Nonequilibrium effects and subgap structure in superconducting contacts

M. A. Peshkin and R. A. Buhrman

School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853

(Received 25 October 1982)

We calculate the current-voltage characteristics of a superconductor—normal-metal—superconductor junction, including the effects of multiple Andreev reflections and of a nonequilibrium quasiparticle distribution in the superconducting regions adjacent to the normal-superconductor interfaces. We find that only ideal point contacts can nonequilibrium effects be neglected. Model calculations are performed in two limiting cases in which the nonequilibrium quasiparticle distributions are readily determined. The results are in good qualitative agreement with experimental observations of subgap structure.

I. INTRODUCTION

Subharmonic gap structure is often observed in the current-voltage ($IV$) characteristics of superconducting microbridges and nonideal tunnel junctions. It is characterized by sudden changes in resistance ($dV/ddI$) occurring at voltages $V = 2\Delta/n$, where $\Delta$ is the gap energy and $n$ is an integer. The search for a satisfactory explanation of subgap structure has a long history.1,2 Most recently Klapwijk, Blonder, and Tinkham (KBT) made a major advance in this area by suggesting that subgap structure is the result of multiple Andreev reflections between the two normal-metal—superconductor ($NS$) interfaces of a superconductor—normal-metal—superconductor ($SNS$) microbridge.3 For the first time a single mechanism has been proposed which can account for both even- and odd-order subharmonics, and which does not predict the higher-order subharmonics to be greatly smaller in amplitude than the lower-order ones. Although this model is strictly applicable only to an $SNS$ microbridge, the calculations may also be applied to superconducting constrictions if the constriction, driven normal by a current, can be considered to be an $SNS$ device. “Leaky” tunnel junction can be modeled as a constriction in parallel with an ideal tunnel junction.

However, careful analysis of the KBT approach reveals several difficulties, as shown below. These include the disappearance of the effect at low temperatures, and the prediction of structures less sharply defined and of smaller amplitude than those generally observed.

We show that these deficiencies can be fully corrected if the KBT approach is extended to include two fundamental effects: the generation of a nonequilibrium electron distribution in the normal-metal microbridge, and, most importantly, the generation of a nonequilibrium quasiparticle distribution (charge imbalance) in the superconducting electrodes. When these effects are included we calculate subgap structure in good qualitative agreement with experimental results.

It has long been known that nonequilibrium effects should be important in microbridges,4,5 so their inclusion in calculations of subgap structure is clearly warranted. Here we show that by treating microbridges as $SNS$ devices we can include nonequilibrium effects in a relatively straightforward and physically apparent manner.

In Sec. II we review the analysis of an $NS$ junction given by KBT. This will familiarize the reader with the phenomenon of Andreev reflection.

In Sec. III we introduce an occupation-of-states factor into the calculation of the $IV$ characteristic of an $NS$ junction, and give a physical justification for it. We show that this modification has no effect on the total current calculated.

In Sec. IV we discuss the generation of charge imbalance at an $NS$ interface and the effect this charge imbalance has on the resistance. We conclude that charge imbalance cannot be neglected unless the contact radius is smaller than the elastic scattering length $l$. In Sec. V we review the KBT analysis of an $SNS$ junction. In Secs. VI and VII we calculate the current through an $SNS$ junction. Our assumptions are as follows:

1. Andreev reflection, a two-particle process, occurs with a probability proportional to the product of the probability of occupation of the two energy levels involved.
2. The probability of Andreev reflection is zero for particles incident on a superconductor outside the gap.
3. There is no potential drop across the normal metal itself. The potential drop occurs solely at the two $NS$ interfaces. Further, we assume that the electron distribution in the normal metal is uniform across the metal.
4. The effects of a nonequilibrium quasiparticle distribution and of ohmic heating on $\Delta$ are ignored. We assume that $\Delta$ does not vary spatially within the superconductors.
5. We consider the limits $a \gg \lambda_{in}$ and $a \ll \lambda_{in}$, where $a$ is the contact radius, and $\lambda_{in}$ is the length scale on which quasiparticles thermalize.

In Sec. VIII we compare our results with experimental data.

II. KBT ANALYSIS OF THE NS INTERFACE

To understand the behavior of an $SNS$ junction it is necessary first to examine the properties of a single $NS$ interface. When an electron is incident on a superconductor from a normal metal, it can be transmitted only if its energy $E$ (measured relative to the pair chemical potential $\mu_S$)
outside the gap, and unity inside the gap.

Current can be measured at any point in the system, so KBT chose to measure it just inside the normal metal. To avoid double counting, we consider only the electron contribution to the current, ignoring that of the holes. A voltage $V$ is applied, raising the Fermi level on the left to an energy $eV$ with respect to the Fermi level on the right.

Four terms then contribute to the electron current. They are shown schematically in Fig. 2. The terms are

$$\int f_0(E-eV)A_{NS}(E)dE,$$
$$\int f_0(E-eV)[1-A_{NS}(E)]dE,$$
$$-\int [1-f_0(E-eV)A_{NS}(E)]dE,$$
$$-\int f_0(E)N_S(E)T_{SN}(E)dE,$$

where $f_0$ is the equilibrium Fermi distribution, $A_{NS}$ is the Andreev-reflection coefficient, and $T_{SN}$ is the superconductor-normal-metal transmission coefficient. $N_S$ is the superconductor density of states, divided by the density of states in a normal metal. For simplicity we take the single-electron density of states to be the same in the normal metal and the superconductor, and independent of energy. We have left out several constant prefactors needed to form true currents throughout this work.

Term (2) is the current due to electrons from $N$ which are Andreev-reflected at the boundary. Term (3) is the current due to electrons from $N$ which are transmitted through the boundary. Term (4) is the current due to holes incident from $N$ which Andreev reflect as electrons at the boundary and thus reduce the total (left to right) current. Term (5) is the current due to electrons from $S$ which are transmitted into $N$.

Detailed balancing requires that

$$N_S(E)T_{SN}(E)=1-A_{NS}(E),$$

with the use of this equation, term (5) can be rewritten

FIG. 2. Four processes at an $NS$ interface. The measurement plane is just to the left of the interface. We consider both incident electrons and holes, but count only electrons that cross the measurement plane.
- \int f_0(E) [1 - A_{NS}(E)] dE \ .

When terms (2)–(4) and (7) are summed and the integral is performed, the results shown in Fig. 3 (curve A) are obtained. These results are in agreement with those of Zaitsev,\textsuperscript{7} calculated by a different method.

III. ANDREEV REFLECTION AS A TWO-ELECTRON PROCESS

We begin by modifying the above (KBT) argument for the SN interface as follows: An incident electron can be reflected into the normal metal as a hole only if the hole state at the required energy is unoccupied. Equivalently, the electron state at that energy must be occupied. If the occupation requirement is not satisfied, an electron cannot Andreev-reflect as a hole, and must instead simply reflect as an electron at the boundary. We assume that the transmission probability \( [1 - A_{NS}(E)] \) is unchanged, but that the probabilities of Andreev and simple reflections are \( A_{NS}(E) f_0(-E) \) and \( A_{NS}(E) [1 - f_0(-E)] \), respectively.

With these modifications, terms (2) and (4) become

\[
\int f_0(E - eV) A_{NS}(E) f_0(-E - eV) dE
\]

and

\[
- \int [1 - f_0(E - eV)] A_{NS}(E) [1 - f_0(-E - eV)] dE \ .
\]

Some consideration will show that the total current contribution from the integrands of terms (8) and (9), for any energy \( E \), is the same as that from the integrands of terms (2) and (4). Care must be taken because processes (4) and (9) contribute their electron current at energy \( -E \). (The equality reflects the symmetry of the Fermi distribution about the chemical potential.) The modification of (2)–(5) to take into account occupation of states thus makes no change in the current-voltage characteristic predicted for an NS junction. Our result is in accord with those of Zaitsev and KBT.

IV. CHARGE IMBALANCE

Consider now a one-dimensional (planar sandwich) SNS junction with an applied voltage \( 2eV > \Delta \), such that a potential difference \( V \) is present between the normal metal and the superconductor on the right. The entire potential may appear across the interface in this way only when the thickness of the normal metal \( (d) \) is small compared to the elastic scattering length \( l \). Otherwise there would be a potential gradient across the normal metal due to scattering. If \( d \ll l \), the current through the normal metal is not the usual one based on bulk conductivity, namely

\[
I/V = (Ne^2\tau_{el}/m)(A/d) \ ,
\]

where \( N \) is the density of electrons, and \( A \) is the area of the interface. Following Sharvin,\textsuperscript{8} we note that no matter how small \( d \) may be, no more electrons may pass through a given area than are incident upon it. It is the number of incident electrons times their charge which is the total forward current \( I_f \). In this kinetic model there may also be reverse current \( I_r \). Thus, for a forward voltage bias, the net current will be given by \( (I_f - I_r) \) and not by Eq. (10).

For the remainder of this section we will consider quasiparticle current only, neglecting the pair current caused by Andreev reflection. This is adequate to calculate the dynamic resistance \( (dI/dV) \) measured at \( eV > 2\Delta \). We will also make the assumption that the density of states is energy independent. To calculate the quasiparticle current through the interface at \( T = 0 \), we need only consider those electron energies which are below the Fermi level of the normal metal, but above the upper gap edge of the superconductor \( (\mu_{S} + \Delta) \). \( I_f \) and \( I_r \) will cancel exactly for all energies below this band, leaving no net current flow. There are no occupied states above the band (Fig. 4). The band is of height \( eV - \Delta \) is thermal equilibrium \( I_r = 0 \), and the band is unoccupied in the superconductor. Approximating the density of states as \( N/mV_f^2 \), we have

\[
I_f = e(N/mV_f)A(eV - \Delta) = (Ne^2/\tau_{el})(A/l)(eV - \Delta) \ ,
\]

which has the same form as Eq. (10) except that the mean free path \( l \) has replaced the actual thickness \( d \). This leads us to define

\[
R_{\text{min}} = (mV_f/Ne^2A) \ .
\]

If elastic scattering in the normal metal is not entirely negligible \( (l - d) \), the actual resistance of the normal metal \( R_N \) to first order is

\[
R_N = R_{\text{min}}(1 + d/l) \ .
\]

All of \( I_f \) as defined above (that is, excluding pairs) contributes to a nonequilibrium accumulation of charge imbalance density \( Q^{\ast} \) in the superconductor.\textsuperscript{9} If the relaxation length for charge imbalance is \( \lambda_Q \), in a one-dimensional system we have a total charge imbalance of

---

\[
\begin{array}{c}
\begin{aligned}
&\begin{array}{c}
\text{FIG. 4. At } T = 0, \text{ the energies which contribute a net current at an NS interface are shown here.}
\end{array}
\end{aligned}
\end{array}
\]

---
\[
\int Q^*(v) dv = Q^*(0) \lambda Q A
\]  

(14)
in the superconductor. We may neglect inelastic scattering, since we are interested only in total charge imbalance and not in its distribution in energy. \(Q^*(0)\) is the charge-imbalance density at the interface \((\nu = 0)\). If the relaxation time of the charge imbalance is \(\tau_Q \gg 1/\nu_F\), equilibrium between the injected current and the decay current will occur when the net current flow into the superconductor satisfies

\[
\tau_Q I_r = \tau_Q (I_r - I_i) = Q^*(0) \lambda Q A .
\]

(15)

We have left out the (energy-dependent) effective charge from the left side of Eq. (15). Its effect will be discussed later. If \(Q^*(0)\) is nonzero, it will create a reverse current

\[
I_r = Q^*(0)A\nu_F/2 ,
\]

(16)
which reduces the net (total) current \(I_i\). Solving for \(I_i\), we find

\[
I_i(1 + \tau_Q/2\nu_F) = (NeA/m_F)(eV - \Delta) ,
\]

(17)
which means that the actual resistance of the NS interface is given by \(R_{\text{min}}(1 + \rho_Q)\), with

\[
\rho_Q = I_r/I_i = \tau_Q \nu_F/2\nu Q ,
\]

(18)
\(\rho_Q\) being a dimensionless quantity that describes the charge-imbalance contribution to the resistance.

We consider \(R_{\text{min}}\) to be the resistance due to the normal region, since that is the total resistance when the superconductor is unperturbed. This leads us to associate

\[
R_S = \rho_Q R_{\text{min}}
\]

(19)
with the dynamic resistance of the superconductor \((eV > \Delta)\). This is the well-known branch imbalance resistance first described by Pippard \textit{et al.}\(^{10}\). If reverse currents due to both NS interfaces in the SNS junction are considered, we find

\[
R_{\text{SNS}} = (1 + 2\rho_Q)R_{\text{min}} .
\]

(20)
If \(d < L\), we have to first order

\[
R_{\text{SNS}} = (1 + d/\tilde{l} + 2\rho_Q)R_{\text{min}} .
\]

(21)
Thus \(\rho_Q\) can be considered a measure of the importance of nonequilibrium effects at the interface. Only if \(\rho_Q << 1\) can such effects be ignored. Equation (18) can easily be evaluated if \(\nu_F\tau_Q >> \tilde{l}\), i.e., in the dirty limit. Then the motion of the quasiparticles away from the interface is diffusive, so

\[
\lambda_Q = (\tau_Q\nu_F/3)^{1/2} ,
\]

(22)
\[
\rho_Q = (3\nu_F\tau_Q/4\tilde{l})^{1/2} >> 1 ,
\]

(23)
showing that nonequilibrium effects cannot be ignored here.

The result of this calculation is that for \(eV > 2\Delta\), the dynamic resistance of an SNS junction has an additional term \(2\rho_Q R_{\text{min}}\) which represents quasiparticle dissipation in the superconducting electrodes. This extra resistance is related to that seen in NS interfaces\(^{10}\) as \(T \rightarrow T_c\), where thermal excitations are responsible for injecting quasiparticles into the superconductor. In our case the applied voltage, rather than thermal excitations, is responsible for the injection.

In general it is not likely that an applied voltage \(eV > 2\Delta\) can be reached in one-dimensional SNS junctions without critical current, Ohmic heating, or other effects causing gap suppression, obscuring effects such as subgap structure which are dependent upon the superconducting gap. Instead, a two- or three-dimensional geometry must be used to gain access to this regime.

The calculation of \(\rho_Q\) can be done in a three-dimensional geometry with slightly more difficulty. This regime is also of interest because many experimental NS interfaces are believed to approximate point contacts or orifices in impermeable barriers between bulk electrodes. In these cases quasiparticles generated at the interface can diffuse away in more than one dimension, which reduces their concentration at the interface relative to that calculated above. This reduces \(I_r\) [Eq. (16)], and therefore \(\rho_Q\) [Eq. (18)].

If the radius of the point contact or orifice \((a)\), is such that \(a > \tilde{l}\), as is often the case, Eq. (14) becomes

\[
\int Q^*(v) dv = \int Q^*(r)V dr = 2\pi Q^*(a) \lambda Q a(\lambda Q + a) ,
\]

(24)
where \(Q^*(a)\) is the charge-imbalance density at the interface, \(r = a\), and we have used the solution to the spherical diffusion problem

\[
Q^*(r) \propto r^{-1} e^{-r/\lambda_Q} .
\]

(25)
In the steady state we must have

\[
I_r \tau_Q = \int Q^*(v) dv ,
\]

(26)
and in analogy to Eq. (16)

\[
I_r = \pi a^2 Q^*(a) \nu_F ,
\]

(27)
making

\[
\rho_Q = (\tau_Q \nu_F/2\lambda_Q) a/(a + \lambda_Q) .
\]

(28)
This reduces for \(a >> \lambda_Q >> \tilde{l}\) to

\[
\rho_Q = (3\nu_F\tau_Q/4\tilde{l})^{1/2} ,
\]

(29)
which is the one-dimensional result. This must be greater than 1, given our restriction \(\tau_Q >> 1/\nu_F\). For \(a << \lambda_Q\) Eq. (28) reduces to

\[
\rho_Q = \frac{a}{\lambda_Q} .
\]

(30)
Thus (at least when \(\tau_Q >> 1/\nu_F\)) \(\rho_Q >> 1\) is not possible at NS interfaces unless \(a << \tilde{l}\), as in an ideal point contact.

While \(\tau_Q\) is strongly temperature dependent,\(^{11}\) \(\tilde{l}\) is not. This means a temperature-independent value of \(\rho_Q\) (and therefore a temperature-independent charge-imbalance dynamic resistance for \(eV > \Delta\)) is obtained if \(a << \lambda_Q\). Otherwise \(\rho_Q\) and the resistance should be temperature dependent.

We saw in Eq. (28) that \(\rho_Q\) becomes independent of \(\lambda_Q\) if \(a << \lambda_Q\). Physically this means that the "diffusion resistance" \(\rho_Q = 3a/2l\) is the bottleneck in the escape of quasiparticles from the interface. Since \(\lambda_Q\) is then irrelevant in determining the quasiparticle density at the interface, it must also be true that the scattering events which contribute to \(\lambda_Q\) are irrelevant to the energy distribution of quasiparticles at the interface. Diffusion rates alone deter-
mine the distribution of quasiparticles. Extending this observation, if $\lambda_{\text{in}} \sim \lambda_Q$, the inelastic events which contribute to $\lambda_{\text{in}}$ also do not affect the quasiparticle distribution. We will call this the “diffusive” regime.

In the opposite extreme $a \gg \lambda_Q$, it is the contact diameter $a$ which becomes irrelevant to the quasiparticle distribution at the interface. Here recombination $(1/\tau_Q)$ rates alone determine the concentration, and inelastic events do not affect the energy distribution. This holds true also for less than three-dimensional fanout of the superconductor, as in a typical thin-film microbridge. If in addition we have $\lambda_{\text{in}} < \lambda_Q$ (for all relevant energies), we would expect complete thermalization of the quasiparticle distribution to occur subject to the constraint of constant $Q^*$. However, this condition holds only near $T_c$. We will call this the “thermalized” regime.

As mentioned above, we have neglect the lower effective charge $q_\text{eff}(E) = (1 - \Delta^2/E^2)^{1/2}$ of quasiparticles injected near the gap edge in Eqs. (15) and (26). We have also neglected, however, the proportionately lower group velocity of quasiparticles near the gap edge. The lower velocity slows the diffusion of quasiparticles from the interface, increasing the effect of each one on $Q^*(0)$. In the diffusive regime ($a < \lambda_Q$) these two effects cancel, leaving a well-defined (that is, energy-independent) $\rho_0 = 3a/2l$. We then have the simple result that an injected quasiparticle at any energy contributes equally to $Q^*(0)$.

In the thermalized regime ($a \gg \lambda_Q$), we find that the effects do not cancel. Rather, $v_p$ in Eq. (29) must be replaced by an (energy-dependent) $v_p = v_p q_k(E)$. In addition $\tau_Q$, which enters Eq. (29), is generally energy dependent. In this regime, however, we know the form of the excess quasiparticles distribution, namely thermal. At any given temperature, a thermally weighted average of $\rho_0(E)$ can be used as a single “effective $\rho_0$” for the superconductor. This single $\rho_0$ describes the decay of the entire thermalized distribution of excess quasiparticles, just as the single energy-independent $\rho_0$ in the diffusive regime describes the decay of the excess quasiparticles at each energy level.

While our main interest here is the SNS contact, we note that for a single NS junction the charge imbalance reduces $I$ for $eV > \Delta$, but has no effect on the excess conductance due to the Andreev-reflection process. Thus the effect of charge imbalance for $kT \ll \Delta$ is to increase the ratio of the low-voltage conductance to the high-voltage conductance, compared to the ratio calculated by Zaitsev and KBT.\textsuperscript{7} Curve $B$ in Fig. 3 shows the $IV$ characteristic of an NS junction for $\rho_0 = 3$, calculated assuming no thermalization of the quasiparticles (i.e., a three-dimensional contact with $a < \lambda_{\text{in}}$). Curve $A$ shows the result calculated neglecting charge imbalance ($\rho_0 = 0$). Note that since $T > 0$, charge imbalance generated by thermally excited electrons entering the superconductor modifies the low-voltage conductance ($eV < \Delta$) as well as the high-voltage conductance as discussed above.

V. KBT ANALYSIS OF THE SNS JUNCTION

KBT consider an SNS junction as two SN interfaces in series, with the $N$ layer between them sufficiently thin that electrons can move ballistically between the two superconductors with no scattering. Owing to the multiple Andreev reflections that can now be initiated by one electron, increased current can flow through the junction for an applied voltage $eV < \Delta$.

In Fig. 5 this process is shown. An electron enters the normal metal from $S1$, Andreev-reflects from $S2$, and a hole recrosses the normal metal. A pair propagates into $S2$ as a consequence of the reflection. Since the hole has opposite charge to that of the electron, it travels right to left again, gaining energy from the applied voltage, reflects from $S1$, and an electron recrosses the normal metal. A pair leaves $S1$ as a consequence of this reflection. After roughly $2\Delta/eV$ passes across the normal metal, the particle will have enough energy to escape into the superconductor. In the process of gaining this energy the particle has transported roughly $\Delta/eV$ pairs across the normal metal. Subgap structure, seen as peaks in the dynamic resistance ($dV/dI$) occurs then at voltages where electrons become able to pass through the normal metal with one fewer reflection.

We now describe the KBT calculation in greater detail. The flux of electrons transmitted into $N$ from $S1$, at energy $E$, is

$$f_0(E - eV)[1 - A_{NS}(E - eV)]$$

(31)

and that of holes is

$$[1 - f_0(E - eV)][1 - A_{NS}(E - eV)].$$

(32)

The electrons of term (31) will always traverse the normal region (Fig. 5) at least once. If they are Andreev-reflect against $S2$, and then again Andreev-reflect
against $S1$, they will traverse it again. The double-reflection process may then be repeated until the electrons are completely depleted by transmission into the superconductors. Similar considerations apply to the holes of term (32).

When these multiple contributions to current are considered, the total currents from incident electrons and from incident holes are found to be

$$\int f_0(E-eV)[1-A_{NS}(E-eV)](1+A_0A_1+A_0A_1A_2A_3 + \cdots) dE$$ (33)

and

$$\int [-1-f_0(E-eV)](1-A_{NS}(E-eV))[A_0+A_0A_1A_2 + \cdots] dE,$$ (34)

respectively. Here $A_m=A_{NS}(E+meV)$, and we have used the property that $A_{NS}(E)=A_{NS}(-E)$. The reverse currents from $S2$ into $N$ must also be considered. They are

$$\int [-f_0(E-eV)](1-A_{NS}(E-eV))$$

$$\times (1+A_0A_{-1}+A_0A_{-1}A_{-2}A_{-3} + \cdots) dE$$ (35)

and

$$\int [-f_0(E-eV)](1-A_{NS}(E-eV))$$

$$\times (A_0+A_0A_{-1}A_{-2} + \cdots) dE.$$ (36)

Terms (33) through (36) are then summed and the integrals performed numerically. Results of the KBT calculation are shown in Fig. 6, showing peaks in dynamic resistance at the correct voltages. However, when these calculations are performed for low temperatures ($T \ll T_c$), the subgap structure is seen to disappear (Fig. 7), whereas experimentally the amplitude increases with decreasing temperature. Furthermore, the calculated structure is neither so sharply defined in voltage, nor so large in amplitude, as the structure experimentally observed. Assuming zero Andreev-reflection outside the gap energies (as we shall do in Sec. VI) does not remove either of these difficulties. It serves merely to create “corners” on the peaks seen in Figs. 6 and 7.

These basic difficulties can be removed by considering the electron distribution in the normal metal, which is neglected by KBT, and by considering the effect of the quasiparticle charge imbalance that accumulates in the superconductors adjacent to the NS interfaces. As shown in Sec. IV, nonequilibrium effects in the superconductors occur in all SNS and similar junctions, and are important unless the geometry or material properties are such that injected (nonthermal) quasiparticles can move away from the interface extremely quickly. They will also be unimportant if a potential barrier greatly reduces the chance of an incident particle crossing the interface, as in a superconductor-insulator-superconductor (SIS) junction.12

**VI. NONEQUILIBRIUM DISTRIBUTION IN THE NORMAL METAL**

We now analyze the SNS case, employing the occupation requirement discussed in Sec. III. We allow Andreev reflection of an electron to occur only when the appropriate hole state is unoccupied. An electron or hole has three choices at each boundary it encounters: transmission, Andreev reflection, or simple reflection. The probabilities of these three are determined jointly by the Andreev-reflection coefficient $A_{NS}$ and by the occupation of the normal-metal state into which the charge may be reflected.

---

FIG. 6. Results of the KBT calculation of total current through an SNS junction at $T/T_c = 0.989$ (Ref. 3), and a typical experimental result (Ref. 10). In the experimental curve, the highest voltage peak occurs somewhat below $eV = 2\Delta$ due to gap suppression, possibly caused by heating.

FIG. 7. Results of the KBT calculation of total current through an SNS junction, for various temperatures.
The Andreev-reflection coefficient [Eq. (1)] was calculated under the assumption that the gap energy $\Delta$ is a step function, going from zero in the normal metal to its full value in the superconductor. This extreme discontinuity results in a higher Andreev-reflection coefficient for electrons incident outside the gap than would result for any less sharp variation of the gap energy. In our calculation we have used the other extreme often assumed for $SN$ interfaces: that the gap energy varies so slowly that the Andreev-reflection coefficient is zero outside the gap, but unity inside the gap. We note that the magnitudes of the subgap structure which our calculations yield would be diminished by using the nonzero reflection coefficient for $E > \Delta$.

We assume that the voltage applied across the junction appears as two equal discontinuities in the potential, one at each $NS$ interface, as shown in Fig. 8. The chemical potential $\mu_N$ of the normal metal is halfway in between those of the superconductors on either side. We have implicitly assumed that there is no spatial variation of $\mu_N$ within the normal metal, which amounts to an assumption of no scattering.

We define an electron distribution function $g^N(E)$ in the normal metal, as a function of $E$ relative to $\mu_N$. This will turn out not to be a Fermi distribution. We assume no spatial variation of $g^N(E)$.

We define currents at each energy of the normal metal, $I_{f\alpha}$, $I_{r\alpha}$, $I_{p\alpha}$, and $I_{h\alpha}$ are electron currents in the forward (left to right), and reverse directions, and hole currents in the forward and reverse directions, respectively. The total current through the normal metal is

$$I_{\text{tot}} = \frac{1}{2} \int \left[ I_{f\alpha}(E) - I_{r\alpha}(E) \right] - \left[ I_{p\alpha}(E) - I_{h\alpha}(E) \right] dE,$$

where $\frac{1}{2}$ is to avoid double counting.

We then set up reflection conditions on the currents. For instance, in Fig. 8,

$$I_{f\alpha}(E_1) = I_{h\alpha}(E_2) \left[ 1 - g^N(E_1) \right] + I_{p\alpha}(E_1) \left[ 1 - g^N(E_2) \right].$$

(38)

The first term is current contributed by the Andreev reflection of left- (reverse) going holes on level $E_2$. $1 - g^N(E_1)$ is the probability that level $E_1$ is unoccupied, so that the holes on level $E_2$ can reflect (as electrons) onto level $E_1$.

The second term is current due to simple reflection of left-going electrons on level $E_1$. $1 - g^N(E_2)$ is the probability that level $E_2$ is unoccupied, which would force the left-going electrons on level $E_1$ to simply reflect rather than Andreev reflecting onto level $E_2$ as holes.

Notice that the Andreev-reflection coefficient itself was unity in each of the above terms. That is because both levels were within the gap of the superconductor on the left. Outside the gap (e.g., on the right side of level $E_1$) the Andreev-reflection coefficient is taken to be zero. We have

$$I_{p\alpha}(E_1) = g^N(E_1 + eV/2),$$

(39)

where $g^N(E)$ is the distribution function of quasiparticles in the superconductor $S_2$, in the region adjacent to the $SN$ interface. In equilibrium $g^N(E)$ would be a Fermi distribution centered at $\mu_S$. Recall that the missing density-off-states factor for the superconductor cancels the transmission coefficient from the superconductor into the normal metal.

Similar equations hold for each of the other currents. The currents on all levels can be calculated self-consistently, once given the distributions $g^N(E)$ and $g^S(E)$.

Not included in the above equations is the effect of elastic scattering, as its inclusion would make interpretation of the equations difficult. Such scattering is of course present in any real system. We have found that failure to include it in numerical calculations leads to unphysical results. In the KBT results (Sec. V), for instance, the lack of scattering makes it possible for particles to make an infinite number of passes across the normal metal, before escaping, when the applied voltage $V \rightarrow 0$. Since with each two traversals of the normal metal two charges are transported through the $SNS$ junction, the KBT calculations yield an (unphysical) nonzero current as $V \rightarrow 0$.

Elastic scattering can be approximated by a one-dimensional model in which an electron has some probability $p_e$ ($p_e << 1$) of being scattered for each traversal of the normal metal. If $l$ is the elastic scattering mean free path and $d$ the thickness of the normal layer, then $p_e = d / l$.

Minimal elastic scattering (typically $p_e = 0.05$) has been included in the results presented in Sec. VIII. This should not seriously violate the assumption of no variation in $\mu_N$ across the normal metal, and in fact the calculated IV curves are insensitive to a reduction of $p_e$ by a factor of 2.

As an example of the inclusion of elastic scattering, consider $I_{f\alpha}(E_1)$, which is created on the left edge of the normal metal according to Eq. (38), $I_{f\alpha}(E_1)$, incident on the right edge, is used in the calculation of some reverse hole current [just as $I_{h\alpha}(E_2)$ is used in calculating $I_{p\alpha}(E_1)$]. When it is so used, the term...
\[ I_{fe}'(E_1) = (1-p_e)I_{fe}(E_1) + p_e g^N(E_1) \]  

(40)

is used in its place. If the normal-metal layer is such that \( d > l \), this simple approximation is expected to fail. In this case more extensive calculations which allow for the possibility of spatial variations in the currents and occupations are necessary.

The electron distribution of the normal metal, \( g^N(E) \), is determined by a Boltzmann equation. In Fig. 8 consider level \( E_1 \). On the right, the net flow of current out of it (electrons minus holes) is

\[ I_{fe}(E_1) - I_{fh}(E_1) \]

and on the left the net current out is

\[ I_{ne}(E_1) g^N(E_1) - I_{nh}(E_1) [1 - g^N(E_2)] \]

\[ - I_{ne}(E_2) [1 - g^N(E_1)] - I_{nh}(E_2) g^N(E_1) . \]

(42)

An equation similar to (41) holds whenever the normal-metal level is at an energy outside the superconductor gap, on either the left or the right. An equation similar to (42) holds whenever the normal-metal level is at an energy inside the gap.

Finally there is a contribution to \( g^N(E_1) \) from inelastic scattering, characterized by a time \( \tau_{in} \) with \( \tau_{in} >> \tau_0 \). This tends to relax \( g^N(E_1) \) toward a Fermi distribution \( f_0(E) \). The current out of level \( E_1 \) due to inelastic scattering is

\[ [g^N(E_1) - f_0(E_1)] \tau_{in}^{-1} . \]

(43)

Minimal inelastic scattering (about \( \tau_0 \) of the strength of ballistic currents) has been included in the results shown in Sec. VIII. As with the elastic scattering, the calculated \( IV \) curves are insensitive to a factor of 2 variation in inelastic scattering. Its inclusion, however, speeds up computation.

The distribution \( g^N(E) \) for all \( E \) can now be found self-consistently by setting the sum of Eqs. (41) and (42) to zero, i.e., no net current into or out of any level. Thus we could find \( g^N(E) \), given \( g^N(E) \) and the currents \( I_{fe}, I_{ne}, I_{fh}, I_{nh} \).

If we set \( g^N(E) \) to its equilibrium value \( f_0(E) \), we can solve for \( g^N(E) \) and the currents self-consistently by numerical methods, for each applied voltage \( V \). The calculated normal electron distribution function \( g^N(E) \) for a voltage \( eV = 2\Delta / 5 \) is shown in Fig. 9. Figure 10(a) shows the \( IV \) characteristic and derivative calculated for \( T = \Delta \). Notice that the function \( g^N(E) \) is far from the equilibrium function \( f_0(E) \), but that despite the nonequilibrium structure in \( g^N(E) \) the structure in the \( IV \) characteristic is less pronounced than that calculated by KBT or generally observed experimentally.

VII. NONEQUILIBRIUM DISTRIBUTION IN THE SUPERCONDUCTOR

We have so far neglected the nonequilibrium effects in the superconducting electrodes. As noted in Sec. IV this approximation is seldom if ever appropriate. To include this effect, we start with a given \( g^N(E) \), and solve for \( g^N(E) \) and the currents self-consistently, as before. We then calculate \( I^*(E) \), the charge-imbalance current injected into the superconductor on the right. We have

\[ I^*(E) = \frac{[I_{fe}(E-eV/2) - I_{fh}(E-eV/2)]}{1 - g^N(E)} \]

\[ - [g^N(E) - 1 - g^N(E)] . \]

(44)

The first term in square brackets is the charge flowing out of the normal metal into the superconductor. The second term in square brackets is the charge flowing from the superconductor into the normal metal. Note that once again the density of states in the superconductor and the \( SN \) transmission coefficient have cancelled.

In the superconductor on the right (Fig. 8), the injected charge is electron-like. The injected charge distorts \( g^N(E) \), which in turn affects \( g^N(E) \) through Eq. (39), affects the currents through Eq. (41), and affects \( I^*(E) \) through Eq. (44).

The injected charge will diffuse away from the \( NS \) interface while decaying with a characteristic time \( \tau_0 \), and thermalizing with characteristic time \( \tau_{in} \). The problem is then to calculate the nonequilibrium function \( g^N(E) \). In principle, \( g^N(E) \) can be calculated from the Boltzmann equation, diffusion rates, and a knowledge of the inelastic scattering and recombination rates as functions of quasiparticle energy. We have not attempted this calculation.

Instead, using insight gained in Sec. IV, we note that in many cases one of two simplifying assumptions can be made. If the contact is three dimensional, satisfying the condition \( a << \lambda_{in} \), then \( g^N(E) \) near the contact is determined solely by the rate of quasiparticle injection and diffusion. This was the assumption used in calculating the \( IV \) characteristic shown in Fig. 3, for an \( NS \) contact. Alternatively, if \( a >> \lambda_{in} \), or the diffusion of injected quasiparticles is restricted to less than three dimensions, then quasiparticles are relaxed toward a thermal distribution by inelastic events. In this case a general Boltzmann equation calculation is required to determine the resulting quasiparticle distribution \( g^N(E) \). However, in the limiting case of
\lambda_Q \gg \lambda_{nn}

which usually applies near \( T_c \), the injected quasiparticles can be assumed to be completely thermalized, while conserving charge imbalance \( Q^* \). We consider both of these limiting cases ("diffusive" and "thermalized") below.

For the diffusive case, \( a \ll \lambda_{nn} \), the deviation of the quasiparticle distribution from thermal equilibrium is given by

\[
\delta g^{S}(E) = \rho_Q I^*(E)/q_k(E),
\]

i.e., by a level-by-level steady state. This is a self-consistent equation, because \( g^S(E) \) appears in the Eq. (44) for \( I^*(E) \). Note that the term \( q_k(E) \) appears in Eq. (45) because (45) is a statement of number conservation, whereas \( I^*(E) \) is a charge current.

To calculate the \( IV \) characteristic we solve self-consistently for the distributions \( g^S \) [Eqs. (41)–(43)] and \( g^R \) [Eq. (45)], for the currents \( I_f, I_r, I_n \) [Eqs. (38)–(40)], and for the injected quasiparticle charge-imbalance current \( I^*(E) \) [Eq. (44)].

For the opposite case (thermalized) \( a \gg \lambda_Q \gg \lambda_{nn} \), it is assumed that the injected quasiparticles quickly relax to a steady-state distribution while conserving quasiparticle charge \( Q^* \), but not conserving the energy of the quasiparticle system. The distribution continues to relax more slowly, reducing \( Q^* \) with time constant \( \tau_Q \).

The distribution we have used is

\[
\delta g^{S}(E) = f_d(E/T),
\]

where \( \delta g^{S} \) is the deviation of the occupation of the electronlike branch from its thermal equilibrium value.

Steady-state equilibrium is established when

\[
\rho_Q \int I^*(E)dE = \int \delta g^{S}(E)dE
\]

Here we do not expect \( \rho_Q \) to be energy independent, but since the nonequilibrium distribution is of a known form (namely, thermal) a single, uniform "effective" \( \rho_Q \) may be used to describe the decay of the entire distribution. To calculate the \( IV \) characteristics we solve self-consistently Eqs. (38)–(44) and (47).

VIII. NUMERICAL RESULTS AND DISCUSSION

Examples of numerically calculated \( IV \) and \( dV/dI \) curves are shown in Figs. 10 and 11. The calculations in Fig. 10 were done assuming \( a \ll \lambda_{nn} \) (diffusive regime), while the calculations in Fig. 11 were done assuming \( a \gg \lambda_Q \gg \lambda_{nn} \) (thermalized regime). In general, the curve show a dynamic resistance which increases with \( V \) until \( eV \approx 2\Delta \). Subgap structure is seen superimposed on this rising resistance. In both regimes the subgap structure is more pronounced and more sharply defined than that cal-
culated ignoring charge-imbalance effects [Fig. 10(a)].

The amplitude of subgap structure depends upon both temperature and $\rho_0$. In Fig. 11(a) ($T < T_c$) the subgap structure is so strong that it results in negative resistance. In a current-biased measurement this would be seen as hysteresis. In Fig. 11(c) ($T = T_c$) the hysteresis has been removed. Sharp peaks changing to hysteretic regions at lower temperature are often observed.\textsuperscript{4,13} Raising $\rho_0$ [Fig. 11(b)] we see a strengthening of subgap structure, almost sufficient to cause hysteresis, while lowering $\rho_0$ [Fig. 11(d)] produces much weaker structure.

Comparing Fig. 11 to the experimental data shown in Fig. 6, we find good agreement in the shape of the resistance peaks at $2\Delta$ and $\Delta$. The higher-order peaks are not quite as distinct in the calculated curves as they are in the experimental curves. In Fig. 6, it may be noted that the $2\Delta$ peak occurs at about 1.6 times the voltage of the $\Delta$ peak. This is commonly attributed to Ohmic heating reducing the value of $\Delta$ as the voltage is increased. We have doubts about the correctness of simple Ohmic heating as a mechanism, but find the suggestion of gap suppression convincing.

The "diffusive" regime calculations shown in Fig. 10 are, as described above, appropriate for a truly three-dimensional contact satisfying $a << \lambda_{_0\text{m}}$. Thus it is perhaps not surprising that these calculations are less like the results generally seen in microbridges, which are typically two-dimensional structures. Note, however, that the higher-order peaks are more pronounced here than in the fully thermalized case.

Octavio et al.\textsuperscript{14} have documented a strong correlation between the decrease or disappearance of hysteresis in subgap structure, and the closeness of approximation to a "well-cooled" microbridge. In general, a well-cooled microbridge is achieved by adopting a variable-thickness bridge geometry which causes the microbridge to resemble, to some degree, a three-dimensional contact. Thus the requirements for good thermal cooling and for the quick dissipation of injected nonequilibrium quasiparticles without thermalization (low $\rho_0$) are likely to be satisfied simultaneously. In accord with Octavio's results, as one goes from a planar two-dimensional contact to a variable-thickness bridge, the system approaches the diffusive regime (Fig. 10) in which hysteresis is not seen, at least at the $\rho_0$ values we have investigated.

In even a qualitative comparison, it is important that a suitable choice of parameters is made. The data shown in Fig. 6 were obtained from an essentially two-dimensional indium microbridge with width and length of order 0.5 $\mu$m, film thickness $\sim 0.1 \mu$m, and mean free path $l \sim 0.1 \mu$m.\textsuperscript{13,15} If $\rho_0$ is calculated in a two-dimensional geometry, we find
\[ \rho_0 = a / l \ln(\Lambda_0 / a) \] (48)

Therefore values of \( \rho_0 \) of the order of 5 are reasonable.

It is interesting to note that in the calculated results for \( a \gg \lambda_m \) (Fig. 11) there is an increase in conductance visible in the curve of \( I \) vs \( eV \) beyond \( 2\Delta \). If observed experimentally, this would likely be taken as an indication of a tunnelling process. Yet our model does not include any tunnelling barriers. This effect is not seen in the diffusive regime (Fig. 10).

As indicated above we have calculated these \( IV \) characteristics assuming minimal scattering in the normal channel. This condition is often not met in experiments with thin-film microbridges. It is not met for the experimental data shown in Fig. 6. To ascertain, at least approximately, the effect of significant elastic scattering in the normal-metal channel, we calculated \( IV \) characteristics in the regime \( d \sim l \). As might be expected, the effect of increasing scattering is to gradually diminish the subgap structure, with the strongest effect in the highest-order peaks. The essentially linear increase in resistance below \( eV = 2\Delta \) is retained from \( d \ll l \) until \( d \sim l \).

For \( d \sim l \) the validity of our simple treatment of elastic scattering becomes questionable. However, in the opposite limit \( (d \gg l) \), a diffusive calculation can be performed in which the occupation of states in the normal metal varies spatially from one normal interface to the other. This calculation yields a much attenuated, but still quite visible, version of the subgap structure calculated for clean SNS contacts. This is in accord with the experimental observations of Warlaumont et al.\(^{16} \) on relatively dirty SNS microbridges.

In conclusion, we have added the effect of charge imbalance to the analysis of SNS junctions pioneered by KBT. This nonequilibrium effect becomes negligible only under ideal point contact conditions with the contact radius \( a \ll l \), where \( l \) is the elastic scattering length. This condition is seldom realized experimentally, especially in thin-film microbridges, as it requires a three-dimensional contact with a radius typically less than 100 nm. By including the effect of charge imbalance in the superconducting electrodes, and by calculating the electron distribution in the normal-metal channel, good qualitative agreement with measured \( IV \) characteristics is obtained.

Our analysis suffers from an oversimplification of the steady-state distribution function and decay mechanisms of the injected quasiparticles. It should be possible to correct this through a proper calculation of the decay of the nonequilibrium distribution similar to that performed by Chi and Clarke\(^{17} \) for tunnel injection, but with the additional complication of a spatial diffusion term.

Our analysis also ignores other nonequilibrium and thermal effects that may well exist in any real microbridge. In particular, the excess injected quasiparticles will cause some suppression of the gap energy as the voltage is increased. We have found in the simpler \( NS \) system that at applied voltages below the equilibrium gap energy \( \Delta_0 \) there is substantial reduction of \( \Delta \) due to this effect. (This will be discussed in a future paper.) The effect is not large enough to destroy superconductivity at the voltages where subgap structure is observed. There may also be effects due to the supercurrent. The fundamental importance of charge imbalance is, however, amply demonstrated.

ACKNOWLEDGMENTS

The authors are pleased to acknowledge fruitful discussion and argument with M. Arai, H. Baranger, U. Eckern, B. D. Hunt, J. Warlaumont, J. W. Wilkins, and especially Charles Rogers. Our thanks go to T. M. Klapwijk, G. E. Blonder, and M. Tinkham for sending us copies of their work prior to publication (Refs. 3 and 12). This work was supported by the Office of Naval Research.