Microscopic Theory of Tunneling: General Theory and Application to the Static Impurity*

C. B. Duke, G. G. Kleiman,† and T. E. Stakelon
Department of Physics, Materials Research Laboratory, and Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801
(Received 11 February 1972)

A microscopic theory of impurity-assisted tunneling is constructed in which the current-carrying (extended) eigenstates of the average one-electron potential in a tunnel junction are utilized as the basis functions which are mixed by the Hamiltonian associated with the presence of a static or dynamic impurity anywhere in the system. In such a system, the one-electron propagator and its concomitant current across the junction can be calculated at zero bias by standard techniques for manipulating temperature Green's functions. The presence of a finite bias across the junction is incorporated into the theory via the principle of rigid occupancy; i.e., the equilibrium occupation of the (current-carrying) many-body eigenstates of the system is taken to be unaffected by the presence of the bias. Therefore the rigid-occupancy hypothesis relates the nonequilibrium current flow to the equilibrium (zero-bias) properties of the junction system. Thus, we obtain a theory of nonequilibrium current flow which is not based on linear-response theory. This hypothesis is incorporated into the Matsubara perturbation theory by treating the chemical potentials of the left- and right-hand electrodes as *separate* Lagrangian mulitpliers determined after the completion of all Matsubara sums to be related by $\mu_R = \mu_L - eV$. Therefore, for purposes of constructing and solving Dyson's equations for the renormalized one-electron propagators, the theory reduces to the conventional equilibrium theory defined using distorted-wave (i.e., non-plane-wave) states. The perturbation theory is shown to yield the conventional one-electron-theory results for the case of a static impurity potential in the barrier. Resonant elastic tunneling through impurity states of energy E_r near the zero-bias Fermi energy ζ causes conductance minima (maxima) for $E_r(0) > \zeta$ [$E_r(0) < \zeta$]. The transfer-Hamiltonian results are recovered by expanding the transmission probability.

I. INTRODUCTION

The transfer-Hamiltonian theory[1,2] of current flow in a tunnel junction originally was developed to explain observations of the superconducting energy gap and bulk density of states in normal-metal–oxide–superconductor tunnel junctions.[3] It provided the theoretical framework within which Josephson[4] made his prediction that phase coherence exists between two weakly coupled superconductors. One of the most recent successful applications of the transfer Hamiltonian has been the interpretation of cusps occurring in the conductance measured in metal-oxide-homopolar semiconductor junctions (e.g., heavily doped In-SiO_2-Si junctions[5,6]) in terms of "self-energy effects"[6-9] due to the interactions of electrons and holes with optical phonons in the bulk semiconductor electrode.

Regardless of its successes the original momentum-representation version of the transfer-Hamiltonian (TH) theory exhibits certain conceptual weaknesses and seems incapable of describing certain recent experiments.[10-18] It is formulated[19] by considering a tunnel junction to consist of two isolated electrodes, between which electrons are interchanged by matrix elements which form the "transfer Hamiltonian." The tunneling current is evaluated by a linear-response analysis in which the transfer Hamiltonian is treated as an "external" source.[19] There are several difficulties with this procedure. First, the set of basis functions used in constructing the linear-response analysis is overcomplete, since these basis functions are direct products of the basis functions of the two isolated electrodes,[19] each of which individually forms a complete set. Second, the TH theory is formulated so that the occupation numbers of electrons in its many-body eigenstates are those of the isolated-electrodes system.[19] Consequently, the linear-response analysis is restricted to a first-order-perturbation-theory treatment[20] of the electron transfer probability from one electrode to the other. A related difficulty with the linear-response analysis is that the diagrammatic expansion of both the elastic and inelastic contributions to the current in the TH theory[19] contains only diagrams which represent the transfer of an electron from one of the electrodes to the other. Thus, the predicted current does not include contributions from "reflection" diagrams, which describe processes in which an electron localized in one of the electrodes is reflected back into the same electrode after interacting in the barrier. Such processes contribute to the total current. They are important when, for example, an electron interacts with a local mode[19] near one of the electrode-barrier interfaces. Also,

the TH theory does not describe processes in which the electrons in both electrodes interact with the same (nonlocal) boson elementary-excitation field.[19]

A final difficulty with the linear-response analysis results from the asymmetric fashion in which it treats electrode and barrier effects.[19] Each of the electrodes is described by a Hamiltonian which contains dynamic interactions of electrons localized in that electrode with, e.g., bulk phonons. Barrier effects, on the other hand, are described solely through potential matrix elements which couple electrons in both electrodes. Thus, the energies of electrons in the electrodes are renormalized by the interactions with bulk phonons, and the electron dispersion relations are reflected in the conductance as self-energy effects.[6] The matrix elements for the transfer of electrons from one electrode to the other, however, are insensitive to the renormalization of the tunneling electrons' energies by many-body interactions.[8] Another consequence of the description of barrier effects solely by transfer matrix elements is the resulting inability of the linear-response treatment to describe processes which involve real intermediate states[15,19] in the barrier. In the case of elastic processes,[19] for example, the current is evaluated to first order in the barrier transmission coefficient $|S_{12}|^2$ and, therefore, is inaccurate in the case of resonant elastic-tunneling processes,[21] for which $|S_{12}|^2 \sim 1$.

The conceptual difficulties noted above are directly related to the fact that the TH theory has proven incapable of describing certain recent experiments. For example, observations of light-sensitive conductance characteristics in GaAs tunnel diodes,[10] doping-dependent reverse-bias phonon line shapes in the conductance measured in a variety of semiconductor junctions,[10-12] and impurity-band effects in metal-insulator-semiconductor junctions[13] suggest that resonant impurity-assisted processes in the barrier are important tunneling mechanisms. Other (related) experimental effects whose description has eluded analysis via the TH are the observations of light-sensitive subharmonic structure in the current in Josephson junctions,[14] real intermediate-state tunneling current[15,16] in metal-insulator-metal junctions in which small metal particles are embedded in the oxide, and resonant field emission[17,18] due to adatoms in metal-vacuum systems.

As one might expect, these fairly well-recognized difficulties with the original formulations[1,2] of the TH theory have motivated several workers to attempt to correct them. We classify these attempts into three broad categories: improved versions of the TH theory, collision theories, and Green's-function theories.

The most recent improved version of the TH theory is that of Appelbaum and Brinkman.[8] They attempted to remedy part of the asymmetry between the description of "barrier" and "electrode" effects by constructing a coordinate representation (Green's-function version) of the TH theory. In their discussion, they formally include the possibility of assisted tunneling through barrier excitations. Their result takes the form of a convolution of the product of the two spectral functions[22] of the isolated electrodes with a transfer matrix element which depends upon both energy and momentum. Therefore, their results retain the form expected for elastic tunneling and, consequently, still exhibit the difficulties associated with overcomplete sets (i.e., the possibility of overcounting) and linear-response theory (i.e., the restriction to first-order perturbation theory).

The only serious attempt to construct a many-electron collision theory is that of Davis.[7] A simpler version, in which no explicit provision is made for the many-body (i.e., exclusion-principle-imposed) aspects of the theory has been given by Adler *et al.*[23] However, Davis's calculation is algebraically quite complicated. It is valid only at zero temperature and restricted to a particular model (a vibrating δ-function impurity). No account is given of resonant-tunneling phenomena which we believe dominate the impurity-assisted inelastic processes he considered.[24] As will be discussed in a following paper,[25] the theory developed in this paper, when applied to his model, predicts results similar but not identical to his.

Turning to the Green's-function theories, the first was that of Zawadowski[26] which led to results identical to Appelbaum and Brinkman's coordinate-representation version[8] of the TH theory. The second, which is conceptually more closely related to our analysis than any of the others, was that of Davis.[9] By deriving a form for the current analogous to that of the TH theory,[19] he was able to show that an electron tunnels with the total renormalized energy and not the bare energy.[22] Most recently, Caroli *et al.*[27] have proposed a nonequilibrium Green's-function theory based on a cellular model in which the hopping matrix element connecting one part of the junction to the next is small. This theory differs substantially from its predecessors. It has not yet been explored in enough depth to recognize any new consequences which it might have and to examine their experimental realization.[28]

The sketch of the literature presented above is designed to indicate the nature of the difficulties with the TH theory, the experimental manifestations of the phenomena whose description is precluded by these difficulties, and the attempts to avoid them which have been presented previously or independently. This is the first in a series of two papers in which we also develop a new micro-

scopic theory of tunneling. In the present paper, we derive the theory and calculate the current for tunnel junctions which contain static impurities. In a subsequent paper, [25] we derive the current for barriers containing vibrating impurities and study the consequences of static impurity potentials which possess "quasibound" states and which produce a resonance in the barrier transmission coefficient. We explore their effects upon both the one-electron and many-body contributions to the current in that paper. The fact that using our theory we have analyzed successfully previously uninterpreted phenomena in semiconductor tunnel diodes has been documented elsewhere. [24]

The physical assumptions from which the theory presented in this paper is derived differ from those underlying the transfer-Hamiltonian theory. These assumptions are: (i) The tunneling electrons are not in states localized in either electrode, [19] but are in doubly degenerate current-carrying eigenstates of the whole junction, [29,30] one of which carries current from left to right (left incident) and the other carries current from right to left (right incident); and (ii) the statistics of current flow in the junction are not determined by transfer of an electron from a filled state in one electrode to an empty one in the other, [19] but are determined by the occupancies of left-incident and right-incident states, which are separately associated with the left and right electrodes, respectively. In this "stationary-state" picture, if both left- and right-incident states of total energy E are occupied, there is no net current contribution from these states because the current from left to right exactly cancels that from right to left. We refer to this assumption as the "rigid-occupancy hypothesis" (ROH). It results in a model of the tunnel junction in which the only effect of an imposed bias (aside from changing the one-electron potential) is that of rigidly translating the energies of the eigenstates associated with the left and right electrodes up and down with respect to each other. In the absence of many-body interactions, this picture is precisely the conventional one-electron stationary-state model. [19]

The ROH is applied via a temperature Green's-function technique[22] in which the chemical potentials μ_L and μ_R of the left and right electrodes, respectively, of a junction with a voltage imposed across it are replaced by Lagrangian multipliers λ_L and λ_R, respectively, whose difference is a boson index[22]: $\lambda_L - \lambda_R = 2\pi m/-i\beta$, in which m is any (i.e., positive or negative) integer and $\beta \equiv 1/k_B T$ (k_B is Boltzmann's constant and T is the temperature). The current is calculated using intermediate steps depending only on equilibrium perturbation theory, after which the Lagrangian multipliers are analytically continued to their physical values so that the total current flowing from left to right is associated with μ_L, while that from right to left is associated with μ_R: $\lambda_L \to \mu_L$ and $\lambda_R \to \mu_R$ (i.e., $\lambda_L - \lambda_R \to eV$ according to the conventions in Fig. 1). The value of this procedure is that the diagram techniques of equilibrium-temperature Green's-function[22] theory can be applied prior to the analytic continuation of the Lagrangian multipliers (because it is not necessary to know the occupation numbers of electrons carrying current either to the right or to the left in this first step of the calculation). Therefore, the central feature of our theory is the replacement of the complete nonequilibrium transport problem with a prescription for using equilibrium propagators plus an analytic continuation procedure to evaluate a nonequilibrium current. The theory is viable because of the special feature of a *tunnel* junction that although current does flow, it is sufficiently small that the externally imposed voltage drop occurs across the barrier region of the junction.

We proceed by specifying some preliminary results about the basis functions in Sec. II. Then in Sec. III, we develop a many-body theory of tunneling in which the currents are calculated in equilibrium, after which the bias is imposed via the ROH by making the analytic continuation $\lambda_L - \lambda_R \to eV$. The theory developed in Sec. III is applied to a simple δ-function model of static impurities in Sec. IV. In that section it is shown that the perturbation theory developed in Sec. III produces the correct expression for the one-electron current in the presence of static impurities. The applications

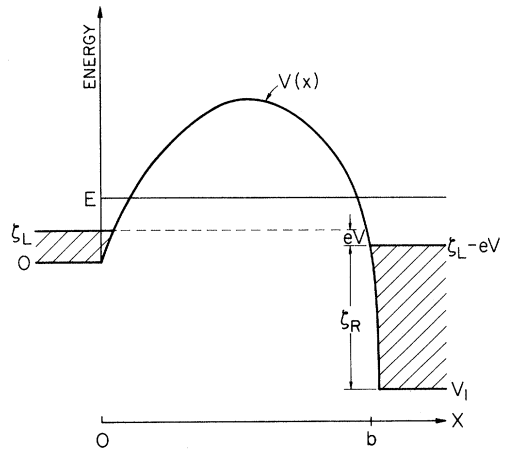

FIG. 1. Schematic typical tunneling junction illustrating the conventions used. The average barrier potential is denoted by $V(x)$ $(0 \le x \le b)$ and all energies are measured from the bottom of the left conduction band. The bottom of the right conduction band is $V_1 \equiv \zeta_L - eV - \zeta_R$. Positive or forward bias is defined by raising the Fermi level of the left electrode relative to that of the right.

of the theory to describe vibrating impurities[25] and experimental data[24] are given elsewhere.

II. EXTENDED BASIS FUNCTIONS

Two fundamental physical assumptions underlie the tunneling theory which is developed in this paper. The first of these is that the electrons in a biased tunnel junction are in states of the whole junction, rather than localized in either electrode. The second assumption is that a tunnel junction across which a voltage has been imposed is a steady-state system. The mathematical formulation of the first of these assumptions is given in this section.

A. Definition and General Properties of Extended-Basis Functions

According to the assumption that the electrons are in extended states of the whole junction, we construct such states as the solutions of the Schrödinger equation for the whole junction. We consider, as our one-electron Hamiltonian, a model consisting of electrodes which can be described by the effective-mass approximation[19] separated by a barrier such as that shown in Fig. 1. In this model, which is convenient, even if there is some question[19] as to its validity, the Schrödinger equation of the eigenstates corresponding to the total energy E is specified by

$$h_0(\vec{r})\,\chi(E,\vec{r}) = E\chi(E,\vec{r})\ . \tag{2.1a}$$

The one-electron Hamiltonian h_0 is defined by

$$h_0(\vec{r}) \equiv \begin{cases} -(\hbar^2/2m_L)\nabla^2, & x<0 \\ -(\hbar^2/2m_B)\nabla^2 + V(x), & 0\le x\le b \\ -(\hbar^2/2m_R)\nabla^2 + V_1, & x>b\ . \end{cases} \tag{2.1b}$$

Equations (2.1b) are derived from the assumption that the effective masses are isotropic in the left (m_L) and right (m_R) electrodes and in the barrier (m_B) region. The barrier potential is represented by $V(x)$ and, from Fig. 1,

$$V_1 \equiv \zeta_L - eV - \zeta_R. \tag{2.1c}$$

The quantities ζ_L and ζ_R are the Fermi degeneracies of the left and right electrodes, respectively. We adopt the conventions, indicated in Fig. 1, that energy is measured from the bottom of the band in the left-hand electrode, and $eV>0$ (forward bias) when the left Fermi level is raised with respect to the right.

Under the assumption of specular tunneling,[19] the eigenstates are specified by

$$\chi(E,\vec{r}) = \psi(E,\vec{k}_{\parallel},x)\,e^{i\vec{k}_{\parallel}\cdot\vec{\rho}} \equiv \chi(E,\vec{k}_{\parallel},\vec{r})\ , \tag{2.1d}$$

in which $\vec{k}_{\parallel}$ and $\vec{\rho}$ are the components of $\vec{k}$ (total momentum) and $\vec{r}$ (position), respectively, in the plane of the junction. The traveling waves in the direction parallel to the planar junction are normalized to $(2\pi)^2\delta(\vec{k}_{\parallel}-\vec{k}'_{\parallel})$.

For $E>0$ and $V_1<0$, the ψ and χ are doubly degenerate.[29] We choose the linearly independent set of solutions defined in the following equations[30]:

$$\psi_L(E,\vec{k}_{\parallel},x) = \frac{A_{EL}}{(2\pi\hbar v_L)^{1/2}} \times \begin{cases} (e^{ikx} + S_{11}e^{-ikx}), & x<0 \\ [C_L^+U_+(x) + C_L^-U_-(x)], & 0<x<b \\ (v_L/v_R)^{1/2}S_{21}e^{iq(x-b)}, & b<x \end{cases} \tag{2.2}$$

$$k^2 \equiv (2m_LE/\hbar^2) - k_{\parallel}^2\ , \qquad v_L \equiv \hbar k/m_L \tag{2.3a}$$

$$q^2 \equiv [2m_R(E-V_1)/\hbar^2] - k_{\parallel}^2, \qquad v_R \equiv \hbar q/m_R \tag{2.3b}$$

$$\left(-\frac{\hbar^2}{2m_B}\frac{d^2}{dx^2} + V(x) + \frac{\hbar^2k_{\parallel}^2}{2m_B} - E\right)U(x) = 0, \quad 0<x<b. \tag{2.4}$$

In the above equations, A_{EL} is a normalization constant and $U_{\pm}(x)$ are decaying (growing) solutions of Eq. (2.4)[31] [for instance, the $U_{\pm}(x)$ are parabolic cylinder functions for Schottky barriers, Airy functions for trapezoidal barriers, and exponentials for square barriers].[19]

The function $\psi_L(E,\vec{k}_{\parallel},x)$ defined by Eq. (2.2) is an extended-basis function (EBF) which carries current from left to right (left incident). Similarly, we define an EBF which carries current from right to left (right incident) for $E>0$ and $V_1<0$:

$$\psi_R(E,\vec{k}_{\parallel},x) = \frac{A_{ER}}{(2\pi\hbar v_R)^{1/2}} \times \begin{cases} (v_R/v_L)^{1/2}S_{12}e^{-ikx}, & x<0 \\ [C_R^+U_+(x) + C_R^-U_-(x)], & 0<x<b \\ [e^{-iq(x-b)} + S_{22}e^{iq(x-b)}], & b<x\ . \end{cases} \tag{2.5}$$

In the case that $0>E>V_1$, we obtain the wave function in this case by letting $ik\to k$ in Eqs. (2.5) (ψ_R does not carry current for $0>E\ge V_1$). Thus, the set $\{\psi_L,\psi_R\}$ is complete for all $E\ge V_1$. In accordance with the effective-mass approximation, the wave functions are obtained by imposing the matching conditions[19] that ψ_i and $\psi_i'/m(x)$ $(i=L,R)$ be continuous across the left-electrode–barrier and barrier–right-electrode interfaces [$m(x)$ represents the effective mass in the different regions of the junction, and ψ' denotes the derivative of ψ].

When we let $A_L=A_R=1$, one can derive[32]

$$\langle\psi_i(E,\vec{k}_{\parallel})\mid\psi_j(E',\vec{k}_{\parallel})\rangle = \delta_{ij}\delta(E-E'), \quad (i,j)=(L,R). \tag{2.6}$$

Since any two of the functions $\{\psi_L,\psi_L^*,\ \psi_R,\psi_R^*\}$ (ψ^* is the complex conjugate of ψ) are linearly independent,[33] we can define Wronskians for $E>0$:

$$W_{ab}(E,\vec{k}_{\parallel},x) \equiv [\psi_a'(E,\vec{k}_{\parallel},x)\,\psi_b(E,\vec{k}_{\parallel},x)$$

$$-\psi_a(E,\vec{k}_{\parallel},x)\psi_b'(E,\vec{k}_{\parallel},x)]\,, \quad (2.7)$$

where $(a,b)=(L,R)$ and a subscript a^* denotes ψ_a^*. From the continuity of $\psi'/m(x)$ throughout the junction, we see that the quantity

$$W_{ab}(E,\vec{k}_{\parallel},x)/m(x)\equiv W_{ab}(E,\vec{k}_{\parallel})/m_i \quad (2.8)$$

is constant throughout the junction, where the subscript i indicates that W_{ab} is evaluated in region i (left, barrier, or right region).

B. Single-Particle Propagator

Turning next to the evaluation of the (three-dimensional) Green's function, we first note that the equations for the basis states can be written

$$[h_0(\vec{r})-E]\chi_i(E,\vec{k}_{\parallel},\vec{r})=0, \quad i=(L,R) \quad (2.9)$$

where χ_i is specified by Eqs. (2.1)–(2.5).

Since the set of functions $\{\chi_L,\chi_R\}$ is complete, we can define a one-electron Green's function

$$G_0(\vec{r},\vec{r}',z)\equiv\sum_{E_i,\vec{k}_{\parallel}}\frac{\chi_i(E,\vec{k}_{\parallel},\vec{r})\,\chi_i^*(E,\vec{k}_{\parallel},\vec{r}')}{z-E_i}\,, \quad (2.10)$$

in which the symbol E_i is the label for both the total energy of a state and the direction in which it carries current. The summation conventions are specified by

$$\sum_{\vec{k}_{\parallel}}\to\int\frac{d^2k_{\parallel}}{(2\pi)^2}, \quad \sum_{E_i}\to\int_0^{\infty}dE+\int_{V_1}^{\infty}dE\equiv\sum_{E_L}+\sum_{E_R} \quad (2.11)$$

for fixed $\vec{k}_{\parallel}$.

The advanced and retarded Green's functions G^A and G^R, respectively, are given by

$$G_0^R(\vec{r},\vec{r}',z)\equiv\lim_{\delta\to 0^+}G_0(\vec{r},\vec{r}',z+i\delta)\,,$$
$$G_0^A(\vec{r},\vec{r}',z)\equiv\lim_{\delta\to 0^+}G_0(\vec{r},\vec{r}',z-i\delta). \quad (2.12)$$

We can Fourier analyze $G_0^{R,A}(\vec{r},\vec{r}',z)$ according to Eq. (2.1), since we assume $\vec{k}_{\parallel}$ conservation, and derive

$$G_0^{R,A}(\vec{r},\vec{r}',z)=\sum_{\vec{k}_{\parallel}}e^{i\vec{k}_{\parallel}\cdot(\vec{\beta}-\vec{\beta}')}G_0^{R,A}(x,x',\vec{k}_{\parallel},z), \quad (2.13a)$$

$$G_0^{R,A}(x,x',\vec{k}_{\parallel},z)\equiv\sum_{E_i}\frac{\psi_i(E,\vec{k}_{\parallel},x)\psi_i^*(E,\vec{k}_{\parallel},x')}{z-E_i\pm i\delta}, \quad i=(L,R)\,. \quad (2.13b)$$

The quantity $G_0^{R,A}(x,x',\vec{k}_{\parallel},z)$ satisfies

$$[h_0(x)-z]G_0^{R,A}(x,x',\vec{k}_{\parallel},z)=-\delta(x-x')\,, \quad (2.14)$$

where $h_0(x)$ is given by

$$h_0(x)\equiv\begin{cases}-\dfrac{\hbar^2}{2m_L}\dfrac{d^2}{dx^2}+\dfrac{\hbar^2}{2m_L}k_{\parallel}^2\,, & x<0\\ -\dfrac{\hbar^2}{2m_B}\dfrac{d^2}{dx^2}+\dfrac{\hbar^2}{2m_B}k_{\parallel}^2+V(x), & 0\le x\le b\\ -\dfrac{\hbar^2}{2m_R}\dfrac{d^2}{dx^2}+\dfrac{\hbar^2}{2m_R}k_{\parallel}^2+V_1\,, & x>b\,.\end{cases} \quad (2.15)$$

A more useful form for the Green's function is derived by solving the differential equation [Eq. (2.14)] directly to give,[33] for $E>0$,

$$G_0^R(x,x',\vec{k}_{\parallel},E)=\frac{2\hbar^2}{W_{LR}(E,\vec{k}_{\parallel})/m_i}\times[\Theta(x-x')\psi_L(E,\vec{k}_{\parallel},x)\psi_R(E,\vec{k}_{\parallel},x')+\Theta(x'-x)\psi_R(E,\vec{k}_{\parallel},x)\psi_L(E,\vec{k}_{\parallel},x')]\,, \quad (2.16)$$

and $G_0^A=(G_0^R)^*$. In Eq. (2.16) we choose the boundary conditions that the outgoing waves are $\psi_L(\psi_R)$ for $x\to+\infty$ $(x\to-\infty)$. The quantity Θ is the unit step function [i.e., $\Theta(|x|)=1$, $\Theta(-|x|)=0$, and $\Theta(0)=\frac{1}{2}$].

III. EXTENDED-BASIS-FUNCTION THEORY FOR CURRENT

The extended-basis functions, which reflect the assumption that tunneling electrons are in states of the whole junction, are specified in Sec. II. In this section, using the assumption that a tunnel junction across which a low voltage is imposed is a steady-state system, we construct a description of the statistical properties of the junction, through the use of the temperature-Green's-function formalism.[22]

We first derive an expression for the current as a functional of a perturbation theory of the temperature Green's function.[22] A prescription is given for the evaluation of the current via perturbation theory. This prescription is applied to the calculation of the current in a junction consisting of noninteracting fermions. We show that, in this analytically solvable case, the calculated current is identical to the result of the conventional one-electron tunneling theory[19] in both an average-barrier model and a model in which a static impurity potential is treated by use of our perturbation theory.

In accordance with the assumption that a tunnel junction is a steady-state system, we construct a theory in which, by definition, we neglect changes in the Fermi distribution function due to nonequilibrium effects, and assume that the statistics for electrons carrying current to either the right or left is determined by the equilibrium occupation of either the left or right electrode, respectively.

We refer to this assumption as the rigid-occupancy hypothesis (ROH). The ROH is the hypothesis basic to the one-electron tunneling theory,[19] and gives the lowest-order contribution to the current expanded as a functional of the distribution function.[20]

The total Hamiltonian is given by

$$\mathcal{H}=\mathcal{H}_0(t)+\mathcal{H}_{int}(t) \ , \tag{3.1a}$$

where the operators are defined in the Heisenberg representation:

$$\mathcal{O}(t)\equiv e^{i\mathcal{H}t}\mathcal{O}e^{-i\mathcal{H}t} \ . \tag{3.1b}$$

In the occupation-number formalism, we write

$$\mathcal{H}_0(t)\equiv\mathcal{H}_{00}(t)+\mathcal{H}_{ph}(t) \ , \tag{3.2a}$$

$$\mathcal{H}_{ph}(t)\equiv\sum_{\vec{p},\lambda} b^{\dagger}_{\vec{p}\lambda}(t)\, b_{\vec{p}\lambda}(t)(\hbar\omega_{\vec{p}\lambda}+\tfrac{1}{2}). \tag{3.2b}$$

$\mathcal{H}_{ph}$ is the Hamiltonian of a gas of bosons (e.g., phonons) of momenta $\vec{p}$ and polarization λ ($b_{\vec{p}\lambda}$ destroys a phonon momentum $\vec{p}$ and polarization λ). $\mathcal{H}_{00}$ is the Hamiltonian of a gas of electrons in states $\chi_i (i=L,R)$ of the type defined in Eqs. (2.5):

$$\mathcal{H}_{00}(t)=\sum_{E_i,\vec{k}_{\parallel}} C^{\dagger}_i(E,\vec{k}_{\parallel},t)C_i(E,\vec{k}_{\parallel},t)E_i \quad i=(L,R). \tag{3.3}$$

The quantities $C_i(C^{\dagger}_i)$ are the annihilation (creation) operators for the electrons.

The interaction Hamiltonian $\mathcal{H}_{int}$ may consist of either electron-electron, electron-phonon, or elastic electron-impurity interactions (or some combination thereof). In this paper we study the last case, for which the interaction Hamiltonian assumes the form

$$\mathcal{H}_{int}(t)=\sum_{E_i,E_j,\vec{k}_{\parallel},\vec{k}'_{\parallel}} C^{\dagger}_i(E,\vec{k}_{\parallel},t)C_j(E',\vec{k}'_{\parallel},t) \times\int d^3r\ \chi^{*}_i(E,\vec{k}_{\parallel},\vec{r})V_I(\vec{r}-\vec{R})\chi_j(E',\vec{k}'_{\parallel},\vec{r}), \tag{3.4}$$

where V_I is the static impurity potential and $\vec{R}$ is the impurity's equilibrium position. Only simple "schematic" models of the impurity potentials are considered in this and subsequent papers.

The discrete Fock representation which provides the basis for the Hamiltonians in Eqs. (3.3) and (3.4) is constructed from the continuum states defined in Eqs. (2.2)–(2.5) in the usual manner by subdividing $(E_i,\vec{k}_{\parallel})$ space (i.e., phase space) into unit cells and by letting each one-electron state in the Fock representation correspond to a unit cell in phase space.[34] The continuum limit is retrieved, at the end of the calculation, by letting the volumes of these unit cells approach zero.

The Fock states, which are eigenstates of the electron Hamiltonian $\mathcal{H}_{00}$ given in Eq. (3.3), satisfy the following equations:

$$\mathcal{H}_{00}\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle=E_\alpha\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle \ , \tag{3.5a}$$

$$N^0_L\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle=n^\alpha_L\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle \ , \tag{3.5b}$$

$$N^0_R\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle=n^\alpha_R\,|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle \ , \tag{3.5c}$$

$$|\,\alpha\rangle\equiv|\,E_\alpha,n^\alpha_L,n^\alpha_R\rangle \ , \tag{3.5d}$$

in which α is an index which distinguishes between different eigenstates of $\mathcal{H}_{00}$ and N^0_L and N^0_R are operators which measure the occupation numbers n^α_L and n^α_R of electrons in left- and right-incident EBF, respectively. The number operator N measures the total number of electrons in the system, and in our basis can be decomposed according to

$$N=N^0_L+N^0_R \ . \tag{3.6a}$$

N is a constant of motion and is independent of the representation in which it is expressed. In the occupation-number formalism,

$$N^0_i\equiv\sum_{E_i,\vec{k}_{\parallel}} N^0_i(E,\vec{k}_{\parallel}) \ , \qquad N^0_i(E,\vec{k}_{\parallel})\equiv C^{\dagger}_i(E,\vec{k}_{\parallel})C_i(E,\vec{k}_{\parallel}), \quad i=L \text{ or } R \ . \tag{3.6b}$$

Therefore, $N^0_i(E,\vec{k}_{\parallel})$ is the operator which measures the number of electrons (0 or 1) in the state $(E_i,\vec{k}_{\parallel})$.

The electron many-body Hamiltonian in Eq. (3.3) can be written, in the absence of interactions, in a time-independent form

$$\mathcal{H}_{00}\equiv\mathcal{H}_{0L}+\mathcal{H}_{0R} \ , \tag{3.7a}$$

$$\mathcal{H}_{0i}\equiv\sum_{E_i,\vec{k}_{\parallel}} E_i N^0_i(E,\vec{k}_{\parallel}), \quad i=L \text{ or } R. \tag{3.7b}$$

In the Schrödinger representation, the density operator[35] $\rho(t)$, which contains the statistical properties of a many-body system, obeys the equation

$$\frac{i\hbar\partial\rho(t)}{\partial t}=[\mathcal{H},\rho(t)] \ , \tag{3.8}$$

where $\mathcal{H}$ is the Hamiltonian of the system. In the unperturbed system, characterized by Hamiltonian $\mathcal{H}_{00}$, a time-independent solution of Eqs. (3.8) is given by the density operator

$$\rho(t)=\rho'_0\equiv e^{\beta(\mu_L N^0_L+\mu_R N^0_R)}e^{-\beta\mathcal{H}_{00}} \ , \tag{3.9}$$

which is in the same form as the density operator of a gas of two different independent species of particles ($\beta\equiv 1/k_BT$, k_B is Boltzmann's constant, and T is the temperature). Thus, in the absence of interactions, the electrons in left- and right-incident states are statistically independent. In this case, which represents the unperturbed system, the electron system can be considered as a composite gas of electrons consisting of two statistically independent subsystems of electrons in left- and right-incident states with chemical potentials μ_L and μ_R, respectively. According to the energy conventions of Fig. 1, $\mu_L=\xi_L$ and $\mu_R=\xi_L-eV$.

In the case that the Hamiltonian includes many-body interactions, such as the electron-phonon interaction, we cannot diagonalize the Hamiltonian as in Eqs. (3.7). Thus, we find neither operators N_L^0, N_R^0, in the form of Eqs. (3.6) which commute with the Hamiltonian, nor the density operator which describes the nonequilibrium system of a tunnel junction with a bias imposed. Finding the appropriate number operators for the general nonequilibrium system (here a tunnel junction) is the central unsolved problem of transport theory. The transfer Hamiltonian theory deals with this problem by applying linear-response theory to the decoupled-electrodes system. The density operator of this system[19] involves the number operators which measure the number of electrons in the isolated electrodes.

Our proposed method for solving this nonequilibrium problem is an essentially heuristic analytic continuation method. ROH states that the currents which flow separately from left to right and from right to left in a tunnel junction are the same in the cases when the junction is in equilibrium and when the junction is biased (i.e., only the one electron potential used in the definitions of the EBF changes in these two cases). However, the net current is different in these two situations; i.e., the separate left and right currents cancel exactly in equilibrium, but they do not cancel when a voltage is imposed. Therefore, if we could compute the many-body wave functions of the system, this model would permit us to evaluate nonequilibrium currents using equilibrium methods. In effect, this procedure is the one we follow using the Green's function theory.

The ROH is our model for the statistics in a tunnel junction. From it, we can derive conditions which the expression we derive for the current must satisfy. These conditions are as follows: (i) Since the ROH model describes a steady-state system, the current does not depend upon time. (ii) The perturbation-theory techniques of the temperature-Green's-function method,[17] which describes the junction in equilibrium, also must describe the junction when it is biased. In particular, the perturbation-theory analysis of the one-electron current for static perturbation potentials must agree with the results of the usual analysis in which the exact one-electron basis states are used *a priori* in evaluating the current. (iii) The statistics of current flow to the right (left) are determined by the statistics of the isolated left (right) electrode with chemical potential μ_L (μ_R).

In order to derive the ROH perturbation theory, we introduce the imaginary-time Heisenberg picture. We develop the algebra in detail to display the differences from the equilibrium theory. An operator in the imaginary-time Heisenberg representation is[22]

$$\mathcal{O}(\tau) \equiv e^{\mathcal{H}\tau} \mathcal{O} e^{-\mathcal{H}\tau}, \quad \mathcal{O}^\dagger(\tau) \equiv e^{\mathcal{H}\tau} \mathcal{O}^\dagger e^{-\mathcal{H}\tau}, \tag{3.10}$$

where $0 \leq \tau \leq \beta$ and $\mathcal{H}$ is the total many-body Hamiltonian. An operator in the interaction picture is defined by[22]

$$\hat{\mathcal{O}}(\tau) \equiv e^{\mathcal{H}_0\tau} \mathcal{O} e^{-\mathcal{H}_0\tau}, \quad \hat{\mathcal{O}}^\dagger(\tau) \equiv e^{\mathcal{H}_0\tau} \mathcal{O}^\dagger e^{-\mathcal{H}_0\tau}, \tag{3.11}$$

where $0 \leq \tau \leq \beta$. The transformation from the imaginary-time Heisenberg picture of Eqs. (3.10) to the interaction picture is indicated by

$$e^{-\mathcal{H}\tau} = e^{-\mathcal{H}_0\tau} S(\tau), \tag{3.12a}$$

where $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{int}}$, and $\mathcal{H}_0$ is given in Eqs. (3.2) and (3.3). Therefore,

$$\mathcal{O}(\tau) = S^{-1}(\tau) \hat{\mathcal{O}}(\tau) S(\tau). \tag{3.12b}$$

A tunnel junction in equilibrium is described by the grand canonical density operator

$$\rho_{\text{eq}} \equiv \frac{e^{-\beta(\mathcal{H}-\mu N)}}{\text{Tr}[e^{-\beta(\mathcal{H}-\mu N)}]} = \frac{e^{-\beta(\mathcal{H}_0-\mu N)} S(\beta)}{\text{Tr}[e^{-\beta(\mathcal{H}_0-\mu N)} S(\beta)]}. \tag{3.13}$$

The expression for ρ_{eq} in terms of the unperturbed equilibrium density matrix and $S(\beta)$ is central to the treatment of an equilibrium system in the temperature-Green's-function perturbation theory.[22] From the ROH, we assume that the current flowing in a biased junction also can be described using a temperature-Green's-function formalism. Thus, using Eq. (3.9), we define a density operator ρ which satisfies the following equations:

$$\rho \equiv e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)} e^{-\beta\mathcal{H}} / \mathfrak{z} = \rho_0 S(\beta), \tag{3.14a}$$

$$\rho_0 \equiv e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)} e^{-\beta\mathcal{H}_0} / \mathfrak{z}, \tag{3.14b}$$

$$\mathfrak{z} \equiv \text{Tr}(e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)} e^{-\beta\mathcal{H}}). \tag{3.14c}$$

The quantities N_L^0, N_R^0 are defined in Eqs. (3.6) and λ_L, λ_R are Lagrangian multipliers which are to be determined from the requirements of the ROH [i.e., (i) and (iii)].

We next introduce the temperature Green's function in the imaginary-time Heisenberg representation[22]

$$\begin{aligned} \mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') &\equiv -\text{Tr}\{\rho T_\tau[\psi(\vec{r}, \tau)\psi^\dagger(\vec{r}', \tau')]\} \\ &\equiv -\langle T_\tau[\psi(\vec{r}, \tau)\psi^\dagger(\vec{r}', \tau')]\rangle, \end{aligned} \tag{3.15a}$$

in which ψ and $\psi^\dagger$ are fermion field operators and T_τ is the τ ordering operator. The "temperature current" is defined by

$$\begin{aligned} \vec{J}(\vec{r}, \tau) &\equiv -\frac{(2s+1)e\hbar i}{2m(x)} (\vec{\nabla}_r - \vec{\nabla}_{r'}) \mathcal{G}(\vec{r}, \vec{r}', \tau, \tau')\Big|_{\substack{\vec{r}=\vec{r}' \\ \tau'=\tau^+}} \\ &= \text{Tr}[\rho \vec{J}_{\text{op}}(\vec{r}, \tau)]/\mathfrak{z}, \end{aligned} \tag{3.16a}$$

where s is the spin and the current operator is given by

$$\vec{j}_{op}(\vec{r},\tau)\equiv\frac{e\hbar(2s+1)}{2im(x)}[\psi^{\dagger}(\vec{r},\tau)\vec{\nabla}\psi(\vec{r},\tau) - \psi^{\dagger}(\vec{r},\tau)\overleftarrow{\nabla}\psi(\vec{r},\tau)] . \quad (3.16b)$$

The arrows on the gradient indicate the operand of differentiation and $m(x)$ is the effective mass at the point where the current is evaluated [i.e., $m(x)$ is not to be construed as a continuous differentiable function of x].

From the ROH, a tunnel junction which is biased is a steady-state system (i.e., the first requirement). The current, therefore must be independent of the time when it is measured, so that, from Eq. (3.16a), we require

$$\mathcal{G}(\vec{r},\vec{r}',\tau,\tau')=\mathcal{G}(\vec{r},\vec{r}',\tau-\tau') . \quad (3.15b)$$

In the Appendix, we prove that $\mathcal{G}$ is translationally invariant in τ if $\Delta\lambda\equiv\lambda_R-\lambda_L=2m\pi/-i\beta$ (m is an integer).

From Eq. (3.6a), we can write

$$\lambda_L N_L^0+\lambda_R N_R^0=\lambda_L N+\Delta\lambda N_R^0=\lambda_R N-\Delta\lambda N_L^0 , \quad (3.17a)$$

$$\Delta\lambda\equiv\lambda_R-\lambda_L=\nu_m , \quad (3.17b)$$

in which ν_m is a boson Matubara index,[22] $\nu_m\equiv 2m\pi/-i\beta$, and m is any integer. The temperature current can be written

$$\vec{J}(\vec{r},\tau)=(1/\mathcal{Z})\mathrm{Tr}\{[e^{\beta\Delta\lambda N_R^0}]e^{-\beta(\mathcal{H}-\lambda_L N)}\vec{j}_{op}(\vec{r},0)\} = (1/\mathcal{Z})\mathrm{Tr}\{[e^{-\beta\Delta\lambda N_L^0}]e^{-\beta(\mathcal{H}-\lambda_R N)}\vec{j}_{op}(\vec{r},0)\}. \quad (3.17c)$$

We have used the τ translational invariance of $\mathcal{G}$ to write Eq. (3.17c). The form of $\vec{J}$ above is analogous to the equilibrium average of the current operator[22] except for the exponentional factors in square brackets. The boson indices in these factors appear also in the linear-response analysis of the TH theory.[19] In that theory, the difference in chemical potentials eV is treated as a boson index ν_m until all Matsubara sums have been performed.[19] At the end of this calculation the boson index ν_m is analytically continued to the physical voltage, i.e., $\nu_m\to eV+i\delta$, $\delta\to 0^+$.

Because $e^{\beta\nu_m}=1$, it is not necessary to know the exact form of N_L^0 and N_R^0 in Eqs. (3.17) in order to calculate $\vec{J}$. It is sufficient to know only that N_L^0 and N_R^0 are the operators which measure the occupation numbers of left- and right-incident EBF in some representation of many-body states of the type defined in Eqs. (3.5). Thus, $\vec{J}$ is independent of the particular representation in which the number operators are diagonal.

It is evident from Eq. (3.17) that we have lost information about the nonequilibrium statistics by making $\Delta\lambda=\nu_m$. By using the ROH model, however, we assume that the statistics of current flow are associated with the directions of current flow (i.e., the third condition). Consequently, these statistics are determined by associating the current which flows from left to right (right to left) with μ_L (μ_R), the chemical potential of the left (right) electrode.

The fact that $\lambda_L=\lambda_R+\nu_m$, which is equivalent to time translational invariance, results in imaginary-time boundary conditions (from the Appendix, theorem 2), which are identical to the usual equilibrium imaginary-time boundary conditions of Kadanoff and Baym.[36] Thus, we can use all of the temperature-Green's-function diagram techniques[22] for equilibrium systems when we evaluate $\mathcal{G}$. We then calculate $\vec{J}$, and get $\vec{j}$, the physical current, by letting $\lambda_L\to\mu_L$, $\lambda_R\to\mu_R$ at the end of the calculation. (In other words, $\lambda_R-\lambda_L=\nu_m\to\mu_R-\mu_L=-eV$. According to the conventions of Fig. 1, $\mu_L=\zeta_L$ and $\mu_R=\zeta_L-eV$.)

Let us next examine the τ-representation continuity equation for the current:

$$e(2s+1)\frac{i}{\hbar}\frac{\partial}{\partial\tau}n(\vec{r},\tau)+\vec{\nabla}\cdot\vec{j}_{op}(\vec{r},\tau)=0 , \quad (3.18a)$$

where

$$n(\vec{r},\tau)\equiv\psi^{\dagger}(\vec{r},\tau)\psi(\vec{r},\tau) , \quad (3.18b)$$

$$\langle n(\vec{r},\tau)\rangle\equiv\mathcal{G}(\vec{r},\vec{r}',\tau-\tau')\big|_{\substack{\vec{r}'=\vec{r}\\ \tau'=\tau+}} . \quad (3.18c)$$

Thus, from Eqs. (3.16) and (3.18), we have

$$\vec{\nabla}\cdot\vec{J}(\vec{r},\tau)=0, \quad (3.19a)$$

or the x component of the temperature current is position independent in a tunnel junction which is translationally invariant in the plane parallel to the junction. Equation (3.19a) is merely a mathematical consequences of the fact that, in steady state, the density is independent of the time at which it is evaluated. The current and density operators $\vec{j}_{op}(\vec{r},t)$ and $n(\vec{r},t)$, respectively, defined in the real-time Heisenberg representation of Eq. (3.1b), obey the continuity equation

$$e(2s+1)\frac{\partial}{\partial t}n(\vec{r},t)+\vec{\nabla}\cdot\vec{j}_{op}(\vec{r},t)=0. \quad (3.19b)$$

The ROH states that a tunnel junction is a steady-state system, even if it is biased. We, thus, expect that the thermal average of the density, $\langle n(\vec{r},t)\rangle$, which is measured at $(\vec{r},t)$, is independent of time. Therefore, the measured current $\vec{j}(\vec{r},t)=\langle\vec{j}_{op}(\vec{r},t)\rangle$, also must obey Eq. (3.19a). Consequently, any analytic continuation of the Lagrangian multipliers which relates the "temperature" current in Eq. (3.17) to the "physical" current in a biased junction must preserve the vanishing of the divergence of the current.

In order to relate the temperature current to the physical current $\vec{j}$, let us write

$$\vec{J}(\vec{r},\tau)=\mathrm{Tr}[e^{\beta\Omega_L N_L^0+\lambda_R N_R^0)}e^{-\beta\mathcal{H}}\vec{j}_{op}(\vec{r},\tau)/\mathcal{Z}] , \quad (3.20a)$$

$$\lambda_R = \lambda_L + \nu_m \ . \tag{3.20b}$$

From the time translationally invariant form of $\mathcal{G}$ (derived as theorem 1 in the Appendix) and Eq. (3.16a), it is clear that $\vec{J}(\vec{r}, \tau) = \vec{J}(\vec{r}, 0)$. Therefore, the prescription for relating $\vec{J}$ and $\vec{j}$ is stated as follows: (i) calculate $\mathcal{G}$ with $\lambda_R = \mu_L + \nu_m$, (ii) evaluate $\vec{J}$ with this $\mathcal{G}$, and (iii) let $\lambda_L \to \mu_L$ (for states carrying current to the right, which are associated with N_L^0) and $\lambda_R \to \mu_R$ (for states carrying current to the left, associated with N_R^0) at the end of the calculation, so that Eq. (3.19a) applies after analytically continuing λ_L and λ_R. From the previous discussion, λ_L and λ_R are defined to within a boson index. For convenience, we set $\lambda_L = \mu_L$ and analytically continue the variable ν_m, i.e., $\nu_m \to 0$ $(-eV)$ for states carrying current to the right (left).

We next derive the perturbation-theory expression for the temperature current. From Eqs. (3.15a) and (3.14), we can write the temperature Green's function in the form[22]

$$\mathcal{G}(\vec{r}, \vec{r}', \tau - \tau') = -\langle T_\tau[S(\beta)\hat{\psi}(\vec{r}, \tau)\hat{\psi}^\dagger(\vec{r}', \tau')]\rangle_0 \ , \tag{3.21}$$

where the subscript 0 indicates an average over ρ_0, the unperturbed density matrix. Before examining the equation of motion of $\mathcal{G}$, let us consider $\mathcal{G}_0$, the unperturbed temperature Green's function (i.e., $\mathcal{H}_{int} = 0$), in the α representation of Eqs. (3.5) (eigenstates of $\mathcal{H}_{00}$, N_L^0 and N_R^0),

$$\mathcal{G}_0(\vec{r}, \vec{r}', \tau - \tau') \equiv -\mathrm{Tr}\{\rho_0 T_\tau[\hat{\psi}(\vec{r}, \tau)\hat{\psi}^\dagger(\vec{r}', \tau')]\} \ . \tag{3.22}$$

In the α representation,[22]

$$\hat{\psi}(\vec{r}, \tau) \equiv \sum_{E_i, \vec{k}_\parallel} \hat{C}_i(E, \vec{k}_\parallel, \tau)\chi_i(E, \vec{k}_\parallel, \vec{r}) \ , \tag{3.23}$$

where $\{\chi_i\}$ $(i = L, R)$ is the set of extended orthonormal one-electron basis functions from which the α representation is constructed. Thus, from theorems 1 and 2 in the Appendix, $\mathcal{G}_0$ can be written in the form[36]

$$\mathcal{G}_0(\vec{r}, \vec{r}', \tau - \tau') = \frac{1}{\beta}\sum_{z_n} e^{-z_n(\tau - \tau')} G_0(\vec{r}, \vec{r}', z_n) \ , \tag{3.24a}$$

$$z_n \equiv \pi(2n+1)/(-i\beta) + \lambda \ . \tag{3.24b}$$

We define the quantity λ by $\lambda \equiv \mu_L + \nu_m$, so that z_n is defined to within a boson index ν_m. The quantity G_0 is defined by

$$G_0(\vec{r}, \vec{r}', z_n) \equiv \int_0^\beta d\tau \mathcal{G}_0(\vec{r}, \vec{r}', \tau) e^{z_n \tau}$$

$$= \sum_{E_i, \vec{k}_\parallel} \frac{\chi_i(E, \vec{k}_\parallel, \vec{r})\, \chi_i^\dagger(E, \vec{k}_\parallel, \vec{r}')}{z_n - E_i} \ , \tag{3.25}$$

which is the expression for the one-electron Green's function in Eq. (2.10).

The unperturbed case provides a good example of use of the prescription which follows Eqs. (3.20). We shall, therefore, trace through the algebra. Using Eq. (3.24a), we perform the Matsubara sum in

$$\mathcal{G}_0(\vec{r}, \vec{r}', \tau) = \frac{1}{\beta}\sum_{z_n} e^{-z_n \tau} \sum_{E_i, \vec{k}_\parallel} \frac{\chi_i(E, \vec{k}_\parallel, \vec{r})\, \chi_i^\dagger(E, \vec{k}_\parallel, \vec{r}')}{z_n - E_i} \ , \tag{3.26}$$

by converting it into a contour integral as shown in Fig. 2,[19] where C is the (counterclockwise) contour of integration:

$$\mathcal{G}_0(\vec{r}, \vec{r}', \tau) = \frac{1}{2\pi i}\oint_C dz\, e^{-z\tau} f(z) \times \sum_{E_i, \vec{k}_\parallel} \frac{\chi_i(E, \vec{k}_\parallel, \vec{r})\, \chi_i^\dagger(E, \vec{k}_\parallel, \vec{r}')}{z - E_i}$$

$$= \sum_{E_i, \vec{k}_\parallel} f(E_i) e^{-E_i \tau} \chi_i(E, \vec{k}_\parallel, \vec{r})\, \chi_i^\dagger(E, \vec{k}_\parallel, \vec{r}') \ , \tag{3.27a}$$

$$f(E_i) \equiv (1 + e^{\beta(E_i - \lambda)})^{-1} \ , \tag{3.27b}$$

and $\lambda = \mu_L + \nu_m$ [i.e., ν_m is a boson index and Eq. (3.27b) is written to indicate that f is defined to within a boson index ν_m]. The temperature current, evaluated using Eq. (3.27a) is given by

$$\vec{J}_0 = -\frac{(2s+1)e\hbar i}{2m(x)} \sum_{E_i, \vec{k}_\parallel} f(E_i)(\vec{\nabla}_r - \vec{\nabla}_{r'})\psi_i(E, \vec{k}_\parallel, x) \times \psi_i^*(E, \vec{k}_\parallel, x')\, e^{i\vec{k}_\parallel \cdot (\vec{\rho} - \vec{\rho}')}\big|_{\vec{r} = \vec{r}'} \ , \tag{3.28}$$

and we have used Eq. (2.13b), for $E_i \geq 0$. The current in the transverse plane vanishes because of translational invariance in that plane ($\vec{k}_\parallel$ and $-\vec{k}_\parallel$ contributions cancel). Thus, normal to the junction

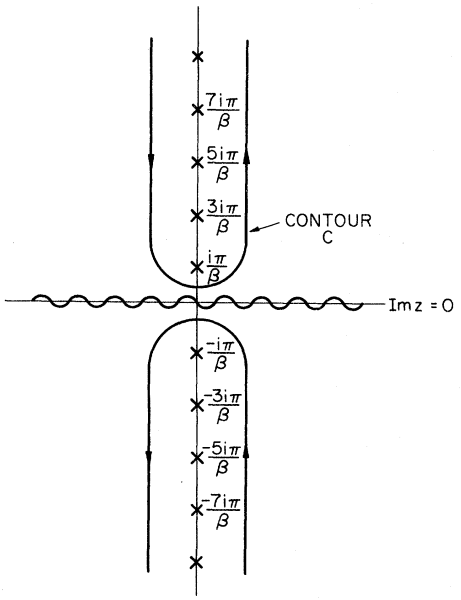

FIG. 2. Counterclockwise contour used in converting the Matsubara sum in Eq. (3.26) into the contour integral in Eq. (3.27a).

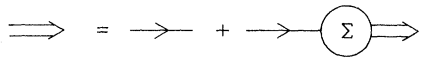

FIG. 3. Schematic Dyson's equation with generalized self-energy.

$$J_0 = -\tfrac{1}{2}[(2s+1)e\hbar i]\sum_{E_i,\vec{k}_\parallel} f(E_i)\, W_{ii}^*(E,\vec{k}_\parallel)/m_j \ . \tag{3.29}$$

In writing Eq. (3.29), we have used the fact that W_{ii}^*/m_j is a constant throughout the junction, so that, for the extended-basis functions of Eqs. (2.2)–(2.5),

$$\frac{W_{LL}^*(E,\vec{k}_\parallel)}{m_i} = \frac{i}{\pi\hbar^2}\,|\,S_{12}(E,\vec{k}_\parallel)\,|^2 = -\frac{W_{RR}^*(E,\vec{k}_\parallel)}{m_i} \ . \tag{3.30}$$

The physical current $\vec{j}_0$ is derived from $\vec{J}_0$ by letting $\nu_s \to 0$ for current flowing to the right and $\nu_s \to -eV$ for current flowing to the left, so that

$$j_0 = \frac{2s+1}{h}\, e\int_0^\infty dE[f_L(E) - f_L(E+eV)] \times \int \frac{d^2\vec{k}_\parallel}{(2\pi)^2}\,|\,S_{12}(E,\vec{k}_\parallel)\,|^2 \ . \tag{3.31a}$$

The quantity $|S_{12}|^2$ is the barrier transmission probability[19] and

$$f_L(E) \equiv (e^{\beta(E-\mu_L)}+1)^{-1}. \tag{3.31b}$$

The current given in Eq. (3.31a) is precisely that predicted by the conventional one-electron tunneling theory.[19]

In order to derive the Dyson's equation for the Green's function, we examine the equations of motion of the temperature Green's function[22] $\mathcal{G}$

$$\left(\frac{\partial}{\partial\tau} + h_0(\vec{r})\right)\mathcal{G}(\vec{r},\vec{r}',\tau-\tau') = -\delta(\tau-\tau')\delta(\vec{r}-\vec{r}') - \langle T_\tau\{[\mathcal{H}_{\rm int}(\tau),\psi(\vec{r},\tau)]\psi^\dagger(\vec{r}',\tau')\}\rangle \ . \tag{3.32}$$

The commutator on the right-hand side of Eq. (3.32) depends on the specific form of the interaction.

The elastic electron-impurity interaction gives, from Eqs. (3.4),

$$[\mathcal{H}_{el}(\tau),\psi(\vec{r},\tau)] = -V_I(\vec{r}-\vec{R})\psi(\vec{r},\tau) \ . \tag{3.33}$$

The static δ-function impurity is treated in Sec. IV.

Writing Eq. (3.32) as an integral equation in the interaction picture, we have Dyson's equation in $(\vec{r},\tau)$ space,[22] which is represented schematically in Fig. 3. The Σ in the diagram is a generalized self-energy whose form depends on the particular interaction. Consequently, Dyson's equation corresponding to Eq. (3.32) is

$$\mathcal{G}(\vec{r},\vec{r}',\tau-\tau') = \mathcal{G}_0(\vec{r},\vec{r}',\tau-\tau') + \int d^3r_2\int_0^\beta d\tau_2\int d^3r_1\int_0^\beta d\tau_1\,\mathcal{G}_0(\vec{r},\vec{r}',\tau-\tau_1) \times\Sigma(\vec{r}_1,\vec{r}_2,\tau_1-\tau_2)\mathcal{G}(\vec{r}_2,\vec{r}',\tau_2-\tau'). \tag{3.34}$$

The imaginary-time boundary conditions of theorem 2 of the Appendix allow us to Fourier transform Eq. (3.34)[36] to give the "energy" representation of Dyson's equation:

$$G(\vec{r},\vec{r}',z_n) = G_0(\vec{r},\vec{r}',z_n) + \int d^3r_2\int d^3r_1\, G_0(\vec{r},\vec{r}_1,z_n) \times\Sigma(\vec{r}_1,\vec{r}_2,z_n)\,G(\vec{r}_2,\vec{r}',z_n) \ , \tag{3.35a}$$

$$G(z_n) \equiv \int_0^\beta d\tau\, e^{z_n\tau}\mathcal{G}(\tau) \ , \tag{3.35b}$$

$$\mathcal{G}(\tau) = \frac{1}{\beta}\sum_{z_n} e^{-z_n\tau}G(z_n) \ . \tag{3.35c}$$

The self-energy Σ involves intermediate-state sums over boson indices in the case of the electron-phonon interaction.[22]

Since λ_L and λ_R differ by a boson index, according to Eq. (3.16), the perturbation expression in Eq. (3.34) is precisely that obtained using the equilibrium temperature-Green's-function method.[22,36] In order to extend our analysis to the nonequilibrium system of a tunnel junction across which a bias is imposed, we must specify a prescription which allows us both to analytically continue the Lagrangian multipliers in Eq. (3.17) and to impose the correct spatial boundary conditions for current flow.

The prescription consists of the following set of instructions: (i) The cuts of the functions $\mathcal{G}_0$, Σ, and $\mathcal{G}$ appearing in Eq. (3.35a) occur in the same order as the $\vec{r}$ arguments: That is, as we proceed from left to right (from $\vec{r}$ to $\vec{r}\,'$) in Eq. (3.35a), the cuts of the associated functions move up in the upper-half of the complex z plane, as illustrated in Fig. 4. Using this ordering, Eq. (3.35a) becomes

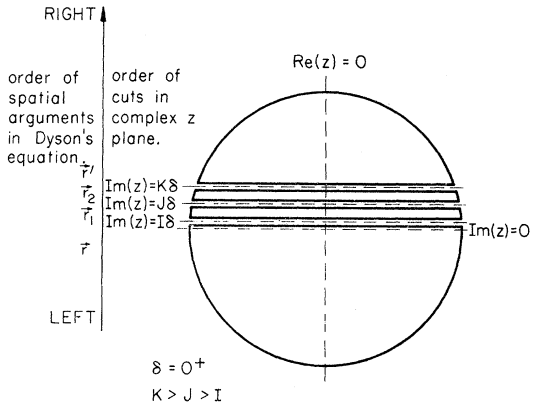

FIG. 4. Schematic illustration of the relation between the order of the cuts in the complex z plane and the spatial arguments in Dyson's equation. This cut ordering is used in order to preserve the correct spatial boundary conditions for the current in the prescription following Eqs. (3.35).

$$G(\vec{r}, \vec{r}', z_n - iK\delta) = G_0(\vec{r}, \vec{r}', z_n - iK\delta) + \int d^3r_1 \int d^3r_2 G_0(\vec{r}, \vec{r}_1, z_n - iI\delta) \times \Sigma(\vec{r}_1, \vec{r}_2, z_n - IJ\delta) G(\vec{r}_2, \vec{r}', z_n - iK\delta) , \quad (3.36a)$$

$I < J < K$, and $\delta = 0^+$. The cuts of the functions within Σ also are ordered according to this prescription. We show, in Sec. IV, that this ordering satisfies the spatial boundary conditions for electron-static-impurity scattering.

(ii) Perform the Matsubara sum as in Eq. (3.27a) with the above prescription.

(iii) Calculate the temperature current in Eq. (3.16a) before analytically continuing the Lagrangian multipliers λ_L and λ_R (i.e., before letting $\nu_m \to -eV$):

$$\vec{J} = -\frac{i\hbar e}{2} \frac{2s+1}{m(x)} (\vec{\nabla}_r - \vec{\nabla}_{r'}) \mathcal{G}(\vec{r}, \vec{r}', \tau - \tau')\big|_{\substack{\tau' = \tau^+ \\ \vec{r}' = \vec{r}}} . \quad (3.36b)$$

The position independence of the temperature current is required by Eq. (3.19a). This position independence must be preserved in the analytic continuation.

(iv) Analytically continue λ_L and λ_R by taking $\nu_m \to 0$ for states carrying current to the right, i.e., terms involving

$$\sum_{E_L, \vec{k}_\parallel} \psi_L \psi_L^*$$

from the discontinuities in G_0 or Σ across the respective cuts, and $\nu_m \to -eV$ for terms involving

$$\sum_{E_R, \vec{k}_\parallel} \psi_R \psi_R^*$$

in the cut discontinuities.

Instructions (i)–(iv) are applied to the calculation of the current in a junction with static impurities in Sec. IV. The results [i.e., Eq. (4.10a)] are shown to agree with those of the one-electron tunneling theory.[19]

A further justification of the ROH method results from studying the relation between the temperature and real-time Green's functions for a system of noninteracting quasiparticles.[32] In other words, we can examine the consequences of the assumption that the many-body Hamiltonian of a tunnel junction (including interactions) of Eq. (3.1a) can be diagonalized in the form of the Hamiltonian of a gas of current-carrying noninteracting quasiparticles whose energies have been renormalized by the interaction. In this model, we can construct the exact many-particle states which correspond to the single-particle EBF, without recourse to perturbation theory. The Lehmann representations[22] of the exact nonequilibrium real-time and temperature Green's functions calculated with these many-particle states can be related with the preceding perturbation-theory prescription.[32] This prescription permits us to decompose the Lehmann representation of the *exact* temperature Green's function into a sum of two terms associated with left- and right-current-carrying states in analogy with our decomposition of the free-particle Green's function in Eqs. (3.27)–(3.31). In all other theories of tunneling,[8,19] except those of Davis,[7,9] this association is made on the basis of defining an "unperturbed" Hamiltonian describing two noninteracting electrodes. In such theories, the tunneling current is calculated by treating the transfer of electrons between electrodes by some variant of linear-response theory. Therefore, the analytic continuation procedure presented above is the central new feature of our theory which permits the association of separate "electrode" chemical potentials with the components of current flowing to the left and right in a biased tunnel junction without the need for treating the interaction between these components via linear-response theory.

IV. APPLICATION TO STATIC IMPURITIES: δ-POTENTIAL MODEL

An examination of the consequences of the electron-static-impurity interaction is interesting for several reasons. First, the theory presented in Sec. III can be applied to analyze the tunneling characteristics due to this interaction, and the results obtained can be compared to those of the conventional versions of one electron-tunneling theory[19]; second, the mechanism of resonant elastic tunneling[15,18,37] can be described within the framework of the static δ-function-impurity model. Finally, our model of resonant elastic tunneling can describe zero-bias anomalies in the conductance which are observed in many types of junctions.[10]

A. Calculation of the Current

Let us examine the model in which we have N static impurities per unit area whose potentials are $V_I(\vec{r} - \vec{R}_n)$, where $\vec{R}_n$ is the equilibrium position of the ion at the site labeled by n. For simplicity, we choose the model potential

$$V_I(\vec{r} - \vec{R}_n) = V_0 \delta(x - d) \delta(\vec{\rho} - \vec{\rho}_n) , \quad (4.1)$$

which describes the localized electronic potential of the ion at $\vec{R}_n = \hat{i}d + \vec{\rho}_n$ ($\hat{i}$ is the unit vector in the x direction and $\vec{\rho}_n$ is the component of the impurity's position in the plane of the junction). Furthermore, let the ions be periodically arranged in a plane parallel to the junction located at $x = d$.

Substituting Eq. (4.1) into Eqs. (3.32) and (3.33) and employing (3.35) yields the following equation:

$$G(\vec{r}, \vec{r}', z_m) = G_0(\vec{r}, \vec{r}', z_m) + \sum_n V_0 G_0(\vec{r}, \vec{R}_n, z_m) \times G(\vec{R}_n, \vec{r}', z_m) , \quad (4.2a)$$

$$z_m \equiv (2m+1)\pi/(-i\beta) + \lambda , \tag{4.2b}$$

from Eq. (3.24b), and $\lambda = \mu_L + \nu_s$, as in Eq. (3.24b), so that z_m is defined to within a boson index.

Since we assume $\vec{k}_\parallel$ conservation throughout the junction, we can Fourier analyze G, in accordance with Eqs. (2.13), and neglect umklapp scattering from the planar array of impurities to give

$$G(x, x', \vec{k}_\parallel, z_m) = G_0(x, x', \vec{k}_\parallel, z_m) + U_0 G_0(x, d, \vec{k}_\parallel, z_m) G(d, x', \vec{k}_\parallel, z_m), \tag{4.3a}$$

$$U_0 \equiv N V_0 . \tag{4.3b}$$

According to our prescription, we must order the cuts of the quantities appearing in Eq. (4.3a) along the imaginary z axis in increasing order from left to right as illustrated in Fig. 4. Equation (4.3a) admits the solution

$$G(x, x', \vec{k}_\parallel, z_m - iK\delta) = G_0(x, x', \vec{k}_\parallel, z_m - iK\delta) + \frac{U_0 G_0(x, d, \vec{k}_\parallel, z_m - iI\delta) G_0(d, x', \vec{k}_\parallel, z_m - iK\delta)}{1 - U_0 G_0(d, d, \vec{k}_\parallel, z_m - iJ\delta)} \tag{4.4}$$

where $I < J < K$, $\delta = 0^+$.

We find $\mathcal{G}(\tau - \tau')$ by performing the Matsubara sum in Eq. (3.35c), where the contour C is given in Fig. 2:

$$\mathcal{G}(x, x', \vec{k}_\parallel, \tau) = -\frac{1}{2\pi i}\int_{-\infty}^{\infty} dz f(z) e^{-st} \{2i\,\mathrm{Im}[G_0^R(x, x', \vec{k}_\parallel, z)] + U_0 \{2i[C_1(R, A, x, x', \vec{k}_\parallel, z) - C_1(A, R, x', x, \vec{k}_\parallel, z)] + C_2(R, A, x, x', \vec{k}_\parallel, z) - C_2(A, R, x', x, \vec{k}_\parallel, z)\}\} , \tag{4.5a}$$

$$C_1(R, A, x, x', \vec{k}_\parallel, z) \equiv \frac{\mathrm{Im}[G_0^R(x, d, \vec{k}_\parallel, z)]\, G_0^A(d, x', \vec{k}_\parallel, z)}{1 - U_0 G_0^A(d, d, \vec{k}_\parallel, z)} , \tag{4.5b}$$

$$C_2(R, A, x, x', \vec{k}_\parallel, z) \equiv \frac{G_0^R(x, d, \vec{k}_\parallel, z) G_0^A(d, x', \vec{k}_\parallel, z)}{1 - U_0 G_0^R(d, d, \vec{k}_\parallel, z)} . \tag{4.5c}$$

We have used the fact that $G_0^{R,A}$ is symmetric under interchange of its spatial arguments in writing Eqs. (4.5) in the compact notation above. The quantity $f(z)$ is the distribution function defined in Eq. (3.27b). The discontinuity in G_0 across its cut is given by

$$\mathrm{Im}[G_0^R(x, x', \vec{k}_\parallel, z)] \equiv (1/2i)[G_0^R(x, x', \vec{k}_\parallel, z) - G_0^A(x, x', \vec{k}_\parallel, z)] = -\pi \sum_{E_i} \psi_i^0(E, \vec{k}_\parallel, x) \psi_i^{*0}(E, \vec{k}_\parallel, x') \delta(z - E_i) , \tag{4.6}$$

in which $i = (L, R)$. We have used Eq. (2.13) and ψ^0 is the unperturbed wave function for the junction indicated in Fig. 1. Inserting Eq. (4.6) into (4.5) and performing the integral over z prior to the sum over E_i gives our final expression for $\mathcal{G}$:

$$\mathcal{G}(x, x', \vec{k}_\parallel, \tau) = \sum_{E_i} f(E_i) e^{-E_i \tau} [A(E_i, \vec{k}_\parallel, x, x') - U_0 F(E_i, \vec{k}_\parallel, x, x')], \quad i = (L, R) \tag{4.7a}$$

$$A(E_i, \vec{k}_\parallel, x, x') \equiv \psi_i^0(E, \vec{k}_\parallel, x) \psi_i^{*0}(E, \vec{k}_\parallel, x') , \tag{4.7b}$$

$$F(E_i, \vec{k}_\parallel, x, x') \equiv \{B_1(E_i, \vec{k}_\parallel, x, x') + B_1^*(E_i, \vec{k}_\parallel, x', x) + U_0 [B_2(E_i, \vec{k}_\parallel, x, x') + B_2^*(E_i, \vec{k}_\parallel, x', x) - B_3(E_i, \vec{k}_\parallel, x, x')]\} [|1 - U_0 G_0^R(d, d, \vec{k}_\parallel, E_i)|]^{-2} , \tag{4.7c}$$

$$B_1(E_i, \vec{k}_\parallel, x, x') \equiv A(E_i, \vec{k}_\parallel, x, d) G_0^A(d, x', \vec{k}_\parallel, E_i) , \tag{4.7d}$$

$$B_2(E_i, \vec{k}_\parallel, x, x') \equiv B_1(E_i, \vec{k}_\parallel, x, x') G_0^R(d, d, \vec{k}_\parallel, E_i) , \tag{4.7e}$$

$$B_3(E_i, \vec{k}_\parallel, x, x') \equiv B_1(E_i, \vec{k}_\parallel, d, x') G_0^R(x, d, \vec{k}_\parallel, E_i) . \tag{4.7f}$$

When we apply

$$J_{op} = -\frac{e\hbar i(2s+1)}{2m(x)} \left(\frac{d}{dx} - \frac{d}{dx'} \right) \Big|_{\substack{x'=x \\ \tau'=\tau^+}}$$

individually to the terms with $i = L$ (or $i = R$) in Eqs. (4.7), we obtain position-independent currents [the quantity $m(x)$ indicates the effective mass in the region of the junction where the current is evaluated, and is *not* a continuous differentiable function of x] . We obtain a particularly simple

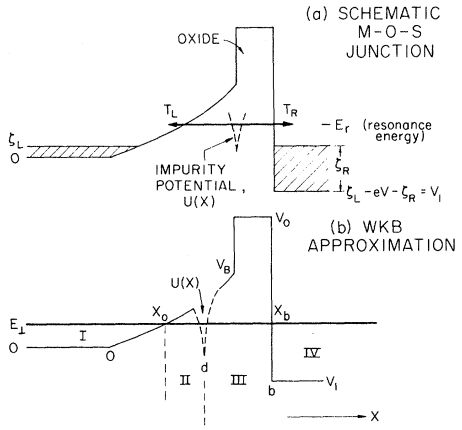

FIG. 5. (a) Schematic metal-oxide-semiconductor (M-O-S) junction with an attractive impurity potential, $U(x) = U_0\,(x-d)$, so that the impurity potential is completely localized at $x=d$. T_L (T_R) is the probability for transmission of an electron from the impurity site to the left (right) electrode. In the drawing, $T_L = T_R$, so that the impurity potential is resonant. (b) The same junction as in (a), except that the conventions of the WKB approximation in Sec. IV B are illustrated.

form if we use Eq. (2.16) to show that, for $E > 0$ (suppressing $\vec{k}_\parallel$ arguments),

$$\psi_L^0(E,d)G_0^R(x,d,E)=\psi_L^0(E,x)G_0^R(d,d,E), \quad x\to\infty$$

$$\psi_R^0(E,d)G_0^R(x,d,E)=\psi_R^0(E,x)G_0^R(d,d,E), \quad x\to-\infty \tag{4.8}$$

$$\psi_L^{*0}(E,d)G_0^A(d,x',E)=\psi_L^{*0}(E,x')G_0^A(d,d,E), \quad x'\to\infty$$

$$\psi_R^{*0}(E,d)G_0^A(d,x',E)=\psi_R^{*0}(E,x')G_0^A(d,d,E), \quad x'\to-\infty .$$

The temperature current is, from Eqs. (3.16a),

$$\vec{J}\equiv-\frac{2s+1}{2m(x)}e\hbar i(\vec{\nabla}_r-\vec{\nabla}_{r'})\mathcal{G}(\vec{r},\vec{r}',\tau-\tau')\Big|_{\substack{\vec{r}=\vec{r}'\\ \tau'=\tau^+}}$$

$$=-\frac{2s+1}{2m(x)}e\hbar i(\vec{\nabla}_r-\vec{\nabla}_{r'})$$

$$\times\sum_{\vec{k}_\parallel}\mathcal{G}(x,x',\vec{k}_\parallel,0^-)e^{i\vec{k}_\parallel\cdot(\rho-\rho')}\Big|_{\vec{r}=\vec{r}'}. \tag{4.9a}$$

Inserting Eqs. (4.7) and (4.8) into (4.9a) and performing some algebraic manipulations yields

$$J=-e\hbar i\,\frac{2s+1}{2m(x)}\left[\left(\frac{d}{dx}-\frac{d}{dx'}\right)\right.$$

$$\left.\times\sum_{E_i,\vec{k}_\parallel}\frac{f(E)\psi_i^0(E,\vec{k}_\parallel,x)\psi_i^{*0}(E,\vec{k}_\parallel,x')}{|1-U_0G_0^R(d,d,\vec{k}_\parallel,E)|^2}\right]_{x'=x}. \tag{4.9b}$$

The current due to the $L(R)$ terms is evaluated at $x\to\infty$ ($x\to-\infty$) in Eq. (4.9b). The physical current j, which flows when a bias is imposed, is derived from Eq. (4.9) by letting $\lambda\to\mu_L$ (μ_R) for the $L(R)$ terms according to our prescription [i.e., $\nu_m\to 0$ ($-eV$) for $L(R)$ terms]. Thus,

$$j=\frac{(2s+1)e}{h}\int_0^\infty dE[f_L(E)-f_L(E+eV)]$$

$$\times\int\frac{d^2k_\parallel}{(2\pi)^2}\,\frac{|S_{12}^0(E,\vec{k}_\parallel)|^2}{|1-U_0G_0^R(d,d,\vec{k}_\parallel,E)|^2}, \tag{4.10a}$$

$$f_L(E)\equiv(1+e^{\beta(E-\mu_L)})^{-1}. \tag{4.10b}$$

In deriving Eqs. (4.10), we have used the wave-function definitions in Eq. (2.2)–(2.5) and the summation convention in Eq. (2.11). The expression in Eq. (4.10a) is the one-electron current for the average barrier shown in Fig. 5. Equations (4.10) agree exactly with the predictions of the ordinary one-electron theory.[19]

B. Some Features of Resonant-Impurity-Assisted Tunneling

Real intermediate states in tunneling processes can produce zero-bias anomalies in the conductance.[10,15,16,24] It is interesting, therefore, to examine the effect of resonant-impurity potentials upon the tunneling characteristics.[37] Since the resonant-barrier-penetration probabilities for different barrier and impurity potentials are Lorentzian in form,[38] we calculate the resonant transmission coefficient for a general barrier potential in the WKB approximation, in which we describe the static impurity potential by a Dirac δ function. The WKB treatment is valid for either high or thick barriers.[21] This approximation permits us to study the general features of resonant-elastic-tunneling line shapes.

A typical barrier containing a resonant impurity is illustrated in Fig. 5. For such a barrier, the WKB barrier-penetration probability is calculated by writing the current-carrying WKB wave function $\psi_{\rm WKB}$ as in Eq. (2.2), and by applying the following matching conditions for a static δ-function potential $U_0\delta(x-d)$ in the barrier:

$$\lim_{\epsilon\to0}[\psi_{\rm WKB}(d+\epsilon)-\psi_{\rm WKB}(d-\epsilon)]=0, \tag{4.11a}$$

$$\lim_{\epsilon\to0}[\psi'_{\rm WKB}(d+\epsilon)-\psi'_{\rm WKB}(d-\epsilon)]=\frac{2mU_0}{\hbar^2}\psi_{\rm WKB}(d). \tag{4.11b}$$

The resulting WKB transmission coefficient is given by[32]

$$|S_{12}(E,\vec{k}_\parallel)|^2\equiv|S_{12}(E_\perp)|^2$$

$$=|S_{12}^0(E_\perp)|^2\,\frac{K_D^2}{(K_D-K_0)^2+\Gamma_{\rm WKB}^2}, \tag{4.12a}$$

$$K_D\equiv\{2m[V(d)-E_\perp]/\hbar^2\}^{1/2}, \tag{4.12b}$$

$$K_0\equiv-mU_0/\hbar^2, \tag{4.12c}$$

$$\Gamma_{\mathrm{WKB}}^2 \equiv \frac{K_0^2}{4}\left[\left(\frac{T_L}{T_R}\right)^{1/2}+\left(\frac{T_R}{T_L}\right)^{1/2}\right]^2 T_L T_R\,, \tag{4.12d}$$

$$|S_{12}^0|^2 = (q/k)\,T_L T_R\,. \tag{4.12e}$$

The wave numbers q and k are defined in Eqs. (2.3), and we have made the effective mass constant throughout the junction.

The quantities T_L and T_R are, respectively, the tunneling probabilities from the impurity site to left and right electrodes, as illustrated in Fig. 5. The transmission coefficient in Eq. (4.12a) is resonant when the width Γ_{WKB} is minimum, which, in turn, requires that

$$T_L \equiv e^{-2J_L} \cong T_R \equiv e^{-2J_R}\,, \tag{4.13a}$$

$$J_L \equiv \int_{x_0}^{d} dx\left(\frac{2m}{\hbar^2}\left[V(x)-E_\perp\right]\right)^{1/2}\,, \tag{4.13b}$$

$$J_R \equiv \int_{d}^{x_b} dx\left(\frac{2m}{\hbar^2}\left[V(x)-E_\perp\right]\right)^{1/2}\,, \tag{4.13c}$$

$$E_\perp \equiv E - \hbar^2 k_\parallel^2/2m\,, \tag{4.13d}$$

$$V(x_0) = V(x_b) = E_\perp\,. \tag{4.13e}$$

The quantities x_0 and x_b are the classical turning points.[21] In deriving Eqs. (4.12) and (4.13), we have made the effective mass constant throughout the junction for simplicity.

Equation (4.12a) can be separated into resonant and nonresonant terms

$$|S_{12}|^2 = |S_{12}^0|^2\left(1+\frac{K_0(2K_D-K_0)-\Gamma_{\mathrm{WKB}}^2}{(K_D-K_0)^2+\Gamma_{\mathrm{WKB}}^2}\right)\,. \tag{4.14a}$$

In order to simplify the calculation, let us note that, for a sharp resonance (i.e., $\Gamma_{\mathrm{WKB}} \ll K_0$),

$$\frac{\Gamma_{\mathrm{WKB}}}{(K_D-K_0)^2+\Gamma_{\mathrm{WKB}}^2} \simeq \pi\delta(K_D-K_0)\,. \tag{4.14b}$$

The current is evaluated by a straightforward substitution of Eqs. (4.14) into Eq. (4.10a):

$$j \cong j_0 + j_r\,. \tag{4.15a}$$

The ordinary one-electron current without resonant effects at zero temperature is given by[19]

$$j_0 \equiv \frac{em}{2\pi^2\hbar^3}\int_{\zeta_L-eV}^{\zeta_L} dE\int_{E_m}^{E} dE_\perp\,|S_{12}^0(E_\perp)|^2\,, \tag{4.15b}$$

$$E_m \equiv \min(0,\,-V_1)\,. \tag{4.15c}$$

The resonant contribution to the zero-temperature current is given in the following equation[32]:

$$j_r \equiv \frac{e}{h}\int_{\zeta_L-eV}^{\zeta_L} dE\,\frac{K_0^3}{\Gamma_{\mathrm{WKB}}}\,|S_{12}^0(E_r)|^2\Theta(E_r-E_m)\Theta(E-E_r)\,, \tag{4.15d}$$

$$E_r \equiv V(d) - \hbar^2 K_0^2/2m\,. \tag{4.15e}$$

The WKB resonance energy E_r is the bound-state energy of a δ-function well with respect to the top of the barrier.

Neglecting the voltage dependence of the term $K_0^3|S_{12}^0|^2/\Gamma_{\mathrm{WKB}}$ in Eq. (4.15d) (which is slowly varying compared with the step function terms), we can derive the resonant conductance

$$G_r(eV) \equiv \frac{dj_r}{d(eV)} \cong G_r^0\left[(R+1)\Theta(\zeta_L-eV-E_r) - R\Theta(\zeta_L-E_r)\right]\,, \tag{4.16a}$$

$$G_r^0 \equiv \frac{e\,K_0^3}{h\Gamma_{\mathrm{WKB}}}\,|S_{12}^0(E_r)|^2\,, \tag{4.16b}$$

$$R \equiv \frac{dE_r}{d(eV)} = \frac{dV(d)}{d(eV)}\,. \tag{4.16c}$$

From the conventions in Fig. 5, $0 \geq R \geq -1$.

In Eq. (4.16a), we have identified the right (left) electrode with a metal (semiconductor), so that $E_m = 0$ in Eqs. (4.15b)–(4.15d).

Equations (4.16) are true in the WKB approximation. In a more general treatment of resonant impurities in junction barriers (i.e., the square barrier,[32] for instance), the barrier-penetration probability is Lorentzian in form[38] and the expression for the resonant conductance in the case of a sharp resonance [i.e., Eq. (4.14b)] is identical in form to Eq. (4.16a). In the general case, however, Eq. (4.16c) is only approximately satisfied. This approximation is adequate for $|S_{12}^0|^2 \ll 1$, however, so that we confine our discussion to junctions which have either high or thick barriers.

If the average barrier potential in an experiment is given, we can deduce the resonant impurity position from the value of R which fits the experimental results. In most cases, however, this potential is not known. We treat R, therefore, as an independent parameter which gives the variation with voltage of the average barrier potential at the position of the impurity. Consequently, we write

$$E_r = E_0 + ReV,\quad 0 \geq R \geq -1\,, \tag{4.17}$$

where E_0 is the position of the resonance energy at zero bias with respect to the bottom of the left-electrode's band (i.e., the semiconductor in a metal-semiconductor contact). The total conductance is

$$G(eV) = G_0(eV) + G_r(eV)\,, \tag{4.18a}$$

$$G_0(eV) \equiv dj_0/d(eV)\,, \tag{4.18b}$$

$$G_r(eV) = \begin{cases} (R+1)G_r^0, & eV < \min\left(\dfrac{\Delta}{R+1}\,,\ \dfrac{\Delta}{R}\right), \\ 0\,, & \dfrac{\Delta}{R+1} < eV < \dfrac{\Delta}{R}\,, \quad \Delta < 0\,, \\ -RG_r^0, & eV > \max\left(\dfrac{\Delta}{R+1}\,,\ \dfrac{\Delta}{R}\right), \\ G_r^0\,, & \dfrac{\Delta}{R} < eV < \dfrac{\Delta}{R+1}\,, \quad \Delta > 0\,, \end{cases} \tag{4.18c}$$

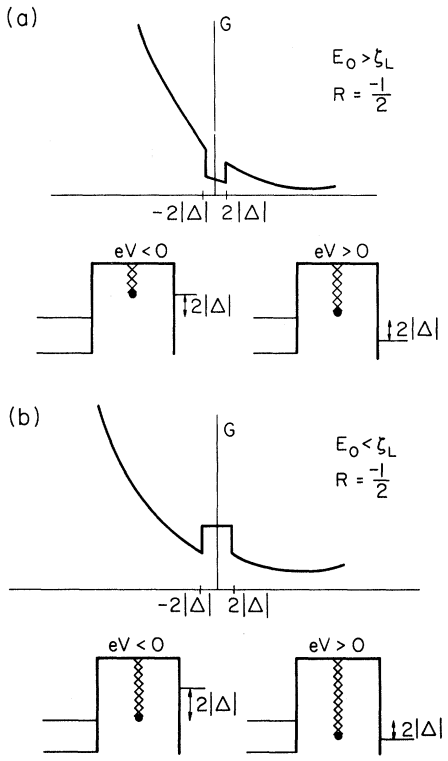

FIG. 6. Resonant-elastic conductance of Eqs. (4.18) illustrated for $R \equiv dE_r/d(eV) = -\frac{1}{2}$ (i.e., E_r is the resonance energy). E_0 is the resonance energy at zero bias (i.e., $E_r = E_0 + ReV$) and $\Delta = \zeta_L - E_0$ in the conventions of Fig. 1. In (a), $E_0 > \zeta_L$ and in (b), $E_0 < \zeta_L$. Therefore, the resonant channel opens in (a) [closes in (b)] when $|eV| > 2|\Delta|$. The schematic tunnel junctions in the figure illustrate the switch concept. The black dot represents the resonance energy and the cross hatching indicates the range of energies for which resonant tunneling can occur. This convention is also used in Fig. 7.

$$\Delta \equiv \zeta_L - E_0 \,. \tag{4.18d}$$

The thresholds in Eq. (4.18c) appear as zero-bias anomalies in the conductance. In Figs. 6 and 7, we illustrate schematically the zero-bias anomalies (ZBA) which occur in resonant elastic tunneling for $R = -\frac{1}{2}$, and $R = 0, -1$, respectively. We can understand these structures by examining Eq. (4.15d): The step function indicates that current flows in the resonant elastic channel when either $\zeta_L > E_r > \zeta_L - eV$ (forward bias, in the conventions of Fig. 1) or $\zeta_L - eV > E_r > \zeta_L$ (reverse bias). This threshold behavior is a reflection of a "switch" effect: The resonance is an extra tunneling channel through the barrier, since $|S_{12}|^2 \simeq 1$, when $E_\perp \simeq E_r$, from Eqs. (4.12) and (4.13). Therefore, tunneling electrons with total energy $E > E_r$ are always able to tunnel through the resonance. When one of the Fermi levels is above the resonance, the resonant elastic channel is opened. Thus, if the resonance energy at zero bias E_0 is above (below) the Fermi level so that $\Delta < 0$ ($\Delta > 0$), resonant-elastic tunneling (RET) does not (does) occur at zero bias. In particular, when $R = -\frac{1}{2}$, E_r moves equally in forward and reverse biases, so that the structure in the conductance near zero bias is symmetric with respect to zero bias. In this case, the conductance appears to have a minimum centered at zero bias when $\Delta < 0$, because the (RET) channel opens only when E_r passes *below* either of the electrode Fermi levels. The threshold for this process is at $|eV| = -2\Delta$. When $\Delta > 0$, the zero-bias anomaly takes the form of a conductance maximum, because the RET channel closes partially when E_r passes *above* either of the electrode Fermi levels. The threshold here is at $|eV| = 2\Delta$. These results are illustrated in Fig. 6.

In the case that $R = -1$ (i.e., the resonance energy moves with the right Fermi level), the contributions to the conductance from RET occur only when $E_r < \zeta_L$, since we have neglected the voltage dependence of the barrier penetration factor in Eq. (4.15d). Therefore, the resonant elastic channel is open when $eV < \Delta$ and closed when $eV > \Delta$ for all Δ. By an analogous argument for $R = 0$ (i.e., the resonance energy moves with the left Fermi level), RET occurs only when $E_r < \zeta_L - eV$. Thus, the RET contribution to the conductance appears for $eV < \Delta$ for Δ positive and negative. These conclusions are illustrated in Fig. 7.

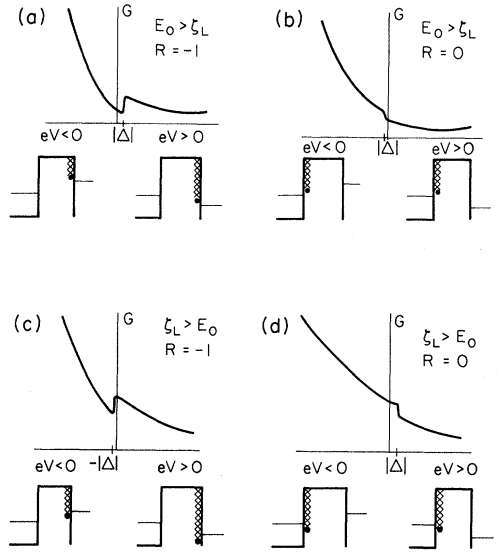

FIG. 7. Resonant-elastic conductance of Eqs. (4.18). In (a) and (c), $R = -1$ (i.e., the resonance energy moves with the right Fermi level). The resonant elastic channel opens when either Fermi level passes above the resonance energy. In (a) and (c), the resonant elastic channel is open only in forward bias, while in (b) and (d), it is open only in reverse bias.

A more realistic treatment of the voltage dependent barrier penetration factor in Eq. (4.15d) and the finite resonance width in Eq. (4.12a) changes the details of the current characteristics, but the qualitative features of our conclusions remain true.[32]

We conclude this section with a brief discussion of the experiments of Zeller and Giaever, who observe resonant (or two-step) tunneling through real (or localized) intermediate states.[15,16] Their system consists of small metal (Sn) particles embedded in the oxide of an Al-insulator-Al junction. Although our model of resonant tunneling makes no explicit provision for multiple-electron occupation of quasibound impurity states and the concomitant electron-electron Coulomb repulsion, it can be related[32] to the (semiclassical) capacitor model of Zeller and Giaever.[15,16] The key to this relation is the fact that both models involve continuous ranges of energies for which two-step tunneling is allowed as well as threshold energies for the two-step processes. In the capacitor model, the one-electron eigenvalue spectrum of a metal particle is assumed to be continuous so that tunneling electrons which "hop" on or off a particle have a continuous range of energies for which this hopping is allowed.[15,16] In addition, in order for two-step tunneling to occur, the battery must supply enough energy (i.e., the activation energy) for the charge on a metal particle to change by e. The ROH model of tunneling through resonant static δ-function-impurity potentials also exhibits a continuous spectrum of total energy E for which resonant tunneling is allowed. The $\vec{k}_{\parallel}$ integral in Eq. (4.10a), which produces the step functions in Eq. (4.15d), is a phase-space integral which results in a "band" of eigenvalues E (i.e., there is a continuous spectrum of $E > E_r$ for which resonant tunneling occurs, so that E_r is a threshold). Averaging over the activation energies in both models produces identical conductance line shapes.[32]

C. Comparison with Transfer-Hamiltonian Theory

In order to derive the transfer-Hamiltonian (TH) expression[8] for the current with our method, we apply the method of "left" and "right" effective Hamiltonians introduced by Bardeen.[1] These Hamiltonians, which are denoted by subscripts l and r, respectively, are defined in the case of a constant effective mass and $k_{\parallel}$ conservation in the following equations:

$$h_i(x) \equiv -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+\frac{\hbar^2}{2m}k_{\parallel}^2+V_i(x)\,, \quad i=(l,r) \tag{4.19a}$$

$$V_l(x) \equiv \begin{cases} 0\,, & x<0 \\ V(x)\,, & 0\le x\le b \\ V(b)\,, & x\le b \end{cases} \tag{4.19b}$$

$$V_r(x) \equiv \begin{cases} V(0)\,, & x<0 \\ V(x)\,, & 0\le x\le b \\ \zeta_L - eV - \zeta_R\,, & b<x\,. \end{cases} \tag{4.19c}$$

The quantity $V(x)$ in Eqs. (4.19) is the exact barrier potential in Eq. (2.15), so that $V_l(x) = V_r(x) = V(x)$, when $0 \le x \le b$. The left, right, and original electrode potentials are illustrated in Fig. 8. According to the TH theory,[1] we must first calculate the nondegenerate eigenfunctions ψ_l and ψ_r of the approximate Hamiltonians h_l and h_r, respectively, and then use perturbation theory to expand the current-carrying eigenfunctions of the exact Hamiltonian $h(x)$ in Eq. (2.15) in terms of ψ_l and ψ_r (which correspond to the localized electrode standing wave functions[19]). The eigenfunctions of the Hamiltonians in Eqs. (4.19) are defined by

$$h_l(x)\psi_{kl}(x) = E(k)\psi_{kl}(x)\,, \tag{4.20a}$$

$$h_r(x)\psi_{qr}(x) = E(q)\psi_{qr}(x)\,. \tag{4.20b}$$

The indices k and q are the wave numbers defined in Eqs. (2.3) for total energy E and $\vec{k}_{\parallel}$.

In the TH method, we are instructed to use perturbation theory to expand the eigenstates of $h(x)$ with fixed $\vec{k}_{\parallel}$:

$$\psi_L(E,\vec{k}_{\parallel},x) \cong \psi_{kl}(x) - \sum_{q^*}\frac{\langle q_r^* \mid \Delta V_l \mid k_l\rangle}{E(q^*)-E(k)-i\delta}\psi_{q^*r}(x)\,,$$

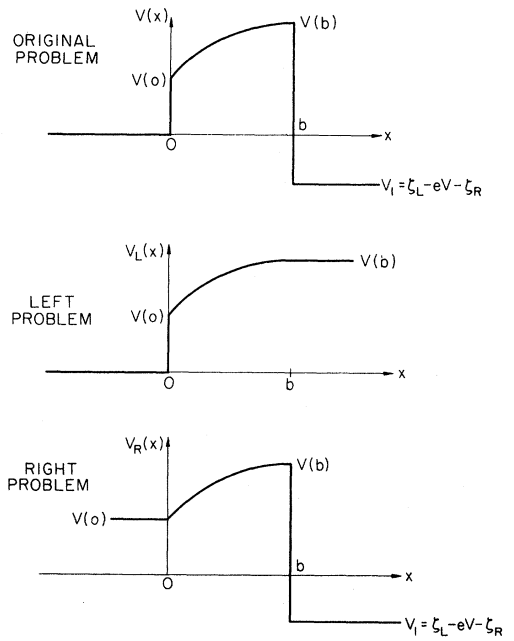

FIG. 8. Schematic illustration of the original and "left" and "right" effective potentials used in the derivation of the transfer-Hamiltonian current in Sec. IV C. The potentials are defined so that $V(x) = V_l(x) = V_r(x)$ for $0 \le x \le b$.

$$E(k) \equiv E \qquad (4.21a)$$

$$\psi_R(E, \vec{k}_\parallel, x) \cong \psi_{qr}(x) - \sum_{k^*} \frac{\langle k_l^* | \Delta V_r | q_r \rangle}{E(k^*) - E(q) - i\delta} \psi_{k^* l}(x) ,$$

$$E(q) \equiv E \qquad (4.21b)$$

in which ψ_{kl} and ψ_{qr} correspond to the same $\vec{k}_\parallel$ as ψ_L and ψ_R, $\Delta V_l \equiv h(x) - h_l(x)$, and $\Delta V_r \equiv h(x) - h_r(x)$. We use the method of Bardeen to rewrite the matrix elements in terms of currents:

$$\langle q_r | \Delta V_l | k_l \rangle \equiv \int_{-\infty}^{\infty} dx\, \psi_{qr}^*(h - h_l)\psi_{kl}$$

$$= \int_b^{\infty} dx\, \psi_{qr}^* \psi_{kl} [E(q) - E(k)] + \frac{\hbar^2}{2m} J_{qk}(b) , \qquad (4.22a)$$

$$J_{qk}(x) \equiv \psi_{qr}^*(x) \frac{d}{dx} \psi_{kl}(x) - \psi_{kl}(x) \frac{d}{dx} \psi_{qr}^*(x) , \qquad (4.22b)$$

$$\langle k_l | \Delta V_r | q_r \rangle \equiv \int_{-\infty}^{\infty} dx\, \psi_{kl}^*(h - h_r)\psi_{qr}$$

$$= \int_{-\infty}^{0} dx\, \psi_{kl}^* \psi_{qr} [E(k) - E(q)] + \frac{\hbar^2}{2m} J_{qk}^*(0) . \qquad (4.22c)$$

Using the forms of the matrix elements in Eqs. (4.22), we obtain

$$\psi_L(E, \vec{k}_\parallel, x) \simeq \psi_{kl}(x)\Theta(b - x) - \frac{\hbar^2}{2m} \sum_{q^*} \frac{J_{q^* k}(b)\, \psi_{q^* r}^*(x)}{E(q^*) - E(k) - i\delta} , \qquad (4.23a)$$

$$\psi_R(E, \vec{k}_\parallel, x) \simeq \psi_{qr}(x)\Theta(x) - \frac{\hbar^2}{2m} \sum_{k^*} \frac{\psi_{k^* l}(x) J_{qk}^*(0)}{E(k^*) - E(q) - i\delta} , \qquad (4.23b)$$

where we have used the closure relation resulting from the fact that $\{\psi_{kl}\}$ and $\{\psi_{qr}\}$ are complete sets of nondegenerate orthonormal functions.

Substituting Eqs. (4.23) into the expression for the unperturbed one-electron Green's function in Eq. (2.13b), we can expand G_0 in terms of ψ_{kl} and ψ_{qr}. The Green's function for $0 \le x,\ x' \le b$ is given by

$$G_0(x, x', \vec{k}_\parallel, z) \cong F_l + F_r , \qquad (4.24a)$$

$$F_l \equiv -\sum_k \frac{1}{E(k) - z} \left(\psi_{kl}(x)\psi_{kl}^*(x') - \frac{\hbar^2}{2m} \psi_{kl}(x) \sum_{q^*} \frac{J_{q^* k}^*(x_0)\psi_{q^* r}^*(x')}{E(q^*) - E(k) + i\delta} - \frac{\hbar^2}{2m} \psi_{kl}^*(x') \sum_{q^*} \frac{J_{q^* k}^*(x_0)\psi_{q^* r}^*(x)}{E(q^*) - E(k) - i\delta} \right),$$

$$0 \le x_0 \le b \qquad (4.24b)$$

$$F_r \equiv -\sum_q \frac{1}{E(q) - z} \left(\psi_{qr}(x)\psi_{qr}^*(x') - \frac{\hbar^2}{2m} \psi_{qr}^*(x') \sum_{k^*} \frac{\psi_{k^* l}(x) J_{qk}^*(x_0)}{E(k^*) - E(q) - i\delta} - \frac{\hbar^2}{2m} \psi_{qr}(x) \sum_{k^*} \frac{\psi_{k^* l}(x') J_{qk}^*(x_0)}{E(k^*) - E(q) + i\delta} \right),$$

$$0 \le x_0 \le b . \qquad (4.24c)$$

In deriving Eq. (4.24), we have replaced the sum over E_L (E_R) states in Eq. (2.13b) by the equivalent sum over k (q) and denoted the total energy for fixed $\vec{k}_\parallel$ by $E(k)$ [$E(q)$], since the ranges of energies of ψ_L and ψ_{kl} (ψ_R and ψ_{qr}) are identical. We have dropped terms of order J^2 and used the fact that $J_{qk}(x_0)$ is position independent for $0 \le x_0 \le b$ [i. e., since $h_l(x_0) = h_r(x_0) = h(x_0)$ for $0 \le x_0 \le b$, ψ_{kl} and ψ_{qr} are linearly independent in the same region]. From the definition of J_{qk} in Eq. (4.22b), we can write

$$G_0(x, x', \vec{k}_\parallel, z) \equiv -\sum_k \frac{1}{E(k) - z} \left[\psi_{kl}(x)\psi_{kl}^*(x') - \frac{\hbar^2}{2m} \psi_{kl}(x) \left(\psi_{kl}^*(x_0) \frac{\partial}{\partial x} G_r(x = x_0, x', \vec{k}_\parallel, E(k) - i\delta) \right.\right.$$

$$\left. - \frac{\partial}{\partial x} \psi_{kl}^*(x_0)\, G_r(x_0, x', \vec{k}_\parallel, E(k) - i\delta) \right) + \frac{\hbar^2}{2m} \psi_{kl}^*(x') \left(\frac{\partial}{\partial x} \psi_{kl}(x_0)\, G_r(x, x_0, \vec{k}_\parallel, E(k) + i\delta) \right.$$

$$\left.\left. - \psi_{kl}(x_0) \frac{\partial}{\partial x'} G_r(x, x' = x_0, \vec{k}_\parallel, E(k) + i\delta) \right) \right] - \sum_q \frac{1}{E(q) - z} [l \to r,\ k \to q], \quad 0 \le x_0 \le b \qquad (4.25a)$$

where we denote the q, r terms schematically, and we define the one-electron Green's functions for the left and right problem as follows:

$$G_l(x, x', \vec{k}_\parallel, E) \equiv -\sum_k \frac{\psi_{kl}(x)\psi_{kl}^*(x')}{E(k) - E} , \qquad G_r(x, x', \vec{k}_\parallel, E) \equiv -\sum_q \frac{\psi_{qr}(x)\psi_{qr}^*(x')}{E(q) - E} . \qquad (4.25b)$$

The retarded and advanced Green's functions are given by $G_i^R(E) \equiv G_i(E + i\delta)$ and $G_i^A \equiv G_i(E - i\delta)$, $\delta = 0^+$, respectively, for $i = l, r$.

According to our method, we calculate the physical (or measured) current via the following steps: (i) Derive the temperature Green's function by performing the Matsubara sum in Eq. (3.35c). (ii) Calculate

the temperature current from Eq. (4.9a). (iii) Make the analytic continuation $\nu_s \to 0$ $(-eV)$ in the terms involving sums over k (q) as stated in the prescription directly following Eqs. (3.35). Denoting the analytic continuation by $\lambda_L \to \mu_L$ $(\lambda_R \to \mu_R)$, the physical current calculated at $x = x_0$ is then

$$j \equiv \frac{(2s+1)e\hbar i}{2m\pi}\,(\vec{\nabla}_r - \vec{\nabla}_{r'}) \sum_{\vec{k}_\parallel} e^{i\vec{k}_\parallel\cdot(\vec{\rho}-\vec{\rho}')} \int_{-\infty}^{\infty} dz\, f(z)\,\mathrm{Im}[G_0^R(x, x', \vec{k}_\parallel, z)]\big|_{\vec{r}=\vec{r}'=\vec{r}_0,\ \lambda_L\to\mu_L,\ \lambda_R\to\mu_R}$$

$$= -\frac{e\hbar^3(2s+1)}{2m^2\pi} \int d^3r \int d^3r'\, \delta(x-x_0)\,\delta(x'-x_0) \int_{-\infty}^{\infty} dE\,[f_L(E) - f_L(E+eV)] \Big(\frac{\partial}{\partial x}\,\mathrm{Im}G_L^R(\vec{r}, \vec{r}', E)\,\frac{\partial}{\partial x'}\,\mathrm{Im}G_R^R(\vec{r}', \vec{r}, E)$$

$$- \frac{\partial^2}{\partial x \partial x'}\,\mathrm{Im}G_L^R(\vec{r}, \vec{r}', E)\,\mathrm{Im}G_R^R(\vec{r}', \vec{r}, E) - \mathrm{Im}G_L^R(\vec{r}, \vec{r}', E)\,\frac{\partial^2}{\partial x \partial x'}\,\mathrm{Im}G_R^R(\vec{r}', \vec{r}, E)$$

$$+ \frac{\partial}{\partial x'}\,\mathrm{Im}G_L^R(\vec{r}, \vec{r}', E)\,\frac{\partial}{\partial x}\,\mathrm{Im}G_R^R(\vec{r}', \vec{r}, E) \Big)\,, \qquad (4.26a)$$

$$f_L(E) = (e^{\beta(E-\xi_L)} + 1)^{-1}\,, \qquad (4.26b)$$

$$G_i^R(\vec{r}, \vec{r}', E) \equiv \sum_{\vec{k}_\parallel} e^{i\vec{k}_\parallel\cdot(\vec{\rho}-\vec{\rho}')}\, G_i^R(x, x', \vec{k}_\parallel, E)\,,$$

$$i = (l, r)\,. \qquad (4.26c)$$

In deriving Eq. (4.26a), we have used the fact that

$$\int_{-\infty}^{\infty} dE\, f(E)\,\mathrm{Im}G_i^R(x, x', \vec{k}_\parallel, E)$$

$$= -\int_{-\infty}^{\infty} dE\, f(E) \textstyle\sum_k \psi_{kl}(x)\,\psi_{kl}^*(x')\,\pi\delta[E - E(k)]\,, \qquad (4.26d)$$

with an analogous term for G_r^R.

Equation (4.26a) is precisely the expression given by Appelbaum and Brinkman[8] for the current, except for the spin degeneracy factor, which these authors omit.

Let us apply Eq. (4.26a) to the square-barrier model of the average barrier [i.e., $V(x) = V_0$, $0 \le x \le b$] in which there is a δ-function impurity potential $U_0\delta(x-d)$. For simplicity, we assume that the junction is symmetric, i.e., $V_1 = 0$ in Fig. 1. In this case the TH current at $T = 0$ is, from elementary calculation (with spin degeneracy included),

$$j_{AB} = \frac{em}{2\pi^2\hbar^3} \int_{\xi_L - eV}^{\xi_L} dE \int_0^E dE_\perp\, \frac{16k^2K^4e^{-2Kb}}{(K^2+k^2)^2}\, D'\,. \qquad (4.27a)$$

$$D' \equiv \left| \left\{ (K - K_0) - K_0 \left[\left(\frac{K+ik}{K-ik} \right) (e^{2K(d-b)} + e^{-2Kd}) - \frac{K_0}{(K-K_0)} \left(\frac{K+ik}{K-ik} \right)^2 e^{-2Kb} \right] \right\} \right|^{-2}, \qquad (4.27b)$$

$$K_0 \equiv -mU_0/\hbar^2\,, \qquad (4.27c)$$

$$K \equiv [(2m/\hbar^2)(V_0 - E_\perp)]^{1/2}\,, \qquad (4.27d)$$

$$k \equiv (2mE_\perp/\hbar^2)^{1/2}\,. \qquad (4.27e)$$

For the same case, the exact current is

$$j = \frac{em}{2\pi^2\hbar^3} \int_{\xi_L - eV}^{\xi_L} dE \int_0^E dE_\perp\, \frac{16k^2K^4e^{-2Kb}}{(K^2+k^2)^2\,|\,1 - e^{-2Kb}(K+ik)/(K-ik)\,|^2}\, D\,, \qquad (4.28a)$$

$$D \equiv \left| \left\{ (K - K_0) - K_0 \left[\left(\frac{K+ik}{K-ik} \right) (e^{2K(d-b)} + e^{-2Kd}) + \left(\frac{K+ik}{K-ik} \right)^2 e^{-2Kb} \right] \right\} \right|^{-2}\,. \qquad (4.28b)$$

From our derivation of the Appelbaum-Brinkman formula for the current in Eq. (4.26a), we observe that the result does not depend upon the explicit form of the average barrier potential. The derivation *does* depend critically upon the assumption that $V_r = V_l = V$ in the barrier and that $|J_{qk}|^2 \ll 1$. The first of these assumptions is necessary to make the left-flowing current subtract exactly from the right-flowing current; i.e., the first assumption is necessary to derive the $f_L(E) - f_L(E+eV)$ form for the current in Eq. (4.26a). It is, therefore, inconsistent to assume *a priori* that the potential of an impurity near the interface between the left electrode and the barrier, for example, can be omitted from the Hamiltonian of the right electrode (as done, e.g., in Ref. 8).

The second assumption (i.e., $|J|^2 \ll 1$) is necessary for the series truncation involved in the step

of the derivation between Eqs. (4.23) and (4.24). In the case of a resonant-impurity potential in the barrier $|J|^2 \simeq 1$, and this truncation is invalid. This point is emphasized in the expressions for the TH and exact currents in Eqs. (4.27) and (4.28), respectively, calculated for the case of a δ-function impurity in a symmetric square barrier. One of the conditions for resonance is that $K \simeq K_0$. In this case, the term multiplying e^{-2Kb} in Eq. (4.27b) becomes very large, which contradicts the requirement that the width of the resonance be very small. By comparing Eqs. (4.27) and (4.28), we see that the TH theory is an inadequate description of resonant tunneling.

V. SUMMARY AND CONCLUSIONS

We have developed a new microscopic theory of tunneling from the two physical assumptions that electrons are in states of the entire junction and that a tunnel junction across which a bias is imposed is a steady-state system. In Sec. II we employed the first assumption in order to define the extended-basis functions (EBF). In Sec. III, we used the second assumption to develop the statistics of current flow. That is, we used the equilibrium temperature Green's function in the derivation of the "temperature current" of Eqs. (3.16), and then obtained the physical current by treating the bias as a boson index ν_s, which is analytically continued: $\nu_s \to 0$ $(-eV)$ for right (left) flowing current. In the discussion following Eq. (3.35), we gave a prescription for calculating the current in perturbation theory.

We displayed explicitly the two simplest applications of the method. The first is the calculation of the average-barrier one-electron current in Eq. (3.31a). The second application, described in Sec. IV, was the calculation of the current in a junction containing static-δ-function impurity potentials. This current was derived in Sec. IV [i.e., Eqs. (4.10)] and the expression shown to agree with the "exact" one-electron current in Eq. (3.31a). Although the static-δ-function model is a special case in which the perturbation-theory series can be summed explicitly, our perturbation-theory prescription obviously agrees with the one-electron theory of tunneling for any model of the static impurity potential.

We next examined the effect of resonant static impurity potentials upon the current. In this case, a new channel for current flow, i.e., a resonant elastic tunneling channel appears. Resonant elastic tunneling can produce zero-bias anomalies (ZBA) in the conductance. The shape of a ZBA depends upon the position of the resonance energy at zero bias relative to the Fermi level as well as the variation of the resonance energy with applied bias.

Finally, for the static-impurity model we displayed explicitly the limits in which our theory reduces to the TH method. As anticipated, we find that our method adequately describes resonant elastic tunneling for all types of junctions whereas the TH does not. We find additional differences between the two methods in our analysis of inelastic tunneling.[25,32]

APPENDIX: GENERAL THEOREMS

Theorem 1

From Eq. (3.15a),

$$\mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') \equiv -\mathrm{Tr}\{\rho T_\tau[\psi(\vec{r}, \tau)\psi^\dagger(\vec{r}', \tau')]\},$$

if

$$\mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') = \mathcal{G}(\vec{r}, \vec{r}', \tau - \tau', 0), \tag{A1}$$

then $\lambda_L = \lambda_R + 2\pi m/(-i\beta)$, where m is a positive or negative integer.

Proof. Let $\tau > \tau'$, then, from Eqs. (3.14) and (3.15),

$$\mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') = -\frac{1}{\mathfrak{d}} \sum_\alpha \langle \alpha | e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)} e^{-\beta\mathcal{K}} e^{\mathcal{K}\tau} \times \psi(\vec{r}) e^{\mathcal{K}(\tau' - \tau)} \psi^\dagger(\vec{r}') e^{-\mathcal{K}\tau'} | \alpha \rangle,$$

where we calculate the trace in the (complete) Fock representation in Eqs. (3.5) (i.e., eigenstates of N_L^0 and N_R^0, but not of $\mathcal{K}$). Invoking cyclic invariance of the trace and inserting a complete set of intermediate states $|\beta\rangle$ in the same occupation number representation yield

$$\mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') = -\frac{1}{\mathfrak{d}} \sum_{\alpha,\beta} \langle \alpha | e^{-\mathcal{K}\tau'} e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)} | \beta \rangle A_{\beta\alpha},$$

$$A_{\beta\alpha} \equiv \langle \beta | e^{-\mathcal{K}(\beta - \tau)} \psi(\vec{r}) e^{\mathcal{K}(\tau' - \tau)} \psi^\dagger(\vec{r}') | \alpha \rangle.$$

Define

$$\Delta\lambda \equiv \lambda_R - \lambda_L. \tag{A2}$$

Thus, $\lambda_L N_L^0 + \lambda_R N_R^0 = \lambda_L N + \Delta\lambda N_R^0$, where $N = N_L^0 + N_R^0$ from Eq. (3.6a). Since $[N, \mathcal{K}] = 0$, we can write

$$\mathcal{G}(\vec{r}, \vec{r}', \tau, \tau') = -\frac{1}{\mathfrak{d}} \sum_{\alpha,\beta} \langle \alpha | e^{-\mathcal{K}\tau'} e^{\beta(\lambda_L N + \Delta\lambda N_R^0)} | \beta \rangle A_{\beta\alpha}. \tag{A3a}$$

At this point, let us note that for $\tau > \tau'$, we can trivially derive that

$$\mathcal{G}(\vec{r}, \vec{r}', \tau - \tau', 0) = -\frac{1}{\mathfrak{d}} \sum_{\alpha,\beta} \langle \alpha | e^{\beta(\lambda_L N + \Delta\lambda N_R^0)} e^{-\mathcal{K}\tau'} | \beta \rangle A_{\beta\alpha}. \tag{A3b}$$

Equations (A3a) and (A3b) are equal only if

$$\langle \alpha | [e^{-\mathcal{K}\tau'}, e^{\beta\Delta\lambda N_R^0}] | \beta \rangle = 0 \tag{A4a}$$

for all $\langle \alpha |$ and $\langle \beta |$. Equation (A4a) is true only if

$$e^{\beta\Delta\lambda(n_R^\beta - n_R^\alpha)} = 1 \tag{A4b}$$

for all $\langle\alpha|$ and $|\beta\rangle$, where, according to Eqs. (3.5), n_R^α is the eigenvalue of N_R^0 in the state $|\alpha\rangle$. The relation in Eq. (A1) is true, therefore, only if

$$\lambda_L = \lambda_R + 2\pi m/(-i\beta)$$

for m an integer. A similar proof applies for $\tau < \tau'$.

In the case that $\mathcal{H}=\mathcal{H}_0$, Eq. (A4a) is true, from Eqs. (3.5), for *all* $|\alpha\rangle$ and $|\beta\rangle$. Therefore, there is *no* restriction upon $\Delta\lambda$ in this situation and the theorem is a truism here.

Theorem 2

Make the definition

$$\mathcal{G}(\vec{r}, \vec{r}', \tau) \equiv \mathcal{G}(\vec{r}, \vec{r}', \tau, 0) .$$

If

$$\lambda_L = \lambda_R + 2\pi m/(-i\beta) ,$$

then

$$\mathcal{G}(\vec{r}, \vec{r}', \tau < 0) = -e^{\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau + \beta) \qquad \text{(A5a)}$$

and

$$\mathcal{G}(\vec{r}, \vec{r}', \tau > 0) = -e^{-\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau - \beta) . \qquad \text{(A5b)}$$

Proof. From cyclic invariance of the trace and the definition of T_τ,

$$\begin{aligned}
&\mathcal{G}(\vec{r}, \vec{r}', \tau) \\
&= -\mathrm{Tr}[\Theta(\tau)\,\psi^\dagger(\vec{r}')\,\rho\psi(\vec{r}, \tau) - \Theta(-\tau)\,\psi(\vec{r}, \tau)\,\rho\psi^\dagger(\vec{r}')] \\
&= -\frac{1}{\mathfrak{Z}}\,\mathrm{Tr}\{e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)}\, e^{-\beta\mathcal{H}}\, e^{\beta\mathcal{H}}\, e^{-\beta(\lambda_L N + \Delta\lambda N_R^0)} \\
&\quad \times [\Theta(\tau)\,\psi^\dagger(\vec{r}')\, e^{\beta(\lambda_L N + \Delta\lambda N_R^0)}\, e^{-\beta\mathcal{H}}\, \psi(\vec{r}, \tau) \\
&\quad - \Theta(-\tau)\,\psi(\vec{r}, \tau)\, e^{\beta(\lambda_L N + \Delta\lambda N_R^0)}\, e^{-\beta\mathcal{H}}\, \psi^\dagger(\vec{r}')]\} .
\end{aligned}$$

For simplicity, we refer again to the Fock representation in Eqs. (3.5), since the trace is invariant. From Eq. (A4a),

$$\langle\alpha|[e^{\beta\Delta\lambda N_R^0}, \psi]|\beta\rangle = \langle\alpha|[e^{\beta\Delta\lambda N_R^0}, \psi^\dagger]|\beta\rangle = 0 ,$$

since we have assumed that $\Delta\lambda$ is a boson index. In addition,

$$e^{-\beta\lambda_L N}\,\psi e^{\beta\lambda_L N} = e^{\beta\lambda_L} , \quad e^{-\beta\lambda_L N}\,\psi^\dagger e^{\beta\lambda_L N} = e^{-\beta\lambda_L} .$$

Therefore,

$$\begin{aligned}
\mathcal{G}(\vec{r}, \vec{r}', \tau) &= -(1/\mathfrak{Z})\,\mathrm{Tr}\{e^{\beta(\lambda_L N_L^0 + \lambda_R N_R^0)}\, e^{-\beta\mathcal{H}}[\Theta(\tau)\, e^{-\beta\lambda_L}\,\psi^\dagger(\vec{r}', \beta)\,\psi(\vec{r}, \tau) - \Theta(-\tau)\, e^{\beta\lambda_L}\,\psi(\vec{r}, \tau + \beta)\,\psi^\dagger(\vec{r}', 0)]\} \\
&= -\Theta(\tau)\, e^{-\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau, \beta) - \Theta(-\tau)\, e^{\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau + \beta, 0) , \\
&= -\Theta(\tau)\, e^{-\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau - \beta) - \Theta(-\tau)\, e^{\beta\lambda_L}\, \mathcal{G}(\vec{r}, \vec{r}', \tau + \beta) .
\end{aligned} \qquad \text{(A6)}$$

Thus, theorem 2 is proved.

*Work supported in part by the Advanced Research Projects Agency, under Contract No. HC 15-67-C-0221.

†Present address: General Motors Research Laboratories, Warren, Mich. 48090.

[1]J. Bardeen, Phys. Rev. Letters 6, 57 (1961).

[2]M. H. Cohen, L. M. Falicov, and J. C. Phillips, Phys. Rev. Letters 8, 316 (1962).

[3]I. Giaever, Phys. Rev. Letters 5, 147 (1960); 5, 464 (1960).

[4]B. D. Josephson, Phys. Letters 1, 251 (1962).

[5]E. L. Wolf, Phys. Rev. Letters 20, 204 (1968).

[6]L. C. Davis and C. B. Duke, Phys. Rev. 184, 764 (1969).

[7]L. C. Davis, Phys. Rev. B 2, 1714 (1970).

[8]J. A. Appelbaum and W. F. Brinkman, Phys. Rev. 186, 464 (1969); Phys. Rev. B 2, 907 (1970).

[9]L. C. Davis, Phys. Rev. 187, 1177 (1969).

[10]C. B. Duke, G. G. Kleiman, A. M. Andrews, R. D. Burnham, N. Holonyak, Jr., and H. W. Korb, in *Proceedings of the Tenth International Conference on the Physics of Semiconductors, Cambridge, Mass.*, 1970 (U. S. AEC, Division of Technical Information, Springfield, Va., 1970).

[11]D. E. Cullen, E. L. Wolf, and W. Dale Compton, Phys. Rev. B 2, 3157 (1970).

[12] M. Mikkor and W. C. Vassell, Phys. Rev. B 2, 1875 (1970).

[13] E. L. Wolf, D. L. Losee, D. E. Cullen, and W. Dale Compton, Phys. Rev. Letters 26, 438 (1971).

[14]I. Giaever and H. R. Zeller, Phys. Rev. B 1, 4278 (1970).

[15]I. Giaever and H. R. Zeller, Phys. Rev. Letters 20, 1504 (1969).

[16]H. R. Zeller and I. Giaever, Phys. Rev. 181, 789 (1969).

[17]E. W. Plummer and R. D. Young, Phys. Rev. B 1, 2088 (1970).

[18]C. B. Duke and M. E. Alferieff, J. Chem. Phys. 46, 923 (1967); D. Penn, R. Gomer, and M. H. Cohen, Phys. Rev. B 5, 768 (1972).

[19]C. B. Duke, *Tunneling in Solids* (Academic, New York, 1969).

[20]S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge U. P., Cambridge, England, 1960).

[21]D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, N. J., 1951).

[22]A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N. J., 1963).

[23]J. G. Adler, H. J. Kreuzer, and W. J. Wattamaniuk, Phys. Rev. Letters 27, 185 (1971).

[24]A. M. Andrews, H. W. Korb, N. Holonyak, Jr., C. B. Duke, and G. G. Kleiman, Phys. Rev. B 5, 2273 (1972).
[25]G. G. Kleiman (unpublished).
[26]A. Zawadowski, Phys. Rev. 163, 341 (1967).
[27]C. Caroli, R. Combescot, P. Nozières, and D. Saint James, J. Phys. C 4, 916 (1971); C. Caroli, R. Combescot, D. Lederer, P. Nozières, and D. Saint James, *ibid.* 4, 2598 (1971); R. Combescot, *ibid.* 4, 2611 (1971).
[28]A preliminary version of a new theory has been presented by G. K. Birkner and W. Schattke [Phys. Letters 36A, 163 (1971)]. Further comment awaits further details of this theory.
[29]L. D. Landau and E. M. Lifschitz, *Quantum Mechanics* (Pergamon, London, 1958).
[30]A. J. Bennett and C. B. Duke, Phys. Rev. 160, 541 (1967).
[31]E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[32]G. G. Kleiman, Ph.D. thesis (University of Illinois, 1971) (unpublished).
[33]F. B. Hildebrand, *Methods of Applied Mathematics* (Prentice-Hall, Englewood Cliffs, N. J., 1965).
[34]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[35]J. M. Ziman, *Elements of Advanced Quantum Theory* (Cambridge U. P., Cambridge, England, 1969).
[36]L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).
[37]J. W. Gadzuk, J. Appl. Phys. 41, 286 (1970).
[38]G. A. Baym, *Lectures on Quantum Mechanics* (Benjamin, New York, 1969).

PHYSICAL REVIEW B VOLUME 6, NUMBER 6 15 SEPTEMBER 1972

Ultrasonic Harmonic Generation in Magnetic Fields

Harold N. Spector
Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616
(Received 1 December 1971)

Second-harmonic generation of ultrasound propagating transverse to a dc magnetic field in a piezoelectric semiconductor is treated using a phenomenological approach which is valid when either $qR \ll 1$ or $ql \ll 1$. We find that the magnetic field can change both the magnitude of the harmonic generation and the frequency at which the harmonic generation has its peak value. It is found that at frequencies near or below the frequency of maximum gain the magnitude of the harmonic generation has a maximum as a function of magnetic field. Therefore, the application of a transverse magnetic field can lead to an enhancement of second-harmonic generation in a piezoelectric semiconductor.

I. INTRODUCTION

When large-amplitude acoustic flux propagates in a piezoelectric semiconductor, the acoustoelectric interaction between the acoustic flux and the conduction electrons leads to frequency mixing of the waves comprising the flux. One such frequency-mixing effect of particular interest is second-harmonic generation. Ultrasonic-second-harmonic generation due to such acoustoelectric interactions was first observed in photoconducting CdS by Tell.[1] Recently, there has been a revival of interest in such frequency-mixing effects, mainly because of the role they play in the growth of domains of acoustic flux under conditions of acoustic amplification.[2-7] Most of the work involving the nonlinear acoustoelectric interactions has concerned the situation in the absence of any external magnetic field. However, in high-mobility semiconductors such as InSb, the application of a strong magnetic field can crucially alter the behavior of the acoustoelectric interaction.[8-12] When such a field is applied transverse to the direction of propagation of the ultrasound, the acoustic gain or loss due to the acoustoelectric interaction is greatly enhanced over its value in the absence of the field. Moreover, the types of acoustoelectric current oscillation which occur in piezoelectric semiconductors are also greatly altered by the application of strong magnetic fields.[13] It is therefore of interest to investigate the effect of a magnetic field on the second-harmonic generation due to the nonlinear acoustoelectric interactions.

In Sec. II we present the theory of second-harmonic generation due to the interaction between the ultrasound and the conduction electrons in the presence of a dc magnetic field. This is done using a phenomenological theory which is valid when the sound wavelength is much greater than the average distance the carrier travels between collisions. In weak magnetic fields $\omega_0\tau \ll 1$, where the carrier does not have much of a chance of being deflected by the field before undergoing a collision, this distance is the carrier's mean free path l. In strong fields $\omega_0\tau \gg 1$, on the other hand, the carrier completes several orbits before being scattered and this distance is of the order of the cyclotron radius R.[14] The theory is then applied to second-harmonic