)(-dP(\mathbf{p}, t)/dt)\, dt$$

$$= \int_0^\infty \psi(\mathbf{p} \mid t)P(\mathbf{p}, t)\, dt/\tau(\mathbf{p} \mid t). \tag{4.9}$$

By (4.7), (4.8) and (4.9)

$$A\psi = (1/\tau(\mathbf{p})) \int_{-\infty}^0 dt\, \psi(\mathbf{p} \mid t) \exp \int_0^t dt'/\tau(\mathbf{p} \mid t') \tag{4.10}$$

with the left-hand side considered as a function of $\mathbf{p}$.

We may define a similar pair of operators

$$\tilde{C} \equiv C\,(1/\tau), \qquad \tilde{A} \equiv \tau A \tag{4.11}$$

by removing the factors $\tau(\mathbf{p}')$ in (4.6) and $1/\tau(\mathbf{p})$ in (4.10). Then by (4.1)[42]

$$f = \tilde{A}\tilde{C}f. \tag{4.12}$$

This is an exact form of the Shockley-Chambers path integral formula. It follows directly from the steady-state Boltzmann equation

$$[1/\tau + \mathbf{F}\cdot(\partial/\partial\mathbf{p})]\, f = \tilde{C}f \tag{4.13}$$

and the fact that the operator on the left is the inverse of $\tilde{A}$:[43]

$$[1/\tau + \mathbf{F}\cdot(\partial/\partial\mathbf{p})]^{-1} = \tilde{A}. \tag{4.14}$$

Equation (4.1) then follows, by (4.11), from comparison of (4.12) with (4.3).

Perturbations of the steady state due to small changes in $\mathbf{F}$, and hence the differential mobility and Hall effect, may be calculated by similar means.[44] By (4.13), the linear response δf to a time-independent change $\delta\mathbf{F}$ is given by

$$[1/\tau + \mathbf{F}\cdot(\partial/\partial\mathbf{p}) - \tilde{C}]\delta f = -(\delta\mathbf{F})\cdot(\partial/\partial\mathbf{p})f \equiv h(\mathbf{p}) \tag{4.15}$$

and hence, by (4.14),

$$\delta f = \tilde{A}h + \tilde{A}\tilde{C}\tilde{A}h + \tilde{A}\tilde{C}\tilde{A}\tilde{C}\tilde{A}h + \cdots$$

$$= \tau(Ah + ACAh + \cdots). \tag{4.16}$$

The coefficient of τ on the right of (4.16) is just equal to $\Sigma_{n=1}^\infty f_b^{(n)}$ for the sequence (4.4), (4.5) when the initial function $f_a^{(0)}$ is h. Since A and C conserve normalization, the convergence of the series depends on the fact that $I\, h = 0$. (The latter is similarly a condition for (4.15) to have a solution.) One might promote numerical convergence by introducing a projection operator Ω, defined by

$$\Omega\psi = \psi - f\, I\psi \tag{4.17}$$

where f is the normalized solution of Eq. (4.13), for example by the substitution $\tilde{A} \to \Omega\tilde{A}$ in (4.16); or by a rearrangement of terms like that in (4.21). (Because of the eigenvalue equation (4.12), there is no formal resolvent for the operator $1 - \tilde{A}\tilde{C}$, and similarly with (4.3); but $1 - \Omega\tilde{A}\tilde{C}$ is not subject to this limitation.)

Since the drift velocity is $\mathbf{u} = I\, f\mathbf{v}$ we have $\delta\mathbf{u} = I\, \mathbf{1}'h = I\, f(\delta\mathbf{F})\cdot(\partial/\partial\mathbf{p})\mathbf{1}'$, and hence the differential mobility is given by

$$\delta\mathbf{u} = (\pm e\delta\mathcal{E})\cdot\langle(\partial/\partial\mathbf{p})\mathbf{1}'\rangle \tag{4.18}$$

where $\mathbf{l}'$, the differential mean free path,[34,35] is given by

$$\mathbf{l}' = Z\tau(\mathbf{v} - \mathbf{u}) + ZSZ\tau(\mathbf{v} - \mathbf{u}) + \cdots \tag{4.19}$$

with Z the adjoint of A as before, and[45]

$$S = C^*. \tag{4.20}$$

The terms explicitly proportional to $\mathbf{u}$, in (4.19), are necessary for convergence of the sum.[46] It may alternatively be expressed as

$$\mathbf{l}' = (Z\tau\mathbf{v} - ZSZ\tau\mathbf{v}) + 2\,(ZSZ\tau\mathbf{v} - ZSZSZ\tau\mathbf{v}) + 3\,(ZSZSZ\tau\mathbf{v} - ZSZSZSZ\tau\mathbf{v}) + \cdots. \tag{4.21}$$

The calculations for time-dependent perturbations are similar. If $\delta\mathbf{F}$ varies as $\exp(i\omega t)$, there is an additional term $+i\omega$ in [ ] on the left of (4.15). This may be taken into account by the substitution $1/\tau \rightarrow i\omega + 1/\tau$ in the definition of $\tilde{A}$. That is, $\tilde{A} \rightarrow \tilde{A}_\omega$ where

$$\tilde{A}_\omega \psi = \int_{-\infty}^{0} dt\ \psi(\mathbf{p} \mid t)\ \exp \int_0^t dt'(i\omega + 1/\tau(\mathbf{p} \mid t')) \tag{4.22}$$

(but still, as in (4.11), $\tilde{A}_\omega = \tau\tilde{A}_\omega$). The *alternative* substitutions (instead of replacing $\tilde{A}$ by $\tilde{A}_\omega$ and accordingly for A and Z)

$$\tilde{C} \rightarrow \tilde{C} - i\omega, \quad C \rightarrow C - i\omega\tau, \quad S \rightarrow S - i\omega\tau \tag{4.23}$$

would generate power series in ω represented by the formal substitution

$$K \rightarrow K + K(-i\omega\tau)K + K(-i\omega\tau)K(-i\omega\tau)K + \cdots, \tag{4.24}$$

where K is $A + ACA + \cdots$ or $Z + ZSZ + \cdots$. The term proportional to $i\omega$ in the expansion for the differential mobility contributes to the zero-frequency dielectric constant.[37]

For $\omega\tau \ll 1$ and fields such that $\tau \simeq A\tau$, we may take

$$A_\omega \approx e^{-i\omega\tau}A \tag{4.25}$$

and similarly for $\tilde{A}$ and Z. The "self scattering" transformation of Rees[28,32,37] and Skullerud[47] makes the substitution

$$W(\mathbf{p}, \mathbf{p}') \rightarrow W(\mathbf{p}, \mathbf{p}') + \left(\frac{1}{\tau_0} - \frac{1}{\tau(\mathbf{p})}\right) \delta^3(\mathbf{p} - \mathbf{p}') \tag{4.26}$$

in the preceding equations, where $\delta^3(\ )$ is the three-dimensional Dirac function and τ_0 is a constant ($1/\Gamma$, in Rees' notation) not greater than the minimum of $\tau(\mathbf{p})$; then all the calculations described here are changed accordingly,[48] but the resulting distribution functions, etc., should be unchanged. If τ_0 is made sufficiently small, for given $\mathbf{F}$ and ω, that (4.25) (with τ_0 in place of τ in the exponent) is admissible, the latter will introduce a constant factor $\exp(-in\omega\tau_0)$ in the nth term of (4.16) and of (4.19). Then the same operators and numerical processes (defined for zero frequency, but containing the Rees substitution (4.26)) will be applicable to all frequencies $\omega \ll 1/\tau_0$, with $n\tau_0$ having the role of the time t in the evolution of f. Extensive calculations of differential mobility versus frequency have been made by Rees[36,37] on this basis. In these calculations the number of iterations, and hence the amount of numerical processing, is increased by a substantial factor of order τ/τ_0. Kwok, Lebwohl, Marcus and Schultz[49] have suggested that one would do better to just integrate the Boltzmann equation (1.2), over the t variable, to obtain $f(\mathbf{p}, t)$ or δf as a function of t. The A operator is then not required in the computations; one has instead the differentiation for the term $\mathbf{F}\cdot(\partial/\partial\mathbf{p})f$. They further propose to calculate $\Pi(-t)f$ instead of f, where the operator $\Pi(t)$ is defined by $\Pi(t)\psi(\mathbf{p}) \equiv \psi(\mathbf{p} \mid t)$. This satisfies the equation

$$(\partial/\partial t)[\Pi(-t)f] = [\Pi(-t)\mathfrak{D}\Pi(t)][\Pi(-t)f] \tag{4.27}$$

so that even the $\mathbf{F}\cdot(\partial/\partial\mathbf{p})$ operator is eliminated and only the modified scattering operator $\Pi(-t)\mathfrak{D}\Pi(t)$ remains in the "equation of motion."

For a system with spherical symmetry, C may be expanded in harmonics corresponding to (2.22), $\Sigma_l C_l P_l$; and the analogous double expansion for A, $\Sigma_m\Sigma_n A_{mn}P_l P_m$, will complete the relations (4.2), with f_a and f_b expanded as in (2.21). (C is "diagonal in this representation," but A is not.) A suitable approximation might be to truncate the series to l, m, $n \leq N$ for a suitable N, leaving an N-component scheme. For isotropic scattering, however, C reduces to C_0; then, in (4.4), A may be replaced by A_{00}. One thereby has an *exact* one-dimension scheme of equations (i.e., with only the energy as independent variable), which was used by Budd.[31] The kernel of A_{00} is unfortunately a complicated function. It has a singularity $(1/\pi\tau vF)\ln(p/2|p - p'|)$ at $E' = E$; but[31] this can be allowed for in the algorithm. For moderate fields one might attempt replacing A_{00} by a simple form based on the first few moments of its kernel.

The iteration scheme (4.4), (4.5) has been implemented and tested for the same n-Ge case,[50] with the same parameter values (T_o, m^*, T, R_1, R_2) as in the computation described at the end of Section 3; and the resulting values of drift velocity and average energy, for fields from 1000 to 5000 V/cm, were in satisfactory agreement. This is consistent with the situation that the Landau truncation (2.27) is a good approximation for this case and parameter values, according to the estimation

$$f_2/f_0 \simeq \tfrac{1}{3}(f_1/f_0)^2 \tag{4.28}$$

(for small f_1/f_0) which with the f_1/f_0 values of Figs. 3 and 4 gives f_2/f_0 less than 3%. Thus, by a choice of parameter values for which good numerical agreement is to be expected, it has been possible to check against each other the procedures described in this Section, for the steady state, and in Section 3. A similar numerical comparison for conditions where f_2/f_0 is not small accordingly would also be of interest, since it should give a measure of the accuracy of the method of Section 3 when streaming is appreciable.

Acknowledgments

I am indebted to the many IBM colleagues who have discussed hot electrons with me; to Mrs. Y. C. Chow for most of the programming work for the computations reported at the end of Section 4; and to N. Friedman for Fig. 2, and for other assistance in the preparation of this paper.

References and footnotes

1. A detailed and extensive account of hot-electron physics is: *High Field Transport in Semiconductors* by E. M. Conwell, Academic Press, New York, 1967.
2. The applied fields and their rates of change, and the carrier densities, are small on the appropriate scale; and energy band widths and scattering lifetimes are large enough for the mobility to be not unduly low.
3. For example, A. Hasegawa and J. Yamashita, *J. Phys. Soc. Japan* **17,** 1751 (1962). Avalanche processes may be included to some extent in the analysis developed here, as formally equivalent to scattering; but if the Auger processes are not negligible they introduce nonlinearity in f.
4. See Ref. 1, Chapter III, Section 5a. The detailed balance relation (2.15), especially, depends on this assumption.
5. The diffusion term $-\mathbf{v}\cdot\text{grad}\, f$ is omitted from (1.2) (but see Footnote 44), and the contribution of position-dependent macroscopic strains (in particular, acoustic waves) from (1.4).
6. With spherical symmetry it is convenient to presume $E(p)$ to be a monotonic function, in (2.19), etc; but no loss of generality is thereby entailed, as one could use p instead of E as the independent variable. In the special situation that $f(\mathbf{p})$ can be taken to be appreciable only within a narrow region in which $\mathbf{p}$ is nearly parallel to $\boldsymbol{\mathcal{E}}$, the single parallel component of momentum is an equivalent variable: See Refs. 15, 33.
7. This still allows W to depend on the angle between $\mathbf{p}_1$ and $\mathbf{p}_2$, as in (2.20). The special case of ellipsoidal surfaces of constant $E(\mathbf{p})$ *and* W a function of $E(\mathbf{p}_1)$ and $E(\mathbf{p}_2)$ only is equivalent to complete isotropy.
8. Otherwise the corresponding equation obtained from (2.10 will not give $f_0 \rightarrow f_B$ when $\mathcal{E} \rightarrow 0$.)
9. L. D. Landau and A. Kompane'ets, *Physik. Zeits. Sowjetunion* **6,** 163 (1934); B. Davydov, *ibid.* **9,** 433 (1936), and **12,** 269 (1937). The first two of these papers contain material algebraic errors.
10. See Footnote 17. For a "tensor mass" valley, strictly speaking all scattering must be isotropic; but for moderate anisotropy there is an approximate applicable generalization of (2.9): see C. Herring and E. Vogt, *Phys. Rev.* **101,** 944 (1956).
11. The maximum energy of the phonon emitted or absorbed is small compared to the electron energy and to kT.
12. B. Davydov and I. Shmushkevich, *J. of Physics* (USSR) **3,** 359 (1940); J. Yamashita and M. Watanabe, *Prog. Theor. Phys.* **12,** 443 (1954); H. G. Reik and H. Risken *Phys. Rev.* **124,** 777 (1961).
13. E. M. Conwell and M. O. Vassell, *Phys. Rev.* **166,** 797 (1968); D. Matz, *Phys. Rev.* **168,** 843 (1968).
14. G. A. Baraff, *Phys. Rev.* **133,** A26 (1964); L. Stenflo, *Proc. IEEE* **54,** 1970 (1966).
15. W. P. Dumke, *Phys. Rev.* **167,** 783 (1968).
16. I. B. Levinson (1964). Translated version: *Soviet Physics—Solid State* **6,** 1665 (1965). Eqs. (3.18) and (3.19)—with (3.4), essentially the result given by him—apply when $2kT$ is appreciably greater than the predominant energy change in scattering.
17. In general for (2.28) to be exact scattering must be isotropic, over a surface of constant energy, for each final energy *not* equal to the initial energy. Then $1 - \tau/\tau'$ is equal to the average, over the final states, of the direction cosine y defined in (2.23).
18. C. Herring, *Bell System Technical Journal* **34,** 237 (1955).
19. Here $m_{\parallel}$ and $m_{\perp}$ are the principal values of $\mathbf{m}$. (The former is in the rotational-axis direction, which for n-Ge is (1,1,1).) The values used here are 1.59 and 0.0815, respectively, times the free electron mass.
20. These values give quite good agreement of calculated with experimental drift velocity versus field [the most recent available data are reported in J. E. Smith, Jr., *Phys. Rev.* **178,** 1364 (1969)], up to about 6000 V/cm; but it should be noted that the agreement is spurious, in that the present results show that before the field exceeds 2000 V/cm there would be substantial spill over of electrons into the (1,0,0) minima, located $\sim 5kT_0$ above the band edge. Also Dumke (private communication) finds that the influence of "nonparabolicity" of the electron energy function, and of the energy of acoustic-mode phonons which contribute to the scattering, should be appreciable in the conditions in question.
21. See the paper by A. D. Falkoff and K. E. Iverson in *Interactive Systems for Experimental Applied Mathematics,* M. Klerer and J. Reinfelds, eds., Academic Press, New York 1968.
22. Their import changes, however. The position of $\langle X\rangle \equiv I\,(E/kT_0)\, f_0$ indicated in Fig. 3 shows that an appreciable fraction of the electrons has energy below the bend value. As $\langle X\rangle$ increases with increasing field $\mathcal{E}$ (by almost a factor of 10, from 1000 to 5000 V/cm) this ceases to be so, and deviations from the description and results given by (3.40) should become unimportant. It was in fact found that down to the lowest fields used in the iterative computation (600 V/cm) the computed values of $\langle X\rangle$ were close to $3T_e/2T_0$ given by (3.40) with the first term on the right replaced by one, i.e., by (3.29). Agreement with the analytical formulas represented by (3.40) is good down to lower fields than one would have anticipated; but there is no assurance that this will hold in other cases, for example at lower temperatures T.
23. The computed drift velocity values were 2% below those given by (3.40), except at the bottom of the field range (600 to 5000 V/cm) of the former. The computed values were accordingly adjusted upward by 2%, in constructing Fig. 6, in lieu of accounting for this disparity. At 1000 V/cm the adjusted computed value agreed with the "high field" value; while by 700 V/cm it lay midway between the "high field" and "low field" values, which differed by 7%.
24. For the opposite limit, of a linear "pencil" distribution,[15] one may obtain an equation like (3.14), but simpler, by applying the truncation procedure of Baraff[14] in lowest order: setting $f_1 = 3f_0$, and substituting this in (3.8). Then (3.5) gives $Fvf_0 = -J_S\{f_0\}$. Solutions of this equation can entail negative T_e values.
25. See for example the chapter by Beeler in *Physics of Many Particle Systems; Methods and Problems,* E. Meeron, ed., Gordon and Breach, New York, 1966.

26. Applications to the analogous problem for ions in gases are described in T. Itoh and T. Musha, *J. Phys. Soc. Japan* **15,** 1675 (1960); M. J. Bell and M. D. Kostin, *Phys. Rev.* **169,** 150 (1968).
27. T. Kurosawa, *Proceedings of the International Conference on the Physics of Semiconductors,* Kyoto 1966, paper X-6.
28. Boardman, Fawcett and Rees, *Solid State Communications* **6,** 305 (1968). Also Fawcett, Hilsum and Rees, *ibid.* **7,** 1257 (1969).
29. This equivalence has a parallel for some calculations of Ohmic conductivity. It appears that for other calculations which have been given a Monte Carlo formulation a transformation to the corresponding iterative procedure could be advantageous.

    Because the distribution function obtained by Monte Carlo is proportional to the number of arrivals in a "cell," N_c say, its error $\sim N_c^{-1/2}$ may be much greater than for an iterative procedure with a comparable amount of numerical processing. On the other hand, in a problem where not all of the particle variables are of interest in the final result one may accomplish a saving in memory requirement, with Monte Carlo, by not storing successive values of such variables after they have served their purpose by influencing the generation of the subsequent values.
30. P. J. Price, *Bull. Am. Phys. Soc.* **4,** 129 (1959).
31. H. F. Budd, *Proceedings of the International Conference on the Physics of Semiconductors,* Kyoto 1966, paper X-5. For the particular case considered by Budd one has an *exact* "one-dimension" equation involving the E variable only, but it is of a more complicated form than the equations of Section 3.
32. H. D. Rees, *Physics Letters* **26A,** 416 (1968); *J. Physics and Chemistry of Solids* **30,** 643 (1969). It should be noted that the normalization of the Legendre polynomials in these papers differs from the $P_l(1) = 1$ used here.
33. P. J. Price, *IBM Research Report RW-98,* published in *Proceedings of the International Conference on the Physics of Semiconductors,* Moscow 1968, paper XV-1. Among misprints in the published version, it should be noted that the operator on the right of Eq. (5) should be "G," not "C"; and in the fifth line after Eq. (2) "states after" should read "states before and after." There should be no minus sign in the exponent of Eq. (7); the exponent will then be negative.
34. P. J. Price, *Proceedings of the International Conference on the Physics of Semiconductors,* Rochester 1958, paper E6. (An error in Eqs. (8) and (9) is corrected in Ref. 35.)
35. P. J. Price, Chapter 8 in *Fluctuation Phenomena in Solids,* R. E. Burgess, ed., Academic Press, New York, 1965.
36. H. D. Rees, *Solid State Communications* **7,** 267 (1969).
37. H. D. Rees, *IBM J. Res. Develop.* **13,** 537 (1969).
38. This peculiar combination of μ-space and γ-space, in which for each particle its collision times are "used as a clock," so that the time evolution of the system can be specified with this synchronization of particle events, has been named by the writer a *synchronous ensemble.*
39. It is not quite necessarily so. The hot-electron system in question may not have a normalizable steady state. (This situation corresponds to convergence of the integral in Eq. (3.4) when the upper limit is infinity.[16]) One does not deal with an infinite energy range in practice; and in lieu of actual departures from the assumed model of the solid (additional bands, scattering processes, avalanche processes, etc., coming in at higher energies) which quench the "escape effect" instability, one might choose to introduce the energy cutoff in such a way as to represent a system having the stability property. One should not exclude the possibility of the thermalization rate being very small for some region of **p**-space (perhaps tending to zero for $E \rightarrow \infty$), with corresponding behavior of the f_a and f_b functions in the iteration scheme (4.4), (4.5). But in any case it is stability on iteration of the actual computational algorithm, imperfectly representing the analytical operations, and not of the latter, which is in question. The general existence, and importance, of the underlying stability are clear, however. They have been particularly stressed by Rees.
40. For an operator G and functions ϕ, ψ in general, the adjoint G^* satisfies $I\phi(G\psi) = I\psi(G^*\phi)$. The adjoint of an integral operator is obtained by interchanging the variables in the kernel, and the adjoint of $\partial/\partial\mathbf{p}$ is $-\partial/\partial\mathbf{p}$.
41. Here, and in the equations following, $\mathbf{F}$ can be the full Lorentz force (1.4), though we are going to be concerned only with the case (1.11).
42. We might base the iteration scheme on (4.12), just as well as on (4.3). In the work of Rees, because of his "self scattering" artifice in the Boltzmann equation, there is no difference between these two procedures.
43. Equivalently, [ ] τ is the inverse of A. For the justification of these statements see P. J. Price, *IBM J. Res. Develop.* **2,** 200 (1958). See also L. V. Keldysh, *Soviet Physics—JETP* **21,** 1135 (1965), in particular Eq. (23).
44. For diffusion, the source term on the right of (4.15) is $-\mathbf{v}\cdot\text{grad}\, f$. For hot electrons the response coefficient, the diffusivity, is not in general given simply by the expectation of some electron variable, as in (4.18) for the differential mobility. The theory of diffusion of hot electrons is discussed in Refs. 34, 35. For calculations of the hot-electron diffusivity tensor, see Patrick Hu and P. J. Price, *IBM Research Report NW-18* (1967) (for the classic case of Ref. 9); G. Persky and D. J. Bartelink, *Physics Letters* **28A,** 749 (1969) (p-germanium); W. Fawcett and H. D. Rees, *Physics Letters* **29A,** 578 (1969) (n-GaAs).
45. In papers of the writer cited here, C is denoted by $\mathfrak{M}$ and its adjoint by $\mathfrak{L}$, and Z is denoted by Q.
46. The nth term of (4.19) is the expectation of $\int\psi(\mathbf{p}|t)dt$ integrated from the $(n-1)$th scattering to the nth scattering following the initial state, with $\psi(\mathbf{p}) = \mathbf{v}(\mathbf{p}) - \mathbf{u}$.
47. H. R. Skullerud, *J. Physics D* (*British Journal of Applied Physics,* Series 2) **1,** 1567 (1968).
48. In particular, the "scattering time" replacing $\tau(\mathbf{p})$ in all functions and operators will be a constant, τ_0. The resulting saving in arithmetical operations in the numerical implementation of A, etc., will be offset by the greater number of iterations required for a given precision.
49. Private communication.
50. The one-dimension scheme of the preceding paragraph is, of course, applicable to this case, and has been applied to it by Budd.[31]

Received June 2, 1969