Microscopic mobility

Morrel H. Cohen, E. N. Economou,* and C. M. Soukoulis†
Corporate Research Science Laboratories, Exxon Research and Engineering Company, Route 22 East, Clinton, Annandale, New Jersey 08801
(Received 4 October 1983; revised manuscript received 24 April 1984)

An explicit formal expression is obtained for the energy-dependent microscopic mobility of a disordered system of independent electrons. It is this mobility which gives the average drift velocity of electrons of a given initial energy in response to an electric field. Both this microscopic mobility and the more familiar microscopic conductivity are local in energy. The microscopic mobility is proportional to the energy derivative of the microscopic conductivity, not to the microscopic conductivity itself, which leads to the scaling relation $s=\tau+1$, where s is the conductivity exponent and τ the mobility exponent. The absurdity of a mobility divergent at the mobility edge leads to the restriction $s\geq 1$. In the presence of interactions, values of $s<1$ can occur. If a quasiparticle representation is valid, $0\leq\tau<s$ and the density of states diverges as the mobility edge is approached from the extended-state side with an exponent $u=-[\tau+(1-s)]$. The microscopic mobility provides a dynamical basis for the introduction of the hole concept into the theory of strongly disordered materials. Comparison with the thermopower then leads to a deeper understanding of the nature of holes in disordered materials. A carrier representation can then be derived for transport in disordered semiconductors.

I. INTRODUCTION

The macroscopic dc electrical conductivity of a disordered system can be written as

$$\sigma_M=\int dE\,\sigma(E)\left[-\frac{\partial f}{\partial E}\right] \tag{1}$$

in the independent-electron picture.[1] In (1) $\sigma(E)$ is an energy-dependent, *microscopic* conductivity and f is the occupation number of a single-particle state of energy E. In obtaining (1) via, e.g., the Kubo formalism,[2] the dynamical effects of the electron-electron interaction, the electron-phonon interaction, and spin-flip scattering have been ignored. In the low-temperature limit, when the conduction electrons are degenerate, Eq. (1) reduces to $\sigma_M=\sigma(E_F)$. $\sigma(E)$ has been studied in great detail in recent years in the context of localization theory,[3] and for disordered metallic systems much progress has been made in understanding the behavior of $\sigma(E)$ within the independent-electron picture[3,4] as well as the modifications of $\sigma(E)$ and reinterpretation of (1) following from the inclusion of the omitted interactions.[5]

Far less attention has been devoted to the fundamental transport theory of disordered semiconductors. In contrast to the metallic case, there are many experiments carried out on semiconductors in which it is more convenient or even necessary to focus on carrier mobility rather than conductivity.[6] Consequently, one finds in the semiconductor literature,[7–10] equations of the form

$$\sigma_M=\sum_\alpha\int dE\,n_\alpha(E)e\mu_\alpha(E)f_\alpha(E)\ , \tag{2}$$

where α is a band index ($\alpha=c$ for the conduction band and v for the valence band), $n_\alpha(E)$ is the density of states per volume, $\mu_\alpha(E)$ the microscopic *carrier* mobility, and $f_\alpha(E)$ the *carrier* occupation number, the carriers being holes in the valence band and electrons in the conduction band. The microscopic carrier mobility gives the average speed $v_\alpha(E)$ obtained by carriers of given energy in linear response to an electric field F,

$$v_\alpha(E)=\mu_\alpha(E)F\ . \tag{3}$$

The *conductivity* mobilities[6] $\bar{\mu}_\alpha$ are the weighted averages of the microscopic mobilities

$$\sigma_M=\sum_\alpha\sigma_\alpha\ , \tag{4a}$$

$$\sigma_\alpha=N_\alpha e\bar{\mu}_\alpha\ , \tag{4b}$$

$$N_\alpha=\int dE\,n_\alpha(E)f_\alpha(E)\ , \tag{4c}$$

$$\bar{\mu}_\alpha=\frac{\int dE\,n_\alpha(E)\mu_\alpha(E)f_\alpha(E)}{\int dE\,n_\alpha(E)f_\alpha(E)}\ . \tag{4d}$$

Drift mobilities measured in injection experiments of various kinds are also defined as in (4), but with $f_\alpha(E)$ a nonequilibrium carrier distribution.[6] This conceptual structure, developed for crystalline semiconductors, has been carried over to the theory of disordered semiconductors.[1] In the process, important issues have not been addressed.

For example, what is the general expression for the microscopic carrier mobility defined through (3) analogous to that obtained through the Kubo formalism for $\sigma(E)$? What is the general relation between the microscopic carrier mobility and $\sigma(E)$? Mott and Davis[12] have proposed for disordered semiconductors, when $f_\alpha(E)$ is the Boltzmann distribution, that

$$\mu_\alpha(E)=\sigma_\alpha(E)/en_\alpha(E)k_BT\ . \tag{5}$$

However, (5) cannot be correct in general because in the absence of interactions, $\mu_\alpha(E)$, $\sigma_\alpha(E)$, and $n_\alpha(E)$ must all be temperature independent. Finally, how does one intro-

 ©1984 The American Physical Society

duce the hole concept for disordered systems to which conventional band theory does not apply?

We give definitive answers to the above questions in the present paper for systems of independent electrons, ignoring by dynamical effects of the Coulomb interaction the electron-phonon interaction and spin-flip scattering. Using the Kubo formalism we are able to define and obtain an exact expression for the microscopic mobility for *electron* states $\mu(E)$, which gives the average velocity $\vec{v}(E)$ induced in electrons occupying states of energy E by a field $\vec{F}$,

$$\vec{v}(E)=\overleftrightarrow{\mu}(E)\cdot\vec{F} . \tag{6}$$

We find the relation

$$n(E)e\overleftrightarrow{\mu}(E)=-\frac{d\overleftrightarrow{\sigma}(E)}{dE} \tag{7}$$

between $\overleftrightarrow{\mu}$ and $\overleftrightarrow{\sigma}(E)$. The relation (7) has important implications for scaling behavior in the vicinity of a mobility edge. One can show, for example, that the conductivity exponent must exceed or equal unity. The microscopic mobility then provides a direct route for the introduction of the hole concept in the presence of arbitrary disorder. We find that states are electronlike when $d\sigma(E)/dE$ is positive and $\mu(E)$ negative, and holelike when $d\sigma(E)/dE$ is negative and $\mu(E)$ positive. Thus one has electrons in extended states above a mobility edge and holes in extended states below a mobility edge, as is relevant to disordered semiconductors. We show that this dynamic criterion for the classification of states is related to the statistical classification provided by the thermopower. Simple rules follow for the relations between quantities defined for electron states and for carriers.

A scaling relation holding at the mobility edge between the exponent for the microscopic mobility and the microscopic conductivity follows from (7). It implies that the conductivity exponent cannot fall below unity. In the presence of interactions, however, it is known to fall below unity.[13] If a quasiparticle representation is valid and the conductivity exponent falls below unity, then one infers from (7) that the density of quasiparticle states must show scaling behavior with a negative exponent.

In sum, the microscopic mobility is a useful and powerful concept in the theory of disordered materials.

II. REVIEW OF CONDUCTIVITY THEORY

Consider a system of N-independent electrons in a region of volume Ω. Its macroscopic electrical conductivity $\sigma_M(\omega)$ at frequency ω can be expressed as[2]

$$\overleftrightarrow{\sigma}_M(\omega)=\frac{e^2}{i\hbar\Omega}\int_0^\infty dt\, e^{-i\omega t}\mathrm{Tr}\{\vec{v}(t)[f,\vec{r}]\} . \tag{8}$$

In (8), $\vec{r}$ is the one-electron position operator, $\vec{v}(t)$ is the one-electron velocity operator at time t in the interaction representation, and f is the occupation operator. If H is the one-electron Hamiltonian, f can be taken as an *arbitrary* function of H. The eigenvalues of f give the occupation numbers of the corresponding eigenstates of H when the electric field is initially turned on. The theory is not constrained to initial states of thermodynamic equilibrium, but holds for any initial state describable by $f=f(H)$ and therefore stationary in the absence of interactions. ω has a small negative imaginary part $-i\alpha$. N and Ω are to be taken to infinity and α to zero in such a way that $n=N/\Omega$ remains fixed and the number of eigenvalues of H in an energy range $\hbar\alpha$ remains of order N. The dc conductivity $\sigma_M(0)=\sigma_M$ is obtained by letting the real part of ω tend to zero before α in (8).

We now write out (8) explicitly in the representation generated by the eigenfunctions of H, ψ_l:

$$H\psi_l=\epsilon_l\psi_l , \tag{9}$$

$$\overleftrightarrow{\sigma}_M=\frac{\pi e^2\hbar}{\Omega}\sum_{k,l}\vec{v}_{kl}\vec{v}_{lk}\delta_\alpha(E_k-E_l)\left[\frac{-\partial f(E_k)}{\partial E_k}\right] . \tag{10}$$

In (10), $f(E_k)$ is the occupation number of state k. The δ function $\delta_\alpha(E_k-E_l)$ is to be treated as a distribution in the sense of Schwartz, that is, as a sequence of good functions (Lorentzians in this case) of width $\hbar\alpha$ decreasing to zero.[14] The formula (10) can be converted to Eq. (1) by inserting the factor

$$\int dE\,\delta_\beta(E-E_k)=1 \tag{11}$$

inside the summation in (10), thus obtaining

$$\overleftrightarrow{\sigma}(E)=\frac{\pi e^2\hbar}{\Omega}\sum_{k,l}\delta_\beta(E-E_k)\vec{v}_{kl}\vec{v}_{lk}\delta_\alpha(E_k-E_l) \tag{12}$$

for $\overleftrightarrow{\sigma}(E)$. It is essential to recognize that the width α of $\delta_\alpha(E_k-E_l)$ in (10) is taken to zero before $\delta_\beta(E-E_k)$ is inserted so that

$$\lim_{\substack{\alpha\to 0\\ \beta\to 0}}\frac{\alpha}{\beta}=0 . \tag{13}$$

The expression (12) for $\overleftrightarrow{\sigma}(E)$, or an equivalent one, has been used as the starting point of many investigations of the influence of disorder on the energy-dependent conductivity.[12,15] The specific results of interest here are that in the isotropic case

$$\overleftrightarrow{\sigma}(E)=\sigma(E)\overleftrightarrow{I} , \tag{14}$$

$\sigma(E)$ shows scaling behavior above (or below) a mobility edge E_c,

$$\sigma(E)\propto(E-E_c)^s , \tag{15}$$

and vanishes below (or above) it in dimension $d>2$, as shown in Fig. 1. The strong energy dependence of $\sigma(E)$ in (15) is associated with the fractal character[16] of the ψ_k with energies E_k near E_c on length scales less than the fluctuation length[3] $\xi(E)$,

$$\xi(E)\propto(E-E_c)^{-\nu} . \tag{16}$$

The conductivity $\sigma(E)$ is related to ξ through[3]

$$\sigma(E)\propto\xi^{-(d-2)} \tag{17}$$

so that the scaling relation[3]

$$s=(d-2)\nu \tag{18}$$

holds for the exponents. Expansion of ν in powers of $d-2$ gives

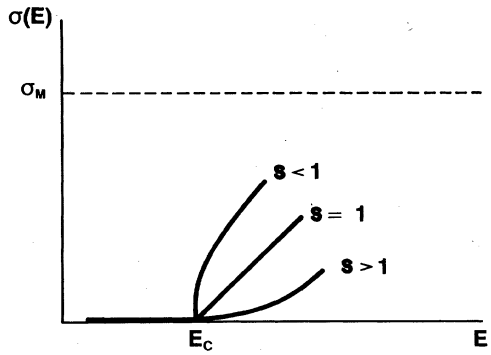

FIG. 1. Energy dependence of the dc conductivity $\sigma(E)$, Eqs. (2), (12), (14), and (15), is shown for $s<1$, $=1$, >1 near the mobility edge E_c.

$$\nu=(d-2)^{-1} \tag{19}$$

through fifth order.[17] Thus our best estimate for the value of s is

$$s=1\ , \tag{20}$$

independent of dimension. However, qualitative arguments and numerical studies of ν give values of s different from (20), both below and above unity.[3,4]

III. MICROSCOPIC MOBILITY

Time-reversal invariance implies that in the absence of an electric field $\vec{F}$ the average velocity of all states ψ_l of energy $E_l=E$ must vanish. Suppose now that $\vec{F}$ is turned on adiabatically. Independent electrons initially in those states would acquire an average velocity linear in $\vec{F}$ if it is sufficiently small,

$$\vec{v}(E)=\overleftrightarrow{\mu}(E)\cdot\vec{F}\ , \tag{6}$$

where $\overleftrightarrow{\mu}(E)$ is the microscopic mobility for electron states of energy E. We shall later derive an explicit expression for $\overleftrightarrow{\mu}(E)$ from linear-response theory, i.e., the Kubo formalism.[2] Now, we explore the consequences of the existence of the relation (6). The total current density induced by the field is obtained from (6) as follows: by multiplying by the electron charge, by the occupation number $f(E)$ for states of energy E which, we stress, can be arbitrary, by the density of states per unit volume $n(E)$, and then by integrating over energy. An expression closely analogous to Eq. (2) follows for $\overleftrightarrow{\sigma}_M$:

$$\overleftrightarrow{\sigma}_M=-\int dE\, n(E)e\overleftrightarrow{\mu}(E)f(E)\ . \tag{21}$$

Equation (21) is of a form familiar to those who have been actively concerned with transport in semiconductors, but it can be puzzling to those familiar primarily with transport in metals. It appears from (21) that in the degenerate case, carriers below the Fermi energy E_F contribute to the conductivity, which is patently absurd. However, as we shall show below, the absurdity is only apparent and not real. An exactly analogous situation occurs in the high-frequency conductivity of a metal in which the plasma frequency enters. The latter can be expressed as proportional to the inverse effective mass averaged over occupied states with the appearance of occupied states contributing to the response of the metal to an applied field. However, because the inverse effective mass is the derivative of the electron velocity with respect to the wave vector, an integration by parts can be made which converts the average of the inverse effective mass over occupied states into an average of the velocity over the Fermi surface, thus showing that only states at the Fermi surface actually contribute to the current, despite appearances. An analogous resolution of the apparent difficulty with (21) is now shown to hold.

Let us return to Eq. (1) and integrate it by parts. Taking the lower limit of integration below the K-shell energy implies that $\overleftrightarrow{\sigma}(E)$ vanishes, and taking the upper limit of integration at infinity implies that $f(E)$ vanishes, with the result

$$\overleftrightarrow{\sigma}_M=\int dE\frac{d\overleftrightarrow{\sigma}(E)}{dE}f(E)\ . \tag{22}$$

Because $f(E)$ can be taken as an arbitrary function of energy, the factors multiplying $f(E)$ in the integrands of (21) and (22) can be equated, yielding (7):

$$n(E)e\overleftrightarrow{\mu}(E)=-\frac{d\overleftrightarrow{\sigma}(E)}{dE}\ . \tag{7}$$

Inserting (7) into (21) and integrating by parts reproduces (1), showing in analogy with the case of the plasma frequency that the contribution of states below E_F to the current in the degenerate case is only apparent and not real. Nevertheless another apparent contradiction follows from (7) which cannot be resolved through the use of properties of the occupation number.

We integrate (7) to obtain

$$\overleftrightarrow{\sigma}(E)=-\int^{E} dE'\, n(E')e\overleftrightarrow{\mu}(E')\ . \tag{23}$$

Equation (12) shows that $\overleftrightarrow{\sigma}(E)$ is a local function of energy, containing contributions only from states on the energy shell E and varying smoothly with E except at a singular point like E_c. Equation (7) implies that $\overleftrightarrow{\mu}(E)$ is also a smooth, local function of energy. Equation (23) then seems to imply that $\overleftrightarrow{\sigma}(E)$ is nonlocal in energy, contradicting both intuition and Eq. (12). Once again the contradiction is only apparent. The integrand on the right-hand side of (23) is a total derivative, so that the only net contributions come from the limits of integration and, in particular, the upper limit because the integral vanishes at the lower.

We now explore these relationships between $\overleftrightarrow{\mu}(E)$ and $\overleftrightarrow{\sigma}(E)$ more explicitly. To derive an explicit microscopic expression for $\overleftrightarrow{\mu}(E)$, we consider the system of N electrons initially in a microcanonical ensemble centered at E,

$$f(H)=\frac{N}{N(E)}\delta_\beta(E-H)\ , \tag{24}$$

where

$$N(E)=\Omega n(E) \tag{25}$$

is the total density of states. Linear-response theory then gives for the average velocity adiabatically induced among

the electrons by a field $\vec{F}$,

$$\vec{v}(E)=\frac{-e}{i\hbar N}\int_0^\infty dt\, e^{-\alpha t}\mathrm{Tr}\{\vec{v}(t)[f,\vec{r}]\}\cdot\vec{F}\ , \tag{26}$$

where $\vec{v}(E)$ is the same quantity that enters (6). The cyclic invariance of the trace permits us to write

$$\overleftrightarrow{\mu}(E)=\frac{e}{i\hbar N}\int_0^\infty dt\, e^{-\alpha t}\mathrm{Tr}[[\vec{v}(t),\vec{r}]f] \tag{27}$$

with f specifically given by (24) in this particular instance. In the representation diagonalizing H, (27) becomes

$$\overleftrightarrow{\mu}(E)=-\frac{2\pi e\hbar}{N(E)}\sum_{k,l}\delta_\beta(E_k-E)\frac{\vec{v}_{kl}\vec{v}_{lk}}{E_l-E_k}\delta_\alpha(E_l-E_k)\ . \tag{28}$$

For the turning of the field to have been adiabatic, $f(H)$ must be smooth on the energy scale defined by $\hbar\alpha$ so that Eq. (13) must hold in the present case as well. Interchanging the dummy indices k and l in (28), making use of time-reversal invariance as well as hermiticity, we can rewrite (28) as

$$\overleftrightarrow{\mu}(E)=-\frac{\pi e\hbar}{N(E)}\sum_{k,l}\frac{\delta_\beta(E_k-E)-\delta_\beta(E_l-E)}{E_l-E_k}\times\vec{v}_{kl}\vec{v}_{lk}\delta_\alpha(E_l-E_k)\ . \tag{29}$$

In view of (13), the summand in (29) is nonvanishing only for values of (E_l-E_k) much smaller than $\hbar\beta$, the width of $\delta_\beta(E_k-E)$. The first factor in the summand therefore reduces to an energy derivative

$$\overleftrightarrow{\mu}(E)=\frac{-\pi e\hbar}{N(E)}\sum_{k,l}\frac{d\delta(E_k-E)}{dE}\vec{v}_{kl}\vec{v}_{lk}\delta_\alpha(E_l-E_k)\ . \tag{30}$$

This is the desired microscopic expression for the microscopic mobility. Comparing the two microscopic expressions, Eq. (12) for the microscopic conductivity and Eq. (30) for the microscopic mobility, one sees that they are related by Eq. (7). Whereas Eq. (7) had previously been derived by comparing two expressions, Eqs. (1) and (21) for the macroscopic conductivity $\overleftrightarrow{\sigma}_M$ holding for arbitrary $f(E)$, we have now demonstrated that Eq. (7) holds at the level of the dynamic processes underlying the linear response to a field.

All that remains to complete our analysis of $\overleftrightarrow{\mu}(E)$ is to derive Eq. (21) via linear-response theory. Returning to Eq. (8), we pass to the dc limit $\mathrm{Re}\omega \ll \alpha$ and once again utilize the cyclic invariance of the trace. The result is

$$\overleftrightarrow{\sigma}_M=\frac{e^2}{i\hbar\Omega}\int_0^\infty dt\, e^{-\alpha t}\mathrm{Tr}[[\vec{r},\vec{v}(t)]f]\ , \tag{31}$$

with f once again an arbitrary function of H. Introducing the representation diagonalizing H and integrating over time leads directly to (21), with $\overleftrightarrow{\mu}(E)$ given by (30). Thus (21), intuitively obvious from (6), follows directly from linear-response theory.

The locality of $\overleftrightarrow{\mu}(E)$ in energy is clearly displayed by Eq. (30) as is that of $\overleftrightarrow{\sigma}(E)$ by (12). Moreover Eq. (7) is a local relation between these local quantities. Substitution of (30) into (23), an apparently nonlocal expression, immediately yields (12), a completely local expression, because the integrand is a total derivative.

IV. SCALING RELATIONS

We have seen in Sec. II that the conductivity $\sigma(E)$ scales with energy near the mobility edge, Eq. (15). If that is true for $\sigma(E)$, it is also true for $\mu(E)$, where

$$\overleftrightarrow{\mu}(E)=\mu(E)\overleftrightarrow{I} \tag{32}$$

in isotropic materials. Let the mobility exponent be τ, that is,

$$|\mu(E)|\sim(E-E_c)^\tau\ . \tag{33}$$

Equation (7) imposes a scaling relation on the exponents s and τ, viz.,

$$s=\tau+1 \tag{34}$$

because $n(E)$ is smooth at E_c for independent electrons.[3] The cases $s>1$, $s=1$, and $s<1$ thus correspond to $\tau>0$, $\tau=0$, and $\tau<0$.

We can immediately rule out the case $s<1$ on physical grounds because $\tau<0$ implies that $\mu(E)$ diverges at E_c. That would imply, according to (6), a divergent carrier velocity for carriers at the mobility edge in a small electric field. Inasmuch as the extended states continuously approach the localized states on the other side of the mobility edge, an infinite drift velocity would be absurd. We therefore arrive at the restriction

$$\tau\geq 0,\quad s\geq 1\ . \tag{35}$$

Arguments or calculations applicable to the present case of a system of noninteracting electrons moving in a static potential which yield values of s less than unity must be incorrect. If a value of $s<1$ were to be observed experimentally at a mobility edge well separated from the Fermi energy, we could infer that electron-electron interactions, electron-phonon interactions, spin-flip scattering, or some combination of these is playing a role in the conductivity. In the presence of interactions, the entire single-particle formalism must be replaced. However, the replacement can occur in stages. In the first stage, a quasiparticle formalism is valid after the introduction of interactions.[18] Both Eqs. (1) and (21) still hold, with $f(E)$ the quasiparticle occupation number and $n(E)$, $\overleftrightarrow{\mu}(E)$, and $\overleftrightarrow{\sigma}(E)$ potentially strongly influenced by the interactions. Equation (7) must also still hold, but Eq. (34) can no longer hold if $s<1$. We immediately infer that when $s<1$, the quasiparticle density of states $n(E)$ must show scaling behavior at E_c, i.e.,

$$n(E)\propto(E-E_c)^u \tag{36}$$

near E_c on the extended-state side. Equation (35) is then replaced by

$$s=u+\tau+1\ . \tag{37}$$

Equation (35) is replaced by

$$\tau\geq 0\ . \tag{38}$$

A value of s less than unity implies a negative value of u,

$$u = -[\tau + (1-s)] , \tag{39}$$

that is a peak in the density of states at E_c. $\sigma(E)$ must remain finite at E_c, so that

$$s \geq 0 , \tag{40}$$

and $n(E)$ must remain integrable at E_c, so that

$$u > -1 . \tag{41}$$

Equations (40) and (41) imply that

$$0 \leq \tau < s , \tag{42}$$

i.e., that s is an upper bound on τ. In the second and final stages, the quasiparticle formalism does not hold, Eqs. (1) and (21) are no longer simple integrals over energy of quantities local in energy, and our fundamental equation, Eq. (7), can no longer hold. As a consequence, neither Eq. (34) nor its generalization, Eq. (37) can be expected to hold.

As yet, no experimental determination exists of any value of s at a mobility edge well separated from any other mobility edge and from the Fermi energy, let alone a value of $s<1$. However, a power-law dependence of $\sigma(E)$ near E_F, $\sigma(E) \propto |E - E_F|^s$ has been inferred in doped crystalline semiconductors at the metal-nonmetal transition. The value $s \cong 0.5$ in uncompensated samples has been imputed to electron-electron interactions.[13] A value of 0.5 would also be obtained for spin-flip scattering.[5] s increases to about 0.8 in compensated samples. It is important to note that at the metal-nonmetal transition, the mobility gap vanishes. The case we have considered, that of a mobility edge well separated from any other mobility edge by a mobility gap, is quite different. The dip in the density of states at the Fermi energy[13] at the transition is presumably the precursor to the Mott-Hubbard gap. The singularity in $n(E)$ at E_c we would predict when $s<1$ would thus seem implausible on the nonmetal side of the transition. There could well be a crossover in the doped semiconductor from $s \cong 0.5$ at the transition to $s \geq 1$ on the nonmetal side.

Returning to the independent-particle picture, Mott has proposed in the past that $\sigma(E)$ has a quasiuniversal step of magnitude[12,3]

$$\Delta\sigma = \frac{1}{3\pi^2} \frac{e^2}{\hbar a} \tag{43}$$

at E_c in three dimensions. Equation (43) implies that $s=0$, violating Eq. (35). A conductivity step is thus impossible in the single-particle picture.[3] However, if we accept the field-theoretic results [Eq. (20)] for s as correct, namely, $s=1$, then $\tau=0$ and the microscopic mobility either has a step or goes as an inverse logarithm. The step would be nonuniversal. For small disorder $d\sigma(E_c)/dE$ is inversely proportional to W^x, where W measures the strength of the disorder potential and the value of x has been estimated to be 2.[3,19] In the case where the mobility edge continuously approaches the band edge as the disorder vanishes, $n(E_c)$ is proportional to W^y, where $y=2$.[19] Thus the mobility step increases as $W^{-(x+y)}$ with decreasing disorder, 4 being a reasonable estimate for $x+y$. In view of Eq. (21), a mobility step is sufficient to produce all the qualitative features produced from (1) by a conductivity step. An inverse logarithm would be essentially as effective.

V. HOLES IN DISORDERED MATERIALS

One must know which of the many beautiful and important results of the electron theory of solids derived explicitly for crystalline materials with at most weak disorder survive the introduction of disorder of arbitrary strength. The hole concept is of particular importance for disordered semiconductors. The introduction of the hole concept proceeds along two avenues for crystalline materials, statistical and kinematic or dynamic. In the statistical considerations, one considers holes simply as unoccupied electron states. This is in itself sufficient for the proof that holes contribute positively to the thermopower and electrons negatively. On the other hand, in metals at low temperatures, where the kinematics and dynamics of the particles at the Fermi surface dominate the properties of the material, the statistical argument becomes ambiguous, as portions of the Fermi surface can bound either regions of occupied states from above or regions of unoccupied states from below. The ambiguity is resolved by kinematic or dynamical considerations which, while remarkably rich in their consequences, depend in an essential way on the periodicity of the crystal. In Sec. V A, we review briefly the known statistical considerations carried out through use of the thermopower. In Sec. V B, we use Eq. (6) as a basis for dynamic considerations and show how it relates to the statistical considerations of Sec. V A.

A. Statistical considerations: Thermopower

Consider a system with a given ground state. An excited state with changes of occupancy in the one-electron states can be said to have holes in the states previously occupied in the ground state and particles in the previously unoccupied states, a notion which goes back to the beginnings of quantum mechanics. This notion found a most direct application in the theory of the thermopower of crystals in which the negatively charged electrons have been shown to contribute negatively and the positively charged holes positively, the distinction being only whether the states in question are above or below the chemical potential or Fermi energy E_F.

Heuristic derivations by Cutler and Mott[7] and by Fritzsche[8] show that the theory of the thermopower is unchanged in strongly disordered systems. They find that the thermopower S is given by

$$S = \frac{-k_B}{e\sigma_M} \int dE\, \sigma(E) \frac{E - E_F}{k_B T} \left[\frac{-\partial f}{\partial E} \right] . \tag{44}$$

In (44), $\sigma(E)$ and $(-\partial f/\partial E)$ are positive (in thermodynamic equilibrium), so that states with $E > E_F$ (electrons) contribute negatively and those with $E < E_F$ contribute positively (holes).

It is a simple matter to derive (44) from rigorous transport theory.[20] The general theory yields

$$S = \frac{1}{T} \left[\frac{L_{12}}{L_{11}} + \frac{E_F}{e} \right] , \tag{45}$$

where L_{ij}, $i=1,2$, are correlation functions which can be written in the energy representation as

$$L_{ij}=\frac{\pi\hbar}{\Omega}\sum_{k,l}(L_i)_{kl}(L_j)_{lk}\delta_\alpha(E_l-E_k)\left[\frac{-\partial f(E_k)}{\partial E_k}\right], \quad (46)$$

$$L_1=-ev_x,\ \ L_2=\tfrac{1}{2}(Hv_x+v_xH)\ . \quad (47)$$

Inserting (47) into (46) yields

$$L_{11}=\sigma_M\ , \quad (48)$$

$$L_{12}=\int dE\,\sigma(E)E\left[\frac{-\partial f}{\partial E}\right] \quad (49)$$

with $\sigma(E)$ given by (10). Inserting (48) and (49) into (46) yields (44).

Mott and Davis[12] have pointed out that (44) reduces to

$$S=\frac{\pi^2}{3}\frac{k_B^2T}{e}\frac{d\ln\sigma}{dE}\bigg|_{E=E_F} \quad (50)$$

in the low-temperature limit when transport occurs only within about k_BT of the Fermi energy. We are not primarily interested in such situations, being concerned more with behavior near the mobility edges of disordered semiconductors. However, (50) is of interest in connection with the dynamic considerations of the next section.

B. Dynamical considerations: Microscopic mobility

Among the very interesting results of the electron theory of solids derived for crystals is that a dc electric field cannot induce a dc current in a single filled band. That results holds also for arbitrarily disordered materials; it follows immediately from (1) because $f(E)$ is unity throughout a filled band, independent of $\sigma(E)$. Equation (21), however, yields the more interesting result that

$$\int dE\,n(E)\mu(E)=0 \quad (51)$$

when the integral is taken over a complete band. Equation (23) yields the still more interesting result that

$$\int_{E_{c1}}^{E_{c2}} dE\,n(E)\mu(E)=0\ , \quad (52)$$

where E_{c1} and E_{c2} are the lower and upper limits to a continuous region of extended states and are therefore mobility edges.

Equation (52) implies that, because $n(E)$ is positive definite in the interval E_{c1},E_{c2}, $\mu(E)$ must change sign within E_{c1},E_{c2}. That this can be the case is obvious from (7) because $d\sigma(E)/dE$ must be positive near E_{c1} and negative near E_{c2}. The "normal" case of a negative $\mu(E)$ implies an electron drift velocity opposite to the direction of the applied field, according to (7). On the other hand, a positive $\mu(E)$ implies an electron drift velocity in the direction of the applied field, which is the direction of response of positively charged particles. Thus, extended states above a mobility edge are electronlike, and extended states below a mobility edge are holelike.

The above exact results have been assumed in applications of the Mott—Cohen-Fritzsche-Ovskinsky model,[21,22] but they had not previously been proved. Moreover, they apply more generally than simply to mobility edges near band edges, as would be the case for amorphous or disordered semiconductors. Whenever $d\sigma(E)/dE$ changes sign from positive to negative within a region of extended states, the states of energy E change from electronlike to holelike, as shown in Fig. 2.

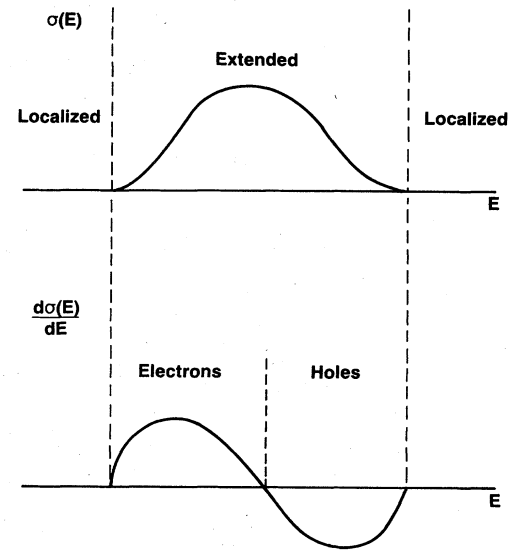

FIG. 2. Energy dependence of the conductivity $\sigma(E)$ within a continuous region of extended states and of $d\sigma(E)/dE$, showing the electron and hole states within that region.

It is usually assumed that the sign of the thermopower reflects the character of the carriers dominating the transport, positive for holes and negative for electrons. As mentioned above, this distinction becomes ambiguous at low temperatures when transport is dominated by carriers at the Fermi energy. Equations (50) and (7), however, show that the thermopower can then be written as

$$S=\frac{\pi^2}{3}k_B^2T\frac{n(E_F)\mu(E_F)}{\sigma(E_F)}\ . \quad (53)$$

Our dynamic criterion based on (6) that electrons have negative $\mu(E)$ and holes have positive $\mu(E)$ is seen to be identical to the statistical criterion as expressed through the thermopower at very low temperatures, (53). When $\mu(E_F)$ is positive, the Fermi energy is to be regarded as a lower bound to a region of states primarily occupied by holes, and, when $\mu(E_F)$ is negative, it is to be regarded as the upper bound to a region of states primarily occupied by electrons, the same as for crystals but without some of their complexities.

C. Carrier representation

We are now in a position to give a rigorous derivation of Eq. (2) without reference to any results derived for crystalline semiconductors in the presence of weak disorder. We suppose that there are clearly distinguishable conduction and valence bands, $\alpha=v$ and c, respectively, so that states of given energy and nonvanishing $\sigma(E)$ or $\mu(E)$ can be uniquely assigned either to the valence or conduction band. The density of states can then be broken down into valence and conduction contributions

$$n(E)=\sum_{\alpha} n_{\alpha}(E)=n_c(E)+n_v(E) \tag{54}$$

with $n_c(E)$ and $n_v(E)$ overlapping at most within the mobility gap between the mobility edges E_v and E_c. The charge carriers are electrons in the conduction band and holes in the valence band. The carrier mobility in the conduction band is defined as

$$\mu_c(E)=-\mu(E)=|\mu(E)| \ , \tag{55}$$

$\mu(E)$ being negative in the conduction band in the relevant energy range. The carrier occupation number is simply the state occupation number,

$$f_c(E)=f(E) \ . \tag{56}$$

In the valence band, the microscopic mobility is already positive in the relevant energy region near E_v, so that the hole mobility is simply

$$\mu_v(E)=\mu(E) \ . \tag{57}$$

On the other hand, the hole occupation number is given by

$$f_v(E)=1-f(E) \ . \tag{58}$$

Insertion of (55)—(58) into (21) leads directly to (2) after use of (51).

In the past it has been supposed that the microscopic carrier mobility $\mu_\alpha(E)$ in (2) is proportional to the microscopic conductivity $\sigma_a(E)$ from incorrect arguments such as the following. For Boltzmann statistics, as is appropriate for disordered semiconductors, Eq. (1) becomes

$$\sigma_M=\sum_{\alpha}\int dE\, \sigma_\alpha(E) f_\alpha(E)/k_B T \ . \tag{59}$$

Comparing (59) with (2) tempts one to equate integrands. If one were to do so, one would obtain

$$\mu_\alpha(E)=\sigma_\alpha(E)/e n_\alpha(E) k_B T \ . \tag{60}$$

This result is patently incorrect because the arguments used to obtain it would also hold in the absence of interactions, in which case $\mu_\alpha(E)$, $\sigma_\alpha(E)$, and $n_\alpha(E)$ are all temperature independent. The flaw lies in equating the integrands when only the integrals are equal. The correct relationship between $\mu_\alpha(E)$ and $\sigma_\alpha(E)$ is given by (7). It can be obtained from (1) and (22) by equation of the integrands because both (1) and (22) hold for arbitrary $f(E)$, whereas (59) holds only for the Boltzmann distribution. Deviations observed experimentally from the predictions of the single-particle theory presented here can with confidence be interpreted as the consequences of interactions. Electron-electron interactions and spin-flip scattering have been shown to be of great importance for localization effects in metals. We have recently shown[23] that the electron-phonon interaction is of great importance in amorphous semiconductors in two distinct ways. First, a mobility tail is induced below the mobility edge at finite temperatures through phonon-assisted hopping. With the proper inclusion of the mobility tail, transport theory has been brought into agreement with the observations for the first time. Second, polaron formation can occur above but near the mobility edge even for weak electron-phonon coupling. Both the Coulomb interaction and the electron-phonon interaction are of importance for doubly occupied localized states, which can be bipolarons. These effects of interaction can be so complex in amorphous semiconductors that a firm grasp of the consequences of the theory without interactions, as provided in part here, is a necessary prerequisite for further progress.

VI. CONCLUSION

Our most significant results are the following: the relation (7) between the energy-dependent conductivity and the microscopic mobility (6), the consequent limitation of the conductivity exponent to unity or larger, (34) or (35), and the dynamic definition of the hole states through the sign of $\mu(E)$ together with its relation to the statistical definition through the thermopower, e.g., (53). These results have all been obtained in the independent-electron picture, ignoring both Coulomb and electron-phonon interactions and spin-flip scattering. The results are of particular usefulness for amorphous semiconductors.

*Permanent address: Department of Physics, University of Crete and Research Center of Crete, Heraklio, Crete.

†Present address: Department of Physics, Iowa State University, Ames, IA 50011.

[1]O. Madelung, *Introduction to Solid State Theory* (Springer, New York, 1981), p. 370.

[2]R. Kubo, Can. J. Phys. **34**, 1274 (1956); D. A. Greenwood, Proc. Phys. Soc. **71**, 585 (1958).

[3]F. Wegner, Z. Phys. B **25**, 327 (1976); **35**, 207 (1979); E. Abrahams, P. W. Anderson, D. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979); Y. Imry, *ibid.* **44**, 469 (1980); D. Vollhardt and P. Wolfle, *ibid.* **48**, 699 (1982); A. Mackinnon and B. Kramer, *ibid.* **47**, 1546 (1981), Z. Phys. B **53**, 1 (1983).

[4]B. Shapiro and E. Abrahams, Phys. Rev. B **24**, 4889 (1981); S. Sarker and E. Domany, *ibid.* **23**, 6018 (1981).

[5]W. L. McMillan, Phys. Rev. B **24**, 2739 (1981); B. L. Altshuler, D. Khmelnitskii, A. I. Larkin, and P. A. Lee, *ibid.* **22**, 5142 (1980); H. Fukuyama, J. Phys. Soc. Jpn. **48**, 2169 (1980); B. L. Altshuler and A. G. Aronov, Solid State Commun. **46**, 429 (1983); P. A. Lee, J. Non-Cryst. Solids **35**, 21 (1980); A. Kawabata, J. Phys. Soc. Jpn. **50**, 2461 (1981).

[6]See, for example, *Amorphous Semiconductors,* edited by M. H. Brodsky (Springer, New York, 1979), p. 113.

[7]M. Cutler and N. F. Mott, Phys. Rev. **181**, 1336 (1969).

[8]H. Fritzsche, Solid State Commun. **9**, 1813 (1971).

[9]M. H. Cohen, J. Non-Cryst. Solids **4**, 391 (1970).

[10]G. H. Döhler, Phys. Rev. B **19**, 2083 (1979); H. Fritzsche, Sol. Energy Mater. **3**, 447 (1980).

[11]One also finds expressions like (2) without the band index but with an additional factor of $[1-f(E)]$ (see Refs. 6 and 12). This is incorrect in independent-particle approximation, as will emerge in Secs. II and III.

[12]N. F. Mott and E. A. Davis, *Electronic Processes in Noncrystalline Material* (Clarendon, Oxford, 1979).

[13]G. S. Grest and P. A. Lee, Phys. Rev. Lett. **50**, 693 (1983) and

references therein.

[14]M. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, New York, 1958).

[15]For a review, see E. N. Economou, *Green's Functions in Quantum Physics* (Springer, New York, 1983), Chap. 7.

[16]C. M. Soulkoulis and E. N. Economou, Phys. Rev. Lett. **52**, 565 (1984).

[17]S. Hikami, Phys. Rev. B **24**, 2671 (1981).

[18]D. Pines, *The Many Body Problem* (Benjamin, New York, 1962).

[19]C. M. Soukoulis, M. H. Cohen, and E. N. Economou (unpublished).

[20]W. Jones and N. H. March, *Theoretical Solid State Physics* (Wiley-Interscience, London, 1973), Vol. 2, p. 685ff, p. 757ff.

[21]N. F. Mott, Philos. Mag. **13**, 989 (1966); Adv. Phys. **16**, 49 (1967); M. H. Cohen, H. Fritzsche, and S. R. Ovshinsky, Phys. Rev. Lett. **22**, 1065 (1969).

[22]D. J. Thouless, in *Ill Condensed Matter,* edited by R. Balian, R. Maynard, and G. Toulouse (North-Holland, Amsterdam, 1978); Phys. Rep. C **13**, 94 (1974).

[23]M. H. Cohen, E. N. Economou, and C. M. Soukoulis, Phys. Rev. Lett. **51**, 1202 (1983).