PHYSICAL REVIEW B VOLUME 25, NUMBER 7 1 APRIL 1982

Transition from metallic to tunneling regimes in superconducting microconstrictions: Excess current, charge imbalance, and supercurrent conversion

G. E. Blonder, M. Tinkham, and T. M. Klapwijk*
Physics Department, Harvard University, Cambridge, Massachusetts 02138
(Received 19 October 1981)

We propose a simple theory for the *I-V* curves of normal-superconducting microconstriction contacts which describes the crossover from metallic to tunnel junction behavior. The detailed calculations are performed within a generalized semiconductor model, with the use of the Bogoliubov equations to treat the transmission and reflection of particles at the *N-S* interface. By including a barrier of arbitrary strength at the interface, we have computed a family of *I-V* curves ranging from the tunnel junction to the metallic limit. Excess current, generated by Andreev reflection, is found to vary smoothly from $4\Delta/3eR_N$ in the metallic case to zero for the tunnel junction. Charge-imbalance generation, previously calculated only for tunnel barriers, has been recalculated for an arbitrary barrier strength, and detailed insight into the conversion of normal current to supercurrent at the interface is obtained. We emphasize that the calculated differential conductance offers a particularly direct experimental test of the predictions of the model.

I. INTRODUCTION

In recent years there has been much experimental and theoretical work on small-area high-current-density superconducting junctions, since such junctions have advantages for certain practical applications. Although the properties (*I-V* curves, etc.) of classic (high-barrier) tunnel junctions have long been understood in great detail,[1] the properties of metallic (no-barrier) junctions have been studied only more recently[2–4] and in less detail, and the transitional case (small barrier) has received practically no attention at all. Here, we present a new unified treatment which is sufficiently general to embrace both limits, as well as the intermediate regime.

In this paper, we confine our attention primarily to the *N-S* interface in the context of a generalized Andreev reflection model. This avoids the complications of the Josephson effect, while allowing complete delineation of *I-V* curves including the so-called "excess currrent" at high voltage, as it increases from zero in tunnel junctions to its limiting value in clean metallic contacts.

The paper begins with a careful clarification of the relation between a general semiconductor model of the superconducting energy levels and the excitations of the BCS model and the Bogoliubov equations. With this in hand, we are quickly led to complete *I-V* curves for the case of arbitrary barrier strength. We also compute the generalization of the tunnel-injection results for the generation of charge imbalance in the superconductor, and give new insight into the detailed processes involved in the conversion from normal to supercurrent at an *N-S* interface. Some of the results which we obtain by our simple technique have been obtained earlier by other more mathematically complex theories, but our method gives greater physical insight. Moreover, the simplicity of our method has allowed us to treat a broader range of problems having direct experimental implications. We have also used this model in a previous publication[5] to explain the origin of the excess current and of the subharmonic gap structure which is observed in metallic *S-S* contacts.

II. GENERALIZED SEMICONDUCTOR SCHEME

It has long been conventional to simplify the calculation of tunnel currents by taking advantage of the fact that the BCS coherence factors u_k and v_k drop out of the computation if excitations with the same energy E_k, but with $k < k_F$ and $k > k_F$, are grouped together. Then a very simple "semicon-

ductor model" can be established, in which only the superconducting density of states $N_s(E)$ distinguishes the superconductor from the situation in the normal state. However, in combining $k_<$ and $k_>$ for each E_k, information is lost. For instance, the quasiparticle charge imbalance Q^* refers specifically to the imbalance between the two branches $k > k_F$ and $k < k_F$ of the quasiparticle spectrum, a distinction which is completely suppressed in the simplest semiconductor model. Another shortcoming of the elementary semiconductor model is that the density-of-states factor makes no reference to the microscopic nature of the states. For example, it is clear that the current must be carried toward and away from the junction by waves traveling in the appropriate directions, yet this physical fact is obscured in the usual tunnel-Hamiltonian model.

Here, we describe a generalized semiconductor model specifically designed to avoid this loss of information. The key innovation is the use of the Bogoliubov equations[6] to handle the interface; this treats all cases from a clean metallic interface to a tunnel junction in the same formal way by matching wave functions at the boundary. These results determine the probability current carried by the various excitations characterized by the semiconductor diagram. For simplicity, we restrict attention to metal clean enough to allow use of momentum as a good quantum number for labeling the various waves.

If one considers only energies $E > \Delta$, the Bogoliubov equation solutions (see Appendix for details) for the incident, transmitted, and reflected particles can be identified with the conventional BCS quasiparticle excitations with energy

$$E_k = (\Delta^2 + \epsilon_k^2)^{1/2} , \tag{1}$$

with respect to the electrochemical potential for condensed pairs, where

$$\epsilon_k = \frac{\hbar^2 k^2}{2m} - \epsilon_F . \tag{2}$$

Because only the square of ϵ_k enters the expression for E_k, there will be a pair of magnitudes of k associated with each energy, namely,

$$\hbar k^{\pm} = (2m)^{1/2}[\epsilon_F \pm (E_k^2 - \Delta^2)^{1/2}]^{1/2} . \tag{3}$$

Moreover, because of the BCS pairing of k and $-k$, one must consider both signs of k together, so that there is a fourfold degeneracy of relevant states for each E, as sketched in Fig. 1. Of course there is an additional twofold spin degeneracy, but in the absense of spin-flip processes this affects only normalizations, and it need not concern us here.

The nature of one type of these BCS quasiparticle states is given by the usual Bogoliubov transformation[7] as

$$\gamma_{k0}^* = u_k c_{k\uparrow}^* - v_k c_{-k\downarrow} , \tag{4}$$

where the BCS coherence factors are given by

$$u_k^2 = 1 - v_k^2 = \tfrac{1}{2}(1 + \epsilon_k / E_k) . \tag{5}$$

Although, in general, Eq. (5) determines the *magnitudes* of u_k and v_k (which can have phase factors), for simplicity we take them to be real and positive. For the *N-S* case treated here, this involves no loss of generality; only in *S-S* devices does phase play a role. The excitations described in Eq. (4) consist partly of electrons created at $+k$ (i.e., $u_k c_{k\uparrow}^*$) and partly of holes created at $-k$ (i.e., $v_k c_{-k\downarrow}$, the destruction of an electron at $-k$). Since either term increases the *system* momentum by $+k$, the subscript k in γ_{k0}^* is a well-defined momentum index suitable for labeling. From Eq. (5) we see that the excitations at $\pm k^+$ are predominantly electronlike, while these at $\pm k^-$ are predominantly holelike. The other members of the degenerate set are described by including the complementary BCS operator γ_{k1} and also including both senses of k.

In preparation for the application of boundary conditions, we note that the excitation described by Eq. (4) is described in the Bogoliubov-equation formalism (see Appendix for details) as a two-element column vector,

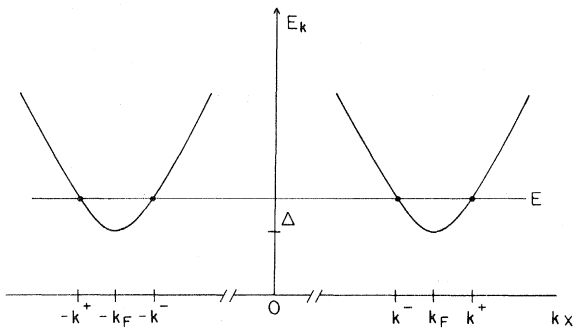

FIG. 1. Schematic plot of excitation energies E_k vs k along an axis passing through the center of the Fermi sphere in the superconducting state. The scale has been greatly expanded near $\pm k_F$.

$$\psi_{k0}=\begin{bmatrix} f_k(x,t) \\ g_k(x,t) \end{bmatrix} . \tag{6}$$

Here f_k and g_k satisfy the differential equations

$$\left[-\frac{\hbar^2}{2m}\nabla^2-\mu(x)+V(x)\right]f+\Delta g = i\hbar\frac{\partial f}{\partial t}\rightarrow Ef , \tag{7}$$

$$-\left[-\frac{\hbar^2}{2m}\nabla^2-\mu(x)+V(x)\right]g+\Delta f = i\hbar\frac{\partial g}{\partial t}\rightarrow Eg ,$$

and they are proportional to u_k and v_k, respectively, when $V(x)$ is constant. The time derivatives may be replaced as indicated in solutions for stationary states of energy E.

As written in Eq. (4), the excitations γ^*_{k0} change the charge of the superconductor by the fractional electronic charge $q_k=u_k^2-v_k^2=\epsilon_k/E_k=\pm N_s(E)^{-1}$. Since physically allowable operators must conserve charge exactly, it is convenient[7] to introduce operators of the form

$$\gamma^*_{ek0}=u_k c^*_{k\uparrow}-v_k S^* c_{-k\downarrow} , \tag{8a}$$

$$\gamma^*_{hk0}=u_k S c^*_{k\uparrow}-v_k c_{-k\downarrow}=S\gamma^*_{ek0} , \tag{8b}$$

where S^* adds a pair to the condensate and S destroys one. These operators create excitations while changing the charge by exactly $\pm e$, corresponding to the subscripts e and h for electron and hole, respectively. Accordingly, they are better adapted to describe the energy levels analogous to those of ordinary electrons. The *excitations* described by Eqs. (8a) and (8b) are identical, and in no sense cancel as one might suppose for electron and hole operators. The systems resulting from (8a) and (8b) differ only by one condensed pair from the reservoir at the chemical potential.

Now for any system with fermion excitations such as these, the total system energy can be written as

$$E=E_G+\sum_k E_k\gamma^*_k\gamma_k+\mu N , \tag{9}$$

where E_G is the ground-state energy, and the sum runs over all the excitations, which have $E_k\geq 0$. The electrochemical potential μ must be introduced explicitly since we are interested in treating processes which change the number N of electrons in a subsystem at given μ by ± 1. In fact, the energy required to make an excitation with charge e is

$$E_{ek}=\mu+E_k , \tag{10a}$$

while that to make an excitation with charge $-e$ differs by 2μ and is

$$E_{hk}=-\mu+E_k=-(\mu-E_k) . \tag{10b}$$

We can now tabulate (see Table I) the possible charge-conserving processes involving subsystems 1 and 2, and write down the corresponding condition set by the conservation of energy. There are four distinct cases, each of which conserves energy equally well in reverse, making eight cases in all. As is evident from Table I, *all* these energetically allowed transitions occur as "horizontal" ones *only* if we plot *both* $\pm E_{ki}$ symmetrically about each μ_i, as is motivated by recalling that $\mu-E_k=-E_{hk}$ (see Fig. 2). Further, the four cases of Table I include *all* the combinations of $\pm E_k$, so that *all* horizontal transitions in such a "reflected" or semiconductor-model diagram are *energetically* allowed.

There is, however, another constraint on the system; namely, a particle incident from the left can only produce transmitted particles with positive group velocity ($dE/d\hbar k$) and reflected ones with negative group velocities. In general, there is a finite probability of all the transfers allowed by these considerations to occur, but in the special case of zero barrier and no discontinuity in electronic-band parameters across the N-S interface, it turns out that there is essentially zero crossover between channels inside and outside the Fermi surface, even near the gap edge. It is here that the Bogoliubov equations prove most useful: The semiconductor diagram shows which transitions are energetically allowed; the Bogoliubov equations provide the

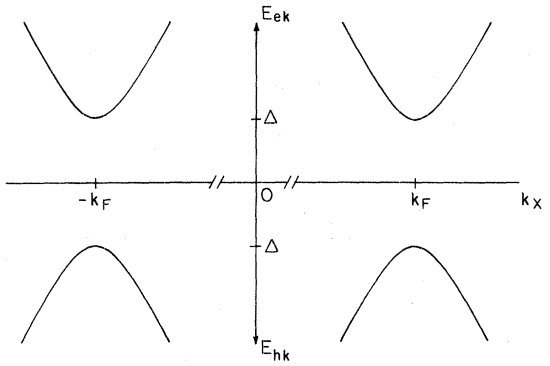

FIG. 2. Semiconductor version of Fig. 1. Inclusion of the quasiparticle branches in the lower half plane allows all the transitions in Table I to be "horizontal."

TABLE I. Processes involving transfer of a single electronic charge.

Process	Energy condition
1. Electron from 1→2 (or reverse)	$\mu_1+E_{k1}=\mu_2+E_{k2}$
2. Hole from 1→2 (or reverse)	$\mu_1-E_{k1}=\mu_2-E_{k2}$
3. Create electron in 1 and hole in 2 (or destroy both)	$\mu_1+E_{k1}=\mu_2-E_{k2}$
4. Create hole in 1 and electron in 2 (or destroy both)	$\mu_1-E_{k1}=\mu_2+E_{k2}$

quantitative probability that each will occur.

As a final preliminary, we consider the special case of the normal metal, obtained by letting $\Delta\rightarrow 0$ in the expressions for the superconducting case. Here the situation simplifies from the superconducting case shown in Fig. 2 because electronlike excitations cannot be made inside the Fermi sphere, since the states there are fully occupied in the ground state, and excitations must be orthogonal to the ground state. Similarly, hole excitations are possible only outside the Fermi sphere. As a result, in the normal metal only the branches shown in Fig. 3 exist.

III. THE *N-S* BOUNDARY IN EQUILIBRIUM

We start by considering the dynamic equilibrium which exists at an *N-S* interface with no applied voltage, as sketched in Fig. 4. In this equilibrium situation, all relevant quaisparticle states (such as those labeled 0, 1, 2, 3, 4, and 5 in the figure) are occupied with the same probability $f_0(E)$, which also gives the probability of a hole in the state 6 at $-E$. Viewed kinematically in a semiclassical approximation, however, particles approach the interface and are transmitted and reflected with certain probabilities, which are given by the products of squared amplitudes of Bogoliubov-equation solutions times their appropriate group velocities. The self-consistency of the equilibrium state provides a useful check on the computed probability coefficients.

To be concrete, consider an electron incident on the interface from the normal state with energy $E>\Delta$, as indicated by the arrow at the state labeled 0 in Fig. 4. By matching slope and value of the wave function across the interface, one finds the probabilities $A(E)$, $B(E)$, $C(E)$, and $D(E)$ for the four processes indicated by those labels in Fig. 4, i.e., outgoing particles at points 6, 5, 4, and 2, respectively. In words, $C(E)$ is the probability of transmission through the interface with a wave vector on the same side of the Fermi surface (i.e., $q^+\rightarrow k^+$, not $-k^-$), while $D(E)$ gives the probability of transmission with crossing through the Fermi surface (i.e., $q^+\rightarrow -k^-$). $B(E)$ is the probability of ordinary reflection, while $A(E)$ is the probability of Andreev[8] reflection as a hole on the other side of the Fermi surface. The latter process

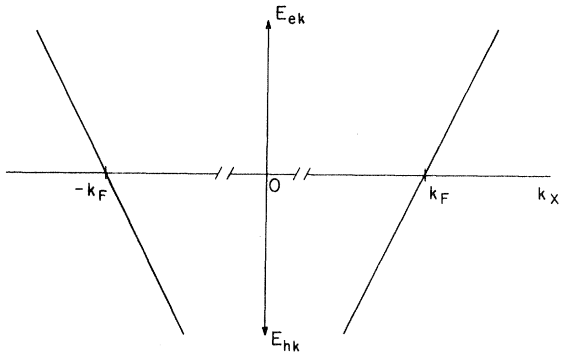

FIG. 3. E_k vs k for the normal state.

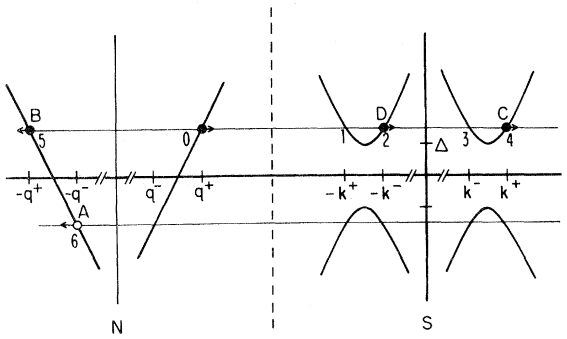

FIG. 4. Schematic diagram of energy vs momentum at *N-S* interface. The open circles denote holes, the closed circles electrons, and the arrows point in the direction of the group velocity. This figure describes an incident electron at (0), along with the resulting transmitted (2,4) and reflected (5,6) particles.

is not included in Table I, since it involves a transfer of a *pair* carrying $2e$ across the interface, while the processes in Table I transfer only a single electronic charge. The energy conservation condition appropriate for Andreev reflection is found by equating the energy of the initial electronlike excitation with that of the final holelike excitation plus a Cooper pair at the chemical potential. That is, $E_{ek}=E_{hk}+2\mu$, from which it follows that

$$|E_e|=|E_h| \ . \tag{11}$$

In other words, the hole is generated as far below μ as the electron was above, as drawn in Fig. 4. Of course, this result is obvious in the excitation representation.

It should be clearly understood that the probabilities A, B, C, and D (and the similar set A', B', C', and D' for quasiparticles incident from the superconducting side) are those appropriate for probability *current,* taking account of the cancelling effects of state density and group velocity. For example, consider the flow of particles across the interface in the energy range E to $E+dE$, and including only states outside the Fermi surface. From left to right, the current from states at 0 to those at 4 in Fig. 4 is proportional to $v_F C(E) f_0(E) dE$; the balancing current from right to left, i.e., from states at 1 to those at 5, is proportional to $v_F C'(E) f_0(E) N_s(E) dE$, where $N_s(E)$ is the normalized density of states in the superconductor. In equilibrium these flows are exactly equal, so that

$$C'(E)N_s(E)=C(E) \ , \tag{12a}$$

and similarly

$$D'(E)N_s(E)=D(E) \ , \tag{12b}$$

governs the branch-crossing transfers. The results of Eqs. (12) can be derived by direct computation, or by appeal to time-reversal symmetry. For simplicity, in the remainder of the paper we write all expressions in terms of $C(E)$ and $D(E)$, avoiding the need for explicit inclusion of the density-of-states factor $N_s(E)$ or the quasiparticle group velocity $v_g=N_s^{-1}v_F$ in integrations over energy.

Recognizing the symmetry of the problem about μ, we see that $A(E)$ and $B(E)$ are even functions of E. Similarly, $C(E)$ and $D(E)$ are also even functions of E, provided one uses the convention that $D(E)$ describes processes in which the transmission involves branch crossing, while $C(E)$ governs transmission processes in which the particle stays on the same branch.

It is important to note that the conservation of probability requires that

$$A(E)+B(E)+C(E)+D(E)=1 \ . \tag{13}$$

This result is particularly useful in simplifying expressions for energies below the gap, $|E|<\Delta$, where there can be no transmitted quasiparticles, so that $C=D=0$. Then, Eq. (13) reduces simply to $A(E)=1-B(E)$, so that a single function of energy is all that is needed.

Now, in general, A, B, C, and D depend on the angle of incidence of the trajectory and on the detailed shape of the scattering potential. For simplicity, we restrict ourselves to a one-dimensional (1D) geometry. For the purposes of this paper, the actual three-dimensional (3D) geometry is invoked only to enforce three boundary conditions on the 1D model: that the junction is well-cooled, and that the energy gap and the electric potential both rise to their full asymptotic values on a scale shorter than ξ on either side of the neck.

To capture the essential effect of any interfacial scattering with a single parameter, we model it by a repulsive potential $H\delta(x)$ located at the interface. This is meant to represent the effect of the typical oxide layer in a point contact, the localized disorder in the neck of a short microbridge, or the intentional oxide barrier in a high-current-density tunnel junction. The calculations with the Bogoliubov equations are quite straighforward, but are given in the Appendix to avoid a digression at this point. The results of the calculations are values of the four coefficients A, B, C, and D as functions of energy and barrier strength. To simplify the formulas, we have introduced a dimensionless barrier strength $Z=k_F H/2\epsilon_F=H/\hbar v_F$. The significance of this Z is illuminated by noting that the transmission coefficient in the normal state is simply $(1+Z^2)^{-1}$, and the corresponding reflection coefficient is $Z^2/(1+Z^2)$.

The expressions for the energy dependences of A, B, C, and D for $E>\Delta$ can be conveniently written in terms of u_0 and v_0, the BCS parameters u and v evaluated on the branch outside the Fermi surface. The results are as given in Table II. For convenience, in addition to the general results we also list the limiting forms of the results for zero barrier ($Z=0$) and for a strong barrier [$Z^2(u^2-v^2)>>1$], as well as for $\Delta=0$ (the normal-metal case). Curves computed using the general results are plotted in Fig. 5 for $Z=0$, 0.3, 1, and 3. In the special case of $E=\Delta$, one notes from Table II that $A=1$ and $B=C=D=0$, independent of Z.

TABLE II. Transmission and reflection coefficients. A gives the probability of Andreev reflection (i.e., reflection with branch crossing), B of ordinary reflection, C of transmission without branch crossing, and D of transmission with branch crossing (see text). ($\gamma^2=[u_0^2+Z^2(u_0^2-v_0^2)]^2$, $u_0^2=1-v_0^2=\frac{1}{2}\{1+[(E^2-\Delta^2)/E^2]^{1/2}\}$, and $N_s(E)=(u_0^2-v_0^2)^{-1}$.)

	A	B	C	D
Normal state	0	$\frac{Z^2}{1+Z^2}$	$\frac{1}{1+Z^2}$	0
General form				
$E<\Delta$	$\frac{\Delta^2}{E^2+(\Delta^2-E^2)(1+2Z^2)^2}$	$1-A$	0	0
$E>\Delta$	$\frac{u_0^2v_0^2}{\gamma^2}$	$\frac{(u_0^2-v_0^2)^2Z^2(1+Z^2)}{\gamma^2}$	$\frac{u_0^2(u_0^2-v_0^2)(1+Z^2)}{\gamma^2}$	$\frac{v_0^2(u_0^2-v_0^2)Z^2}{\gamma^2}$
No barrier ($Z=0$)				
$E<\Delta$	1	0	0	0
$E>\Delta$	v_0^2/u_0^2	0	$1-A$	0
Strong barrier $[Z^2(u^2-v^2)>>1]$				
$E<\Delta$	$\frac{\Delta^2}{4Z^2(\Delta^2-E^2)}$	$1-A$	0	0
$E>\Delta$	$\frac{u_0^2v_0^2}{Z^4(u_0^2-v_0^2)^2}$	$1-\frac{1}{Z^2(u_0^2-v_0^2)}$	$\frac{u_0^2}{Z^2(u_0^2-v_0^2)}$	$\frac{v_0^2}{Z^2(u_0^2-v_0^2)}$

Accordingly, except for $Z=0$, $A(E)$ has a sharp peak at the gap edge, although this peak becomes unobservably narrow for large values of Z.

Although we have developed this technique quite independently, other workers have also employed the Bogoliubov equations, and along similar lines. Andreev[8] in 1964 used them to calculate the thermal-boundary resistance at an N-S interface. Thus, he was essentially concerned with the restriction of quasiparticle flow from N to S caused by the gap. Kümmel[9] expanded the equations to include the condensate phase, and then used them to calculate the momentum balance of electrical-current flow near a vortex core. He emphasized the interchange of current between quasiparticles and the pairs in geometries where the gap varied with position. Then, in 1971 Demers and Griffin[10] and Griffin and Demers,[11] calculated the transmission coefficients in N-S-N and S-N-S geometries and made the extension to a δ-function barrier at the interface. Surprisingly, this important work remained essentially unnoticed for a number of years but recently (1977) Entin-Wohlman[12] has drawn on it to develop boundary conditions that she used in solving the Gor'kov gap equation at an N-S interface. Thus, use of the Bogoliubov equations to determine the fraction and type of reflect-

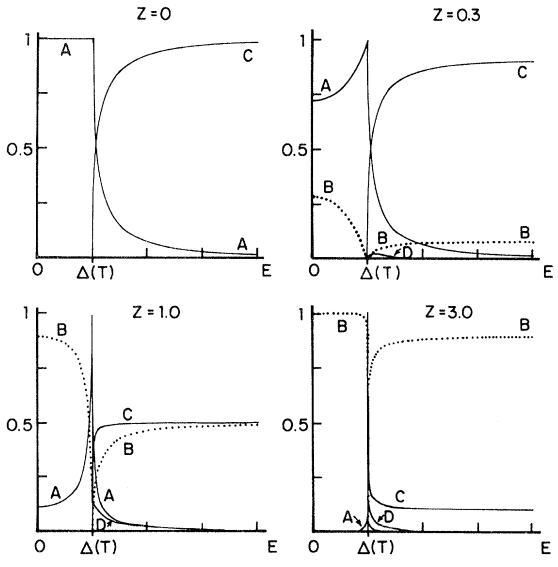

FIG. 5. Plots of transmission and reflection coefficients at N-S interface. A gives probability of Andreev reflection, B gives probability of ordinary reflection, C gives transmission probability without branch crossing, and D gives probability of transmission with branch crossing. The parameter Z measures the barrier strength at the interface.

ed particles is not new, though perhaps the technique remains underutilized. What *is* new is our extension of the method to calculations at finite bias voltage (as described in the next section) and our development of a framework that allows the physical consequences of these processes to emerge more clearly.

IV. THE *N-S* BOUNDARY AT FINITE VOLTAGE: *I-V* CURVES

When a voltage is applied, nonequilibrium quasiparticle populations will be generated, which can in general be found only by obtaining a self-consistent solution to a suitable Boltzmann equation. This solution is greatly simplified if one can assume ballistic acceleration of the particles without scattering. This is a good approximation for our case of a small orifice connecting massive electrodes, so long as the diameter of the orifice is small compared to a mean-free path. Moreover, in this regime, Bogoliubov-equation solutions neglecting scattering potentials except at the interface should also be usable.[13] The physical assumption which we use to define the problem is then that the distribution functions of all *incoming* particles are given by equilibrium Fermi functions, apart from the energy shift due to the accelerating potential. It is convenient to choose the electrochemical potential of the pairs in the superconductor as our reference level, since that remains a well-defined quantity even when the quasiparticle populations are far from equilibrium. Moreover, it is the quantity which defines the location of the "mirror plane" of the semiconductor diagram. With this convention, all *incoming* electrons from the S side have the distribution function $f_0(E)$, while those coming in from the N side are described by $f_0(E-eV)$.

Since the current must be conserved, it can be calculated in any plane. It is particularly convenient to do so on the N side of the interface, where all current is carried by single particles, and none as a supercurrent. To find the current in our 1D model, we take the difference between $f_{\rightarrow}(E)$ and $f_{\leftarrow}(E)$, the distribution functions at points such as 0 and 5 on Fig. 4, and integrate over E. That is,

$$I=\mathscr{A}J = 2N(0)ev_F\mathscr{A}\int_{-\infty}^{\infty}[f_{\rightarrow}(E)-f_{\leftarrow}(E)]dE\ , \tag{14}$$

where $\mathscr{A}$ is an effective-neck cross-sectional area, including a numerical factor for angular averaging which will depend on the actual 3D geometry. For example, in the orifice model of a point contact, $\mathscr{A}=\pi a^2/4$, where a is the radius of the orifice.[14] As usual, $N(0)$ refers to the one-spin density of states at ϵ_F.

With our assumption about incoming populations, it follows that

$$f_{\rightarrow}(E)=f_0(E-eV)\ , \tag{15}$$

while

$$f_{\leftarrow}(E)=A(E)[1-f_{\rightarrow}(-E)]+B(E)f_{\rightarrow}(E)+[C(E)+D(E)]f_0(E)\ . \tag{16}$$

In writing Eq. (16), we have chosen to write $f_{\rightarrow}(E)$ instead of $f_0(E-eV)$ in anticipation of later dealing with more complicated problems where $f_{\rightarrow}(E)$ and $f_{\leftarrow}(E)$ must be found self-consistently. Even in such cases, Eq. (16) remains valid as it stands, and serves as a boundary condition for a Boltzmann equation, whose solution would be inserted into the general expression (14) to find the current. However, in the present simple case, we can simply substitute Eqs. (15) and (16) into Eq. (14), obtaining

$$\begin{aligned} I_{NS} &= 2N(0)ev_F\mathscr{A}\int_{-\infty}^{\infty}(f_0(E-eV)-\{A(E)f_0(E+eV)+B(E)f_0(E-eV)\\ &\qquad +[1-A(E)-B(E)]f_0(E)\})dE\\ &= 2N(0)ev_F\mathscr{A}\int_{-\infty}^{\infty}[f_0(E-ev)-f_0(E)][1+A(E)-B(E)]dE\ . \end{aligned} \tag{17}$$

In obtaining the final form, we have used the properties that $A+B+C+D=1$, $A(E)=A(-E)$, and $f_0(-E)=1-f_0(E)$. The quantity $[1+A(E)-B(E)]$ in Eq. (17) can be referred to as the "transmission coefficient for electrical current." Its form shows that while ordinary reflection [described by $B(E)$] reduces the current, Andreev reflection [described by $A(E)$] increases it by giving up to two transferred electrons (a Cooper pair) for one incident one.

If both sides of the interface are normal metal, $A=0$ since there is no Andreev reflection, and

$1-B=C=(1+Z^2)^{-1}$ as noted above, so that Eq. (17) reduces to the simple form

$$I_{NN}=\frac{2N(0)e^2v_F\mathcal{A}}{1+Z^2}V\equiv\frac{V}{R_N}\ . \tag{18}$$

Note that even in the absence of a barrier ($Z=0$), there is still a nonzero normal-state resistance. This is the spreading resistance typical of a two-dimensional (2D) or 3D geometry, often referred to as the Sharvin resistance[14] in the present case of a point-contact geometry with infinite mean free path relative to orifice size.

For the more interesting *N-S* case, we have integrated (17) numerically to obtain *I-V* curves for a number of barrier strengths at $T=0$. The results are shown in Fig. 6. The effect of changing the barrier strength is even more dramatic in the differential conductance dI/dV, plotted versus V in Fig. 7, again at $T=0$. From Figs. 6 and 7, it is clear that barrier strengths $Z\gtrsim 10$ give results essentially indistinguishable from those for classical tunnel junctions, although, as our model demonstrates, there is a continuous variation from the metallic to the tunneling limit. To facilitate the interpretation of experimental dI/dV curves, it is useful to note at $T=0$, Eq. (17) implies that dI/dV is proportional to the transmission coefficient for electrical current, $1+A(eV)-B(eV)$, while at $T\neq 0$, dI/dV is proportional to a thermally smeared version of the same quantity.

V. RELATION TO TRANSFER HAMILTONIAN APPROACH

With the arbitrary barrier case well in hand, we are in a position to contrast our technique with the more conventional transfer Hamiltonian approach. If one assumes a thick oxide between two banks, to lowest order one can consider the two sides as noninteracting systems, containing electrons in standing-wave states labeled by quantum numbers k_L or k_R. Of course, they are not truly unconnected and there is a small probability that an electron from one side can tunnel to the other side; thus the wave function must actually penetrate through the barrier. Since this penetration will change the energy of the system, many workers have tried to model the resulting change by adding a "transfer" term to the Hamiltonian, although this phenomenological model has never been fully justified.[15] The tunneling transfer matrix element T is tricky to evaluate, since it is a function of the amplitude of the small tail on the penetrating wave and requires some microscopic knowledge of the system. Nonetheless, many workers have calculated T, and the result, for eV much smaller than the barrier height, is an essentially featureless matrix element much like that for an *N-N* contact.

By treating the tunneling as a perturbation, one can calculate the current by using the Golden Rule, and find the electron-transfer rate proportional to a squared matrix element times the number density of unoccupied final states. This is in the origin of the BCS density-of-states factor times the thermal weighting factor $[1-f(E-eV)]$

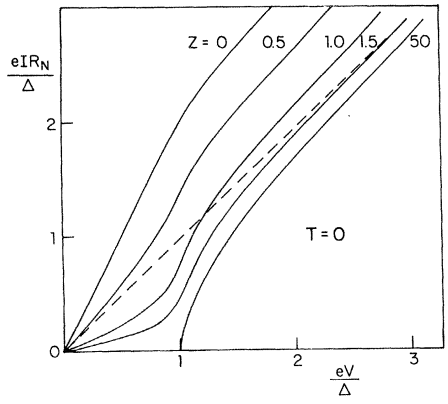

FIG. 6. Current vs voltage for various barrier strengths Z at $T=0$. These curves attain their asympotic limits only for very high voltages. For example, the tunnel junction ($Z=50$) curve will be within 1% of the normal-state curve (dotted line) only when $eV\geq 7\Delta$.

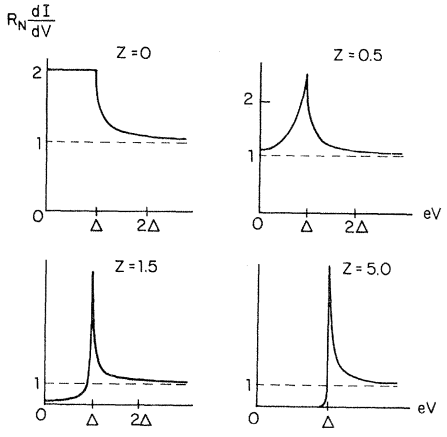

FIG. 7. Differential conductance vs voltage for various barrier strengths Z at $T=0$. This quantity is proportional to the transmission coefficient for electric current for particles at $E=eV$.

which appears in the tunneling current, for particles from L to R. When we integrate over all states on the left, an additional $N_L(E)f(E)$ appears, for a total probability (per unit energy) of

$$|T|^2 N_L(E)N_R(E-eV)f(E)[1-f(E-eV)] .$$

The latter two factors reflect the view that this is a stochastic process involving two independent banks, so that the joint probability for tunneling is proportional to the probability of the initial states being occupied times the probability of final states being unoccupied (i.e., unblocked). When we combine currents in both directions, the cross terms involving products of Fermi functions cancel, and we are left with the usual tunneling result in which the current is proportional to an integral over

$$N_L(E)N_R(E-eV)[f(E-eV)-f(E)] .$$

The Bogoliubov equation approach is quite different in spirit. As we saw above, the current transmission coefficient (in contrast to the matrix element squared in the tunnel calculation) varies rapidly for $|E| \approx \Delta$. This is the origin of the density-of-states factor in the final result. In no sense are the two banks decoupled; we associate a given set of wave vectors on the left $(q^+, -q^+, q^-)$ with a specific set on the right (k^+, k^-), regardless of the barrier height. q^+ is not an "initial" state from which transfer can be blocked by an occupied "final" state k^+. Rather, a wave packet describing an incident particle in q^+ evolves continuously into transmitted and reflected wave packets in the completely deterministic way described by Liouville's theorem without scattering. Thus, we associate one distribution function with this entire set of wave vectors. Combining currents originating from both sides gives a difference in Fermi functions just as the transfer Hamiltonian did (after cancellation of cross terms), but the rationale is entirely different in detail. Since we can achieve the same result with a model that has a larger range of validity and usefulness, we have chosen this latter point of view.

VI. EXCESS CURRENT

Although the high-voltage ($eV >> \Delta$) portion of each I-V characteristic in Fig. 6 is linear with slope R_N, it does not in general fall on the normal-state curve $V = IR_N$ (as it does in a tunnel junction). Rather, it is displaced by a constant amount referred to as the excess current, $I_{\rm exc}$, which can be found by extrapolation back to the $V=0$ axis. By manipulation of Eqs. (16) and (17), this $I_{\rm exc}$ can be written as

$$I_{\rm exc} \equiv (I_{NS} - I_{NN})\,|_{eV>>\Delta} = \frac{1}{eR_N[1-B(\infty)]} \times \int_0^\infty [A(E)-B(E)+B(\infty)]dE , \quad (19)$$

where from Table II, $B(\infty) = Z^2/(1+Z^2)$ is the reflection coefficient at high energy or in the normal state. Using this expression and those in Table II, we have computed numerically the dependence of $I_{\rm exc}R_N$ (the "insufficient voltage" referred to by Likharev[2]) upon barrier strength Z. Since this high-voltage limit depends on T only through the magnitude of Δ, the normalized quantity $I_{\rm exc}R_N/\Delta$ is a function only of barrier strength. This dependence is plotted in Fig. 8. Note the rapid falloff with increasing Z; clearly, excess current will be seen only in relatively barrier-free contacts.

More generally, one can define a voltage-dependent excess current,

$$I_{\rm exc}(V) = I_{NS}(V) - I_{NN}(V) , \quad (20)$$

which simply gives the extra current due to the superconductivity. Since $I_{NN}(V)$ is simply V/R_N, finding $I_{\rm exc}(V)$ is equivalent to giving the entire I-V curve, as was done in Fig. 6. However, it is worth noting that for $Z=0$ and $\Delta << kT$, the calculation can be carried out analytically, with the result,

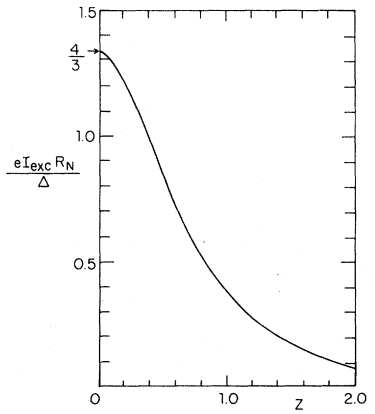

FIG. 8. Excess current (in units of Δ/eR_N) as a function of barrier strength Z. The temperature dependence of the curve is entirely contained within $\Delta(T)$.

$$I_{\rm exc}(V)=(4\Delta/3eR_N)\tanh(eV/2kT),\quad \Delta<<kT\ . \tag{21}$$

As noted in our previous work, this result agrees in detail with that obtained by Zaitsev[4] for a clean metallic *NS* contact, after correcting a factor of 2 in his result.

We can also calculate the excess current for an *S*-*S'* contact, by using the *N*-*S* result above. As we explained in a previous publication,[5] in the high-voltage limit, the *S*-*S'* device can be thought of as a series combination of an *S*-*N* and a *N*-*S'* microconstriction. The "insufficient voltages" of the two simply add, so that the excess current is the sum of the contributions from the two sides. For example, in the clean contact with $Z=0$, we find

$$I_{\rm exc}=4(\Delta+\Delta')/3eR_N\ .$$

While the excess current has a clear mathematical definition, it may be a difficult quantity to measure reliably, for at least two reasons. First, for voltages greater than the gap, heating will generally distort the *I*-*V* curve away from its ideal isothermal shape.[16] There may, in fact, be no linear region at high voltages until heating has driven the contact fully normal. Secondly, even in the absence of heating, care must be taken to carry the measurement to sufficiently high voltages to obtain a true asymptote. For example, the tunnel junction curve in Fig. 6 (i.e., $Z=50$) appears to have a significant negative (rather than zero) excess current, even at a voltage as high as $eV=3\Delta$. Of course, if one assumes our computed *I*-*V* curves are correct and applicable to the contact at hand, one can determine a value of Z by fitting the measured curves at lower voltages.

VII. ANALYSIS OF CURRENTS AND CHARGE-IMBALANCE GENERATION

In this section we establish a formalism for separating the total current I_{NS} into parts differentiated according to the mechanism of charge transfer. In particular, we distinguish the part of the current I_{NS}^* associated with the creation of quasiparticle charge imbalance Q^* in the superconductor from the part that is converted directly into supercurrent. We also define linear-response coefficients giving the normalized differential conductance Y and its counterpart Y^*, describing quasiparticle charge injection, and evaluate them as functions of Z and T.

For *tunnel* junctions, the rate of generation of quasiparticle charge branch imbalance has been discussed in some detail by Tinkham and Clarke[17] (TC), by Tinkham,[18] by Pethick and Smith[19] (PS), and by Clarke, *et al.*[20] (CESST). The latter authors reviewed carefully the interrelations of the definitions and conventions of the various treatments including the rather different conceptual approach of Schmid and Schön.[21] Using our present method, we can generalize the tunnel results to arbitrary barrier strength, providing a description which covers all cases from the tunnel limit to the ideal metallic contact, and which gives very detailed insight into the process.

To minimize repetition, we refer the reader to CESST for an index of the earlier work, and only summarize the essential points. Noting that quasiparticles at k have a charge

$$q_k=u_k^2-v_k^2=\epsilon_k/E_k=\pm N_s(E_k)^{-1}\ , \tag{22}$$

the total quasiparticle charge (per unit volume, in units of e) is

$$Q^*=\sum_k q_k f_k\ , \tag{23}$$

where f_k is the actual occupation number in a nonequilibrium system. To retain electrical neutrality, there must be an equal and opposite change in the condensate charge, produced by a shift in μ_s. [The quantity Q defined earlier by TC differs from Q^* by replacement of q_k in Eq. (23) by $\mathrm{sgn}q_k=\pm1$, and represents the *numerical* population imbalance between the branches. Since Q^* appears to have more general physical significance, we confine our attention to it in this paper.]

To illuminate this problem, it is helpful to rewrite the expression (17) for the current in a form emphasizing the transmitted (rather than reflected) particles by using the sum rule (13) to replace $1+A-B$ by $2A+C+D$. Thus, we write the total current as

$$I_{NS}=2N(0)ev_F\mathscr{A}\times\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[2A(E)+C(E)+D(E)]dE\ . \tag{24}$$

This can be decomposed into two parts by recognizing that the term $2A(E)$ represents the current

$$I_{NS}^{A}=2N(0)ev_F\mathscr{A}\times\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[2A(E)]dE\ , \tag{25}$$

due to Andreev reflection, in which the current in the superconductor is carried entirely by pairs. The remainder of I_{NS} is the current

$$I_{NS}^{qp}=2N(0)ev_F\mathscr{A}\times\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[C(E)+D(E)]dE\ , \tag{26}$$

due to quasiparticle transfer. But even I_{NS}^{qp} does not represent the rate of generation of *quasiparticle charge* at the interface, which we will call I_{NS}^{*}. To compute that, we must distinguish quasiparticles in the two channels, k^+ and k^-, and also take account of their fractional charge $q_k=\pm N_s(E)^{-1}$. Thus,

$$I_{NS}^{*}=2N(0)ev_F\mathscr{A}\times\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[C(E)-D(E)]N_s^{-1}(E)dE\ . \tag{27}$$

Finally, since experiments usually involve current bias rather than voltage bias, it has been useful to determine the ratio

$$F^{*}\equiv I^{*}/I\ , \tag{28}$$

giving the charge imbalance generated as a fraction of the injected current. These quantities have been evaluated in the past for tunnel junctions, but now we can evaluate them for arbitrary barrier strength, as a function of V and T.

To make contact with the earlier work on tunnel injection, we note that for strong barriers the results in Table II show that $A\propto Z^{-4}<<1$, so I_{NS}^{A} is negligible and $I_{NS}\approx I_{NS}^{qp}$ reproduces the form of the usual expression for tunnel current, namely,

$$I_{NS}^{qp}=\frac{1}{eR_N}\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times N_s(E)dE\ ,\quad Z^2>>1\ . \tag{29}$$

The density-of-states factor $N_s(E)$ is *not* inserted arbitrarily at this point; rather, it arises from the limiting form of $C+D$ given in Table II, namely $Z^{-2}(u^2-v^2)^{-1}$. As is well known, Eq. (29) shows no excess current; the Andreev term is essential for that, but it is important only for weak barriers. Again taking the strong barrier limit to evaluate I_{NS}^{*}, we note from Table II that $(C-D)=Z^{-2}$, so that Eq. (27) becomes

$$I_{NS}^{*}=\frac{1}{eR_N}\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times N_s^{-1}(E)dE,\ Z^2>>1 \tag{30}$$

again exactly equivalent to the standard tunneling result for Q_{inj}^{*}.

For comparison, we now consider the same quantities in the other limiting case of no barrier, $Z=0$. Then, as noted in Table II, $B=D=0$ and $C=1-A$. In this case, the total current is

$$I_{NS}=\frac{1}{eR_N}\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[1+A(E)]dE,\ Z=0\ . \tag{31}$$

The Andreev portion of the current I_{NS}^{A} is still given by Eq. (25), but I_{NS}^{qp} can be rewritten in terms of $A(E)$ as

$$I_{NS}^{qp}=\frac{1}{eR_N}\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[1-A(E)]dE,\ Z=0 \tag{32}$$

while

$$I_{NS}^{*}=\frac{1}{eR_N}\int_{-\infty}^{\infty}[f_0(E-eV)-f_0(E)]\times[1-A(E)]\times N_s^{-1}(E)dE,\ Z=0\ . \tag{33}$$

Note that in Eqs. (32) and (33) there is no contribution to the integrals for $|E|<\Delta$, where $1-A=0$; but in Eq. (31), $1+A=2A=2$ in the gap, where there is a large current derived entirely from the Andreev-reflection process. By contrast, in the strong-barrier limit, there is no current at all in the gap, all current being carried by quasiparticles above the gap. Since Q^{*} is generated only by quasiparticles, the ratio $F^{*}=I^{*}/I$ will in general be smaller for metallic contacts than for tunnel junctions.

In the important special case of very low voltages across the contact ($eV << kT$), we can replace the difference of Fermi functions in the expressions (24)–(27) and (29)–(33) by $eV(-\partial f_0/\partial E)$ and define generalized linear-response functions analogous to the Y and Z of CESST. These two functions are the ratios of I_{NS} and I^*_{NS}, respectively, to the current I_{NN} at the same small voltage. We now generalize these quatities to include a dependence on barrier strength as well. In doing so, we also change the notation from Y and Z to Y and Y^* to gain in suggestiveness and to avoid confusion with the use of Z for normalized barrier strength in this paper. Specifically, we define

$$Y(Z,T)=\left.\frac{I_{NS}(Z,T)}{I_{NN}(Z)}\right|_{eV\to 0}=(1+Z^2)\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right][2A(E)+C(E)+D(E)]dE\ , \tag{34}$$

and

$$Y^*(Z,T)=\left.\frac{I^*_{NS}(Z,T)}{I_{NN}(Z)}\right|_{eV\to 0}=(1+Z^2)\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right][C(E)-D(E)]N_s^{-1}(E)dE\ . \tag{35}$$

In the strong barrier limit, $A\sim Z^{-4}\sim 0$, and using the limiting forms of C and D from Table II as above, we obtain

$$Y(Z>>1,T)=\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right]N_s(E)dE\ , \tag{36}$$

and

$$Y^*(Z>>1,T)=\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right]N_s^{-1}(E)dE\ , \tag{37}$$

which are exactly the standard tunneling results given by CESST and others. In the zero barrier limit, $Z=0$, $B=D=0$, $C=1-A$, and we have

$$Y(Z=0,T)=\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right][1+A(E)]dE\ , \tag{38}$$

$$Y^*(Z=0,T)=\int_{-\infty}^{\infty}\left[-\frac{\partial f_0}{\partial E}\right][1-A(E)]\times N_s^{-1}(E)dE\ . \tag{39}$$

These coefficients Y and Y^*, and their ratio $F^*=Y^*/Y$, are plotted versus T for various values of Z in Figs. 9–11. Note that $Y^*(Z,T)$ is almost independent of the barrier strength Z.

VIII. APPLICATIONS AND DISCUSSION

A direct application of our results is to fit the temperature dependence of the differential conductance of a point contact at zero voltage, which should be simply $Y(Z,T)/R_N$. This quantity is easily measured, and, unlike the excess current, it is not subject to problems with heating effects.

Although the parameter Z plays a fundamental role in our theory, it cannot be independently determined, and must be inferred from the I-V curve. However, since both $I_{\rm exc}$ and $Y(Z,T)$ depend on Z, we can test the model by comparing these two experimentally accessible quantities against each other. Figure 12 shows $Y(Z,T=0)$ vs $eR_NI_{\rm exc}/\Delta$. They exhibit an approximately linear relationship to each other, as might be expected physically, since both depend linearly on the Andreev-reflection probability.

In addition to our theoretical curve, two other points are displayed in Fig. 12. The point marked Zaitsev-clean is the result found by Zaitsev[4] for a microconstriction with no scattering in the bulk or in the neck, and it agrees exactly with our calculation for $Z=0$. The point labeled AVZ-dirty is the the result of a calculation by Artemenko, Volkov, and Zaitsev[3] for a microconstriction in the dirty limit ($\tau kT << \hbar$) where the neck is short compared to the nonequilibrium length $\xi(T)(1-t)^{1/4}$. Although our model is not intended to describe diffusive transport in the presence of distributed bulk scattering, a suitable value ($Z\sim 0.55$) of the strength of our 1D δ-function barrier potential gives results very close to those obtained by their

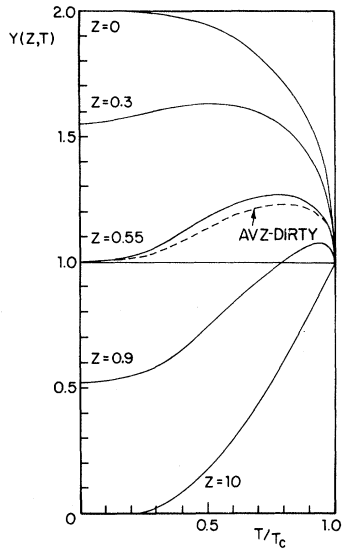

FIG. 9. Y, the zero-voltage differential conductance in units of the normal-state conductance, vs T, for various barrier strengths Z. The curve AVZ-dirty is described in the text.

Green's-function method.

Our charge-imbalance results may also provide a useful perspective on the work of Hsiang and Clarke[22] on resistance due to charge-imbalance relaxation near an N-S interface. By a physical argument, they reasoned that the appropriate F^* for a clean *metallic* interface could be approximated by what we call here $Y^*(Z >> 1)$ as computed for a *tunnel* junction. Our calculation would instead give $F^* = Y^*(Z=0)/Y(Z=0)$. As just noted above, Y^* is almost independent of barrier strength Z, so $Y^*(Z >> 1) \approx Y^*(Z=0)$. But the denominator factor $Y(Z=0)$ varies from 1 at T_c to 2 at $T=0$, so the resulting T dependence of our F^* is somewhat different from that used by Hsiang and Clarke, although it has the same qualitative shape. Because the interfaces studied by Hsiang and Clarke were not point contacts of the sort we have been treating, however, it is not clear that our results should be expected to describe their data in detail.

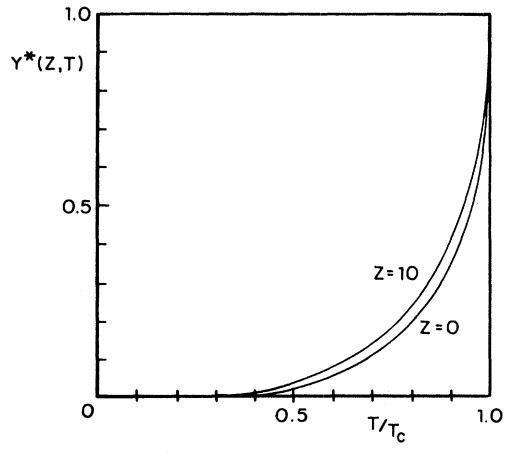

FIG. 10. Y^*, the normalized differential charge imbalance conductance at zero voltage, vs T, for various barrier strengths Z.

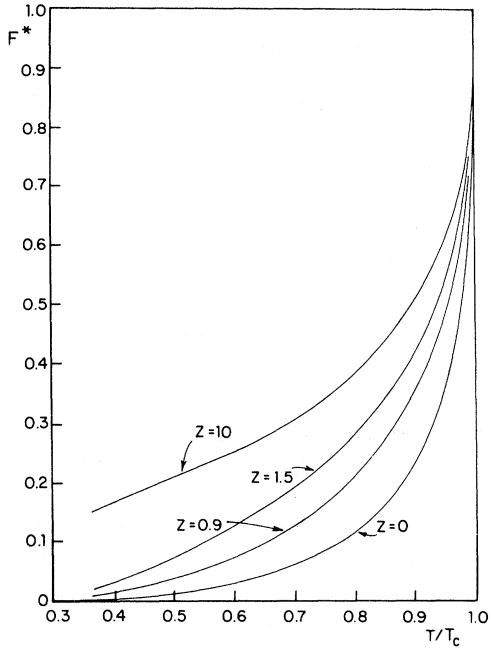

FIG. 11. F^*, the ratio of quasiparticle charge injected to total injected current, as a function of T for various barrier strengths Z.

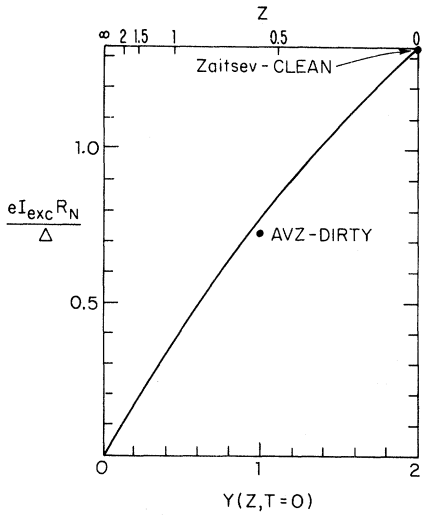

FIG. 12. $eI_{exc}R_N/\Delta$ vs $Y(T=0)$ as the barrier strength parameter Z ranges from ∞ to 0. The points AVZ-dirty and Zaitsev-clean are discussed in Sec. VIII.

As noted above, the total current I_{NS} can be expressed as the sum of three terms: I_{NS}^{A}, the Andreev term with no quasiparticle transfer; I_{NS}^{*}, the injection rate of quasiparticle charge; and $(I_{NS}^{\rm qp} - I_{NS}^{*})$, the current carried by the pairs that must flow to compensate for the fractional charges of the quasiparticles. The current I_{NS}^{*} relaxes to supercurrent over a distance $\Lambda_{Q^*}=(D\tau_{Q^*})^{1/2}$, a diffusion length determined by τ_{Q^*}, the charge-imbalance relaxation time.

The picture described in the previous paragraph corresponds to the viewpoint of TC and PS, in which the current $(I-I^*)$ is viewed as being converted discontinuously to supercurrent at the interface.[23] If one examines the problem more deeply, however, one can obtain a more microscopic view of the conversion process. The key point is that there are solutions of the Bogoliubov equations even for $|E|<\Delta$, but they are evanescent waves which decay in a distance

$$\hbar v_F/2(\Delta^2-E^2)^{1/2}\sim\xi(T)\ .$$

(See the Appendix for details.) Thus, even the Andreev current I_{NS}^{A} is carried for $\sim\xi$ as a quasiparticle current before decaying into a pair current, and the interference of k^+ and k^- waves causes a similar effect also for the current $I_{NS}^{\rm qp}-I_{NS}^{*}$ above the gap. This two-step conversion process is shown schematically in Fig. 13. This point of view is reminiscent of that of Schmid and Schön, who treat the entire injected current as a source of charge imbalance, but combine "conversion" terms with τ_{Q^*} relaxation to find a shorter effective relaxation time $\tau_R<\tau_{Q^*}$, which reproduces the same steady-state Q^* as found by the two-step analysis of Tinkham and Clarke[17,18] and of Pethick and Smith.[19] There seems to be a gain in physical insight, however, in distinguishing the two types of processes.

Recognition of the existence of a penetration depth of order ξ for the evanescent quasiparticle waves in the gap also provides a more clear-cut explanation for one aspect of the observations of Octavio *et al.*[16] on subharmonic gap structure (SGS) in tin variable-thickness bridges (VTB's). They observed voltage-dependent shifts in the positions of the SGS which they attributed to heating. Using the gap shift as a thermometer, with $\Delta_{\rm BCS}(T)$ as the calibration curve, they could infer temperature rises. These temperature increases were roughly proportional to the dissipated power P, but much

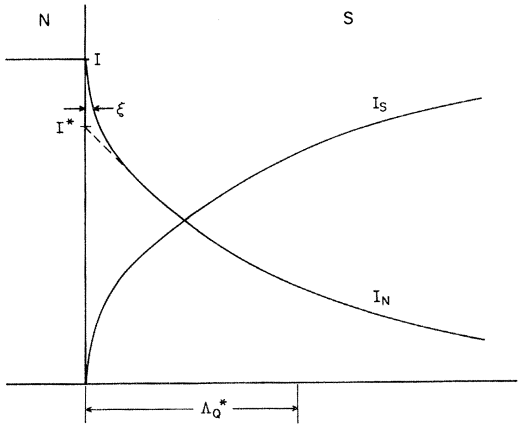

FIG. 13. Schematic diagram of conversion of normal current to supercurrent at N-S interface. The evanescent waves due to Andreev reflection die out over a distance of order ξ, while the injected quasiparticle charge imbalance relaxes over a charge diffusion length Λ_{Q^*}.

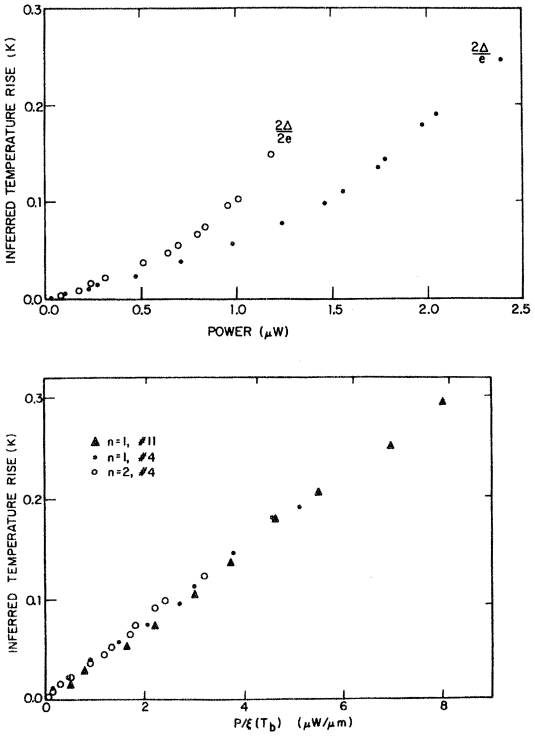

FIG. 14. Temperature rise in tin variable-thickness bridge inferred from depression of energy gap as found from subharmonic gap structure (after Octavio *et al.*). (a) depicts temperature rise vs power dissipated, P. (b) depicts temperature rise vs $P/\xi(T)$. The superior linear fit in (b) is consistent with observed gap being sensed at a distance $\sim\xi$ from center of bridge.

more precisely proportional to $P/\xi(T)$, as shown by the comparison of Fig. 14(a) and 14(b). Given the 3D cooling of these VTB's, in which the temperature rise is proportional to P/r (where r is the distance from the center of the contact), this observation was interpreted as showing that the energy gap giving the SGS was that in the superconductor at a distance $\sim\xi$ from the center. In a recent paper,[5] we have argued that the SGS arises from multiple Andreev reflections together with the sharp change in $A(E)$ at the gap edge. Given the penetration of the evanescent waves to a depth $\sim\xi(T)$, it is plausible that the gap reflected in the SGS should be that at a similar depth into the bank, as assumed by Octavio *et al.* on the basis of somewhat different reasoning.

IX. CONCLUSION

We have proposed a simple theory for the *N-S* microconstriction or point contact that is general enough to describe the complete crossover from metallic to small-area tunnel-junction behavior. From this model, a deeper understanding of the semiconductor picture, excess currents, current conversion, and charge-imbalance processes have emerged, in addition to calculation of *I-V* curves for arbitrary barrier strength. The latter results are directly accessible to experiment, and should be helpful in interpreting the wide variety of *I-V* curve shapes observed in the laboratory. The application of this model to the *S-N-S* case has been presented elsewhere,[5] where it has been shown to offer an explanation of the SGS observed in metallic weak links.

ACKNOWLEDGMENTS

One of us (M.T.) would like to acknowledge stimulating discussions at the initial stage of this work with J. Clarke, U. Eckern, A. Schmid, G. Schön, and A. F. Volkov. The work was supported in part by the National Science Foundation, the Office of Naval Research, and the Joint Services Electronics Program. One of us (T.M.K.) would like to thank the Nederlandse Organisatie voor Zuiver Wetenshappelijk Onderzoek for a grant.

APPENDIX

The Bogoliubov equations may be written

$$i\hbar\frac{\partial f}{\partial t}=\left[-\frac{\hbar^2\nabla^2}{2m}-\mu(x)+V(x)\right]f(x,t)+\Delta(x)g(x,t)\,, \tag{A1a}$$

$$i\hbar\frac{\partial g}{\partial t}=-\left[\frac{-\hbar^2\nabla^2}{2m}-\mu(x)+V(x)\right]g(x,t)+\Delta(x)f(x,t)\,, \tag{A1b}$$

where $\Delta(x)$ is the energy gap and $\mu(x)$ is the chemical potential (specifically that of the condensate, in case of nonequilibrium conditions where μ is not uniquely defined). In the normal metal [$\Delta(x)=0$], Eq. (A1a) is the Schrödinger equation for electrons, while Eq. (A1b) is the time-reversed Schrödinger equation for electrons. Since an electron satisfying the time-reversed Schrödinger equation behaves in many ways like a hole, we will adopt that terminology here.

For $\Delta\neq 0$, the electron and hole wave functions couple together, with two major consequences. First, a gap appears in the E vs k relation, just as it does in elementary BCS theory. To see this result more clearly, we solve Eq. (A1) for a specific geometry. The simplest example is one where $\mu(x)$, $\Delta(x)$, and $V(x)$ are all constant, and the solutions to Eq. (A1) are time-independent plane waves. Taking as our trial solutions $f=\tilde{u}e^{ikx-iEt/\hbar}$ and $g=\tilde{v}e^{ikx-iEt/\hbar}$, we find for $V=0$

$$E\tilde{u}=\left[\frac{\hbar^2k^2}{2m}-\mu\right]\tilde{u}+\Delta\tilde{v}\,, \tag{A2a}$$

$$E\tilde{v}=-\left[\frac{\hbar^2k^2}{2m}-\mu\right]\tilde{v}+\Delta\tilde{u}\,. \tag{A2b}$$

Solving Eqs. (A2a) and (A2b) for E we obtain

$$E^2=\left[\frac{\hbar^2k^2}{2m}-\mu\right]^2+\Delta^2\,. \tag{A3}$$

There are two roots for E. We are concerned with excitations above the ground state ($E\geq 0$), and in this case

$$\tilde{u}^2=\frac{1}{2}\left[1\pm\frac{(E^2-\Delta^2)^{1/2}}{E}\right]=1-\tilde{v}^2\,, \tag{A4}$$

$$\hbar k^{\pm}=\sqrt{2m}\,[\mu\pm(E^2-\Delta^2)^{1/2}]^{1/2}\,. \tag{A5}$$

We also define

$$1-v_0^2=u_0^2=\frac{1}{2}\left[1+\frac{(E^2-\Delta^2)^{1/2}}{E}\right].$$

For $|E|<\Delta$, $\tilde{u}$ and $\tilde{v}$ will be complex conjugates. This definition differs from the usual BCS convention in which u_k and v_k are not defined for $E<\Delta$. Using the notation

$$\psi=\begin{bmatrix}f(x)\\g(x)\end{bmatrix},$$

the four types of quasiparticle waves for a given energy take the forms

$$\psi_{\pm k^+}=\begin{bmatrix}u_0\\v_0\end{bmatrix}e^{\pm ik^+x} \tag{A6a}$$

and

$$\psi_{\pm k^-}=\begin{bmatrix}v_0\\u_0\end{bmatrix}e^{\pm ik^-x}. \tag{A6b}$$

The ψ's are the position space representation of the $\gamma_{k0}^*=u_k c_{k\uparrow}^*-v_k c_{-k\downarrow}$ since $c_{k\uparrow}^*$ produces the wave e^{ikx}, while $c_{-k\downarrow}$ destroys an electron at $-k$, leaving the system with a net positive momentum represented by the creation of a hole wave e^{ikx}.

A second consequence, which follows from the $E(k)$ relation (A3), is that the group velocity $v_g=dE/d\hbar k$ goes to zero at the gap edge, even while the phase velocity remains near the normal-state value. This has some interesting consequences, as can best be illustrated by deriving conservation laws for probability and charge.

We define $P(x)$ to be the probability density for finding either an electron or a hole at a particular time and place. Thus, for a solution to Eq. (A1), $P(x,t)=|f|^2+|g|^2$. Using Eq. (A1) to evaluate $\partial P/\partial t$, we find

$$\frac{\partial P}{\partial t}+\vec{\nabla}\cdot\vec{J}_P=0, \tag{A7}$$

where

$$\vec{J}_P=\frac{\hbar}{m}[\mathrm{Im}(f^*\nabla f)-\mathrm{Im}(g^*\nabla g)]. \tag{A8}$$

Note that the hole current enters Eq. (A8) with a sign opposite to that of the electron part, just as one expects for a "time-reversed" particle. One can also show that $J_P\propto v_g=\partial E/\partial\hbar k$. Thus, at the gap edge, it is zero.

We can also derive a conservation law for quasiparticle *charge*. Assigning a unit charge $+e$ to the electron and $-e$ to the hole, the net quasiparticle charge density in one of these excitation waves is $Q=e(|f|^2-|g|^2)$. Using Eq. (A1), it is easy to show that

$$\frac{\partial Q}{\partial t}+\vec{\nabla}\cdot\vec{J}_Q=\frac{4e\Delta}{\hbar}\mathrm{Im}(f^*g), \tag{A9}$$

where

$$\vec{J}_Q=\frac{e\hbar}{m}[\mathrm{Im}(f^*\nabla f)+\mathrm{Im}(g^*\nabla g)]. \tag{A10}$$

The term on the right in Eq. (A9) is a source (or drain) term connecting the quasiparticles with the condensate, and we will see an explicit interchange of currents when treating the *N-S* case below. However, for now we concentrate on Eq. (A10). Both the electron and hole contribution to the quasiparticle current enter with the same sign, and a moment's thought shows this to be reasonable. An electron ($f\propto e^{ik^+x}$), moves in the positive direction and carries a positive current. A hole, ($g\propto e^{ik^+x}$), moves in the negative direction, but also carries a positive current due to the sign of its charge. Thus, while $J_P\propto v_g=0$ at the gap edge, $J_Q=ev_F$, so one can view the charge current as traveling with the Fermi velocity. Of course, for the normal metal nothing so dramatic occurs, and we always find simply $J_Q=eJ_P$.

Starting from the basic Eq. (A1), it is a simple matter to work out the boundary conditions on steady-state plane-wave solutions at an *N-S* interface. To model the elastic scattering that usually occurs in the orifice, we include a δ-function potential in Eq. (A1), i.e., $V(x)=H\delta(x)$.

The appropriate boundary conditions, for particles traveling from *N* to *S* are as follows.

(i) Continuity of ψ at $x=0$, so $\psi_S(0)=\psi_N(0)\equiv\psi(0)$.

(ii) $(\hbar/2m)(\psi_S'-\psi_N')=H\psi(0)$, the derivative boundary condition appropriate for δ functions.

(iii) Incoming (incident), reflected, and transmitted wave directions are defined by their group velocities. We assume the incoming electron produces only outgoing particles.

Thus,

$$\psi_{\mathrm{inc}}=\begin{bmatrix}1\\0\end{bmatrix}e^{iq^+x},\quad \hbar q^{\pm}=\sqrt{2m}\sqrt{\mu\pm E},$$

$$\psi_{\mathrm{refl}}=a\begin{bmatrix}0\\1\end{bmatrix}e^{iq^-x}+b\begin{bmatrix}1\\0\end{bmatrix}e^{-iq^+x},$$

$$\psi_{\mathrm{trans}}=c\begin{bmatrix}u_0\\v_0\end{bmatrix}e^{ik^+x}+d\begin{bmatrix}v_0\\u_0\end{bmatrix}e^{-ik^-x}.$$

Applying the boundary conditions, and carrying out an algebraic reduction, we find

$$a=\frac{u_0 v_0}{\gamma}\ , \tag{A11a}$$

$$b=-\frac{(u_0^2-v_0^2)(Z^2+iZ)}{\gamma}\ , \tag{A11b}$$

$$c=\frac{u_0(1-iZ)}{\gamma}\ , \tag{A11c}$$

$$d=\frac{iv_0 Z}{\gamma}\ , \tag{A11d}$$

where

$$Z=\frac{mH}{\hbar^2 k_F}=H/\hbar v_F\ ,$$

and

$$\gamma=u_0^2+(u_0^2-v_0^2)Z^2\ .$$

For simplicity, we have let $k^+=k^-=q^+=q^-=k_F$, whenever that substitution did not lead to a qualitative charge. Note that in the absence of a barrier (i.e., $Z=0$), $b=d=0$. Physically speaking, this means that all reflection is Andreev reflection and all transmission occurs without branch crossing.

The quantities $A(E)$, $B(E)$, $C(E)$, and $D(E)$ used in the main body of the paper are actually the probability currents, for the particle, measured in units of v_F. For example, $A=J_P^A/v_F=a^*a$, and $D=d^*d(u_0^2-v_0^2)$. Since plane-wave currents are spatially uniform, we need not specify the position at which they are evaluated. However, for $E<\Delta$, k^+ and k^- in the superconductor have small imaginary parts which lead to an exponential decay on a length scale λ, where

$$\lambda=\frac{\hbar v_F}{2\Delta}\left[1-\left[\frac{E}{\Delta}\right]^2\right]^{1/2}\ . \tag{A12}$$

Although right at the gap edge the length diverges, the characteristic length is $\hbar v_F/2\Delta=1.22\xi(T)$. Thus, in order of magnitude terms, one can say that the particles penetrate a depth $\sim\xi(T)$ before the current is converted to a supercurrent carried by the condensate. For clarity, we have defined $C(E)$ and $D(E)$ as the probability currents at $x>>\lambda$, measured in units of v_F. The changeover from full to zero current is described next.

As a simple example of quasiparticle current being converted to condensate current, we consider the case $H=0$. Then, $b=d=0$, $a=u_0/v_0$, and $c=1/u_0$. For $E<\Delta$, $a^*a=1$, which means the incident electron is totally reflected as a hole. Thus, the total charge carried in the normal metal equals $2ev_F$, but in the superconductor J_Q is exponentially small for $x>>0$. Explicitly,

$$J_Q^C=(e\hbar/m)\mathrm{Im}[(ce^{ik^+x})^*\nabla(ce^{ik^+x})\times(u_0^*u_0+v_0^*v_0)]\ .$$

Letting $k^+\approx k_F+i\eta/\hbar v_F$, where $\eta=(\Delta^2-E^2)^{1/2}$, we have

$$J_Q^C=\frac{|u_0|^2+|v_0|^2}{|u_0|^2}ev_F e^{-2\eta x/\hbar v_F}\ .$$

Below the gap, $|u_0|^2=|v_0|^2$, so

$$J_Q^C=2ev_F e^{-2\eta x/\hbar v_F}\ .$$

The "disappearing current" reappears as current carried by the condensate. Rewriting the drain term in Eq. (A9) as

$$-\vec{\nabla}\cdot\vec{J}_s\equiv(4e\Delta/\hbar)\mathrm{Im}(f^*g)\ , \tag{A13}$$

then

$$\begin{aligned}J_s&=\frac{4e\Delta}{\hbar}\int_0^x\mathrm{Im}(c^*v_0^*e^{-ik^+x'^*}cu_0e^{ik^+x'})dx'\\&=\frac{4e\Delta}{\hbar}\mathrm{Im}\left[\frac{v_0^*u_0}{u_0^*u_0}\right]\int_0^x e^{-2\eta x'/\hbar v_F}dx'\\&=2ev_F(1-e^{-2\eta x/\hbar v_F})\ .\end{aligned} \tag{A14}$$

This is the desired result, explicitly showing the supercurrent increasing to an asymptotic value as $x\to\infty$, at the same rate as the quasiparticle current dies away.

---

*Permanent address: Laboratorium voor Technische Natuurkunde, Technische Hogeschool Delft, Delft, The Netherlands.

[1]See, for example, W. L. McMillan and J. M. Rowell, in *Superconductivity,* edited by R. D. Parks (Marcel Dekker, New York, 1969), Vol. 1, p. 561.

[2]K. K. Likharev and L. A. Yakobson, Zh. Eksp. Teor. Fiz. 68, 1150 (1975) [Sov. Phys.—JETP 41, 2301 (1975)].

[3]S. N. Artemenko, A. F. Volkov, and A. V. Zaitsev, Pis'ma Zh. Eksp. Teor. Fiz. 28, 637 (1978) [JETP Lett. 28, 589 (1978)]; Zh. Eksp. Teor. Fiz. 76, 1816

(1979) [Sov. Phys.—JETP 49, 924 (1979)]; Solid State Commun. 30, 771 (1979).

[4]A. V. Zaitsev, Zh. Eksp. Teor. Fiz. 78, 221 (1980) [Sov. Phys.—JETP 51, 111 (1980)].

[5]T. M. Klapwijk, G. E. Blonder, and M. Tinkham, in Proceedings of the 16th International Conference on Low Temperature Physics, [Physica B + C (in press)].

[6]P. G. DeGennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966), Chap. 5.

[7]See, for example, M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1975), Chap. 2.

[8]A. F. Andreev, Zh. Eksp. Teor. Fiz. 46, 1823 (1964) [Sov. Phys.—JETP 19, 1228 (1964)].

[9]R. Kümmel, Z. Phys. 218, 472 (1969).

[10]J. Demers and A. Griffin, Can. J. Phys. 49, 285 (1970).

[11]A. Griffin and J. Demers, Phys. Rev. B 4, 2202 (1971).

[12]O. Entin-Wohlman, J. Low Temp. Phys. 27, 777 (1977).

[13]Scattering in the bulk near an *S-N* boundary introduces some interesting complications based on the coherence on the Andreev reflection process. These have been discussed qualitatively by A. B. Pippard, in *Nonequilibrium Superconductivity, Phonons, and Kapitza Boundaries,* edited by K. Gray (Plenum, New York, 1981), Chap. 12, and references therein.

[14]Yu. V. Sharvin, Zh. Eksp. Teor. Fiz. 48, 984 (1965) [Sov. Phys.—JETP 21, 655 (1965)].

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

[16]M. Octavio, W. J. Skocpol, and M. Tinkham, IEEE Trans. Magn. MAG-13, 739 (1977); M. Octavio, Technical Report No. 13, Div. of Appl. Sciences, Harvard University (unpublished).

[17]M. Tinkham and J. Clarke, Phys. Rev. Lett. 28, 1366 (1972).

[18]M. Tinkham, Phys. Rev. B 6, 1747 (1972).

[19]C. J. Pethick and H. Smith, Ann. Phys. (N.Y.) 119, 133 (1979).

[20]J. Clarke, U. Eckern, A. Schmid, G. Schön, and M. Tinkham, Phys. Rev. B 20, 3933 (1979).

[21]A. Schmid and G. Schön, J. Low Temp. Phys. 20, 207 (1975).

[22]T. Y. Hsiang and J. Clarke, Phys. Rev. B 21, 945 (1980).

[23]In the notation of Pethick and Smith, $\vec{J}_n^Q$ corresponds to our I^*, and their $\vec{J}_s^Q$ corresponds to our I-I^*. Similarly, their $\vec{J}_n$ corresponds to our J_Q as defined in the Appendix.