Tunneling from a Many-Particle Point of View[1]

R. E. PRANGE

University of Maryland, College Park, Maryland

In this book are mentioned a number of applications of the theory of tunneling, and there are numerous others, ranging from nuclear pickup processes to the Esaki diode. However, we shall keep in mind, metal-to-metal tunneling through an oxide layer. In the courses in quantum mechanics which we had in our youth, we solved the problem of the tunneling of a single electron. This we did by considering an incident wave, a reflected wave, and a transmitted wave. The transmission coefficient can be calculated in some approximation (usually the WKB approximation). As is well known, the transmission coefficient depends chiefly upon an exponential factor

$$\exp \left[-2 \int_{x_0}^{x_1} \varkappa(x)\, dx\right]$$

where $\varkappa(x)$ is $\sqrt{2m\,[V(x) - \varepsilon_k]}$. Here we have already adopted the potential barrier model of the tunneling junction, and the WKB approximation, but the main features of this result are certainly of more general validity. However, we shall not discuss the calculation of the transmission coefficient further, because we are interested here in many-particle effects. Besides, most of the time, the experimenter has no independent measure of the thickness of the barrier, which usually is a horrible mess.

The simplest model of the metal system is that of independent quasi particles, in which we take into account only the Pauli principle. The results for this model can be obtained easily if we know the individual

[1] Research supported in part by U. S. Air Force Office of Scientific Research.

transmission coefficients. We have to make sure that there is an empty state available into which the electron tunnels, so that at zero temperature, for example, we can have a current only when we have applied a potential difference across the barrier. The effect of interaction in this case is obscured. However, for most purposes, this model provides excellent framework within which one can understand experimental results.

Recently, Bardeen (*1*) pointed out that it was superior to recognize the smallness of the parameter $|T|^2$, the tunneling coefficient, at the outset. Instead of treating the tunneling as a sort of scattering process, he treated it as a transition process between nearly stationary states. Let Φ_0 be the state at the initial time. In the absence of any tunneling, this state will evolve in time according to $\Phi_0(t) = \Phi_0 \exp(-iE_0 t)$. However, if one allows tunneling, the state will evolve into $\Psi(t) = \Psi_0(t) + \Sigma_\nu a_\nu(t)\Phi_\nu \exp(-iE_\nu t)$. The Φ_ν are states in which an electron has passed across the barrier. We remark that $(\Phi_\nu, \Phi_0) \neq 0$, but this sort of expansion is well understood. Let us apply time-dependent perturbation theory. We calculate

$$\left(i\frac{\partial}{\partial t} - H\right)\Psi(t) = (E_0 - H)\Phi_0 \exp(-iE_0 t) + \sum_\nu i\,\dot{a}_\nu(t)\Phi_\nu \exp(-iE_\nu t) + \sum_\nu a_\nu(E_\nu - H)\Phi_\nu \exp(-iE_\nu t) = 0\,.$$

Thus

$$i\frac{\partial}{\partial t} a_\nu = (\Phi_\nu, (H - E_0)\Phi_0) \exp[-i(E_0 - E_\nu)t]\,.$$

Then,

$$\frac{d}{dt}|a_\nu(t)|^2 = 2\pi\,|(\Phi_\nu(H - E_0)\Phi_0)|^2\,\delta(E_0 - E_\nu)\,. \quad (1)$$

This is the rate at which the initial state is transformed into the state Φ_ν. Bardeen has shown how one can rewrite the matrix element above in a more natural form. Consider the x-coordinate (perpendicular to the barrier) of the electron which has tunneled. Call the midpoint of the barrier $x = 0$. Then, for positive x, say Φ_0 is practically an eigenstate of H, so we may write

$$(\Phi_\nu, (H - E_0)\Phi_0) = \int \dots \int^0 dx\, [\Phi_\nu^* (H - E_0)\, \Phi_0(x_1 \dots x \dots)]$$

$$= \int \dots \int^0 dx\, [\Phi_\nu^*(H - E_0)\Phi_0 - \Phi_0(H - E_\nu)\Phi_\nu^*]$$

since the second term vanishes for negative x. Since $E_\nu = E_0$ we have

$$(\Phi_\nu(H - E_0)\Phi_0) = \int \dots \int^0 dx \left(-\frac{1}{2m}\frac{\partial}{\partial x}\right)\left(\Phi_\nu^* \frac{\partial}{\partial x}\Phi_0 - \Phi_0 \frac{\partial}{\partial x}\Phi_\nu^*\right)$$

$$= -i[J_x(0)]_{\nu 0}\,. \qquad (2)$$

Here $J_x(0)$ is the operator for the total current passing across the plane $x = 0$.

This is Bardeen's calculation, with which it is hardly possible to quarrel. Various people (*2–4*) have immediately jumped to the conclusion that it should be possible to represent the Hamiltonian in the presence of the barrier by

$$H = H_R + H_L + T + \dots \qquad (3)$$

where

$$T = \sum_{m,n} T_{mn} a_m^\dagger b_n + \text{Hermitian conjugate.}$$

Here $a_m^\dagger$ creates electrons in a state on the right and $b_m^\dagger$ creates electrons on the left. The labels m, n refer to some set of single particle states, not necessarily plane waves. The dots refer to terms of order T^2 and smaller. It is also assumed that T_{mn} does not depend on interaction to any appreciable degree.

This is also a conclusion which we do not wish to dispute. However, it might be valuable to examine in some model, in exactly what sense Eq. (3) is correct.

In the following, I shall describe the contents of a note of my own (*5*).

Let us adopt the single model of a tunnel junction described by a potential $V(x)$ (Fig. 1).

It is hard to believe that the symmetry, the simple barrier, or the specular nature of the transmission will change the results much, but I certainly have not examined in detail what errors these simplifications may have introduced. If we stick to this model then, the Hamiltonian is

$$H = H_0 + V + W \tag{4}$$

where H_0 is the independent electron Hamiltonian, $V(x)$ is graphed in Fig. 1, and W is the interaction between electrons, or between electrons and phonons. H_0 does not have to be the free-electron Hamiltonian.

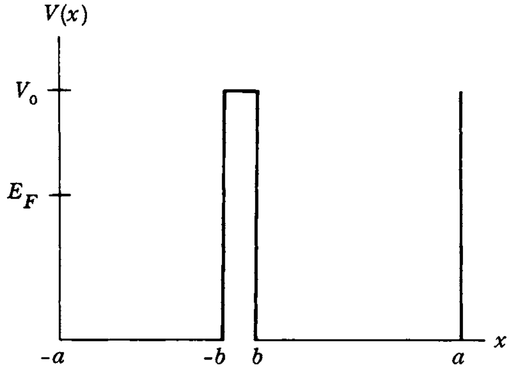

FIG. 1

The problem we wish to solve is this: Express the H of Eq. (4) in the form of Eq. (3). To solve this problem we have to decide what the labels m, n appearing in Eq. (3) really mean. What we would *really* like heuristically is to express the annihilation operator $\psi(x)$ as

$$\psi(x) = \sum_m \varphi_m'(x)a_m + \chi_m'(x)b_m \tag{5}$$

where (a) the φ', and χ' together form a complete orthonormal set, (b) any electron confined to the right (left) of the barrier can be expressed solely in terms of the states $\varphi_n'(\chi_m')$, (c) the single-electron Hamiltonian is a well-defined operator on all φ', χ' functions.

If we had such states, life would be simple, since we could substitute Eq. (5) into (the second quantized form of) Eq. (4), segregate the a and b operators into H_R and H_L, and what is left over would be T.

This explains why we need assumption (c) above, since without it, the substitution of (5) into (4) is not possible, or at best it is tricky. In fact, one does not want to deal with states φ', χ', which involve single-particle energies which are too great, otherwise H_R, for example, is too unlike the barrierless Hamiltonian we want. It is unfortunate, because states which vanish identically for one or the other sign of x are thereby ruled out and in consequence, we cannot satisfy both assumptions (a) and (b).

Since we have to give up one of these two assumptions, let us examine which one we should give up. I want to argue that assumption (a) must be foregone, at least the part which says "orthonormal."

I have not worked out a formal proof, but it appears to be true that complete orthonormal sets which do not involve high single-particle energies always have this condition: although they can be more or less confined to one side of the barrier, there is a long tail which leaks through. This tail is not negligible, because it really describes the effect which is under study. If we inject an electron into the right side of the tunnel junction, the tunneling rate is just given by the amount of "left-hand" wave function which is needed to form the initial wave packet.

An example of this state of affairs is shown in our model. The single particle eigenstates are even or odd in x. A complete orthonormal set is therefore

$$\Phi_k'[x_k'] = 2^{-1/2}\ [e_k(x) \pm o_k(x)]$$

where $e_k(o_k)$ are the even (odd) eigenstates of $H_0 + V$. However,

$$\frac{d}{dt} N_R = i \left\langle [\sum_k a_k^+ a_k, H] \right\rangle$$

is not constant for small times, as we expect on physical grounds. Instead, N_R oscillates in time, a behavior which is well known in elementary quantum mechanics, but is not observed in tunnel junctions.

The way out is to introduce c_n and d_n operators, (so no confusion arises with a_n and b_n), where

$$\psi(x) = \sum_n c_n \varphi_n(x) = \sum d_n \chi_n(x)\,. \tag{6}$$

In the symmetrical model, we may take $\varphi_n(x) = \chi_n(-x)$ and we suppose they satisfy, say,

$$[H_0 + V_1(x)]\,\Phi_n(x) = \varepsilon_n \varphi_n(x)$$

where $V_1(x)$ is graphed in Fig. 2.

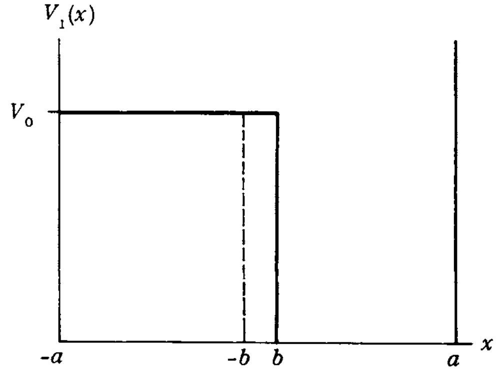

FIG. 2

The only important feature is that we somehow keep the leakage through the barrier of φ, χ to a minimum. *Both* $\{\varphi_n\}$, $\{\chi_n\}$ are complete othonormal sets. We want to express the Hamiltonian in terms of the c's and d's, as before, but in such a way that we do not mix up the single-particle energies more than is necessary. (This additional condition is necessary, because we have two orthonormal complete sets to do the job of one.)

The prescription is simple and obvious for the symmetrical case. I have never worked it out in general, but I do not doubt it can be done. The recipe is this: Express H in terms of the even and odd states $e_k(x)$, $o_k(x)$. Express these in turn according to

$$\begin{aligned} e_k(x) &= \sum_n \, [\varphi_n(x) + \chi_n(x)]\,\lambda_{nk} \\ o_k(x) &= \sum_n \, [\varphi_n(x) - \chi_n(x)]\,\mu_{nk}\,. \end{aligned} \tag{8}$$

It is completely straightforward, even if a little tedious, to evaluate the expansion coefficients λ, μ. The result is

$$\begin{aligned}\lambda_{kl} &= 2^{-1/2}\,[\delta_{kl} - T_{kk}/(\varepsilon_k - \varepsilon_l) - \tfrac{1}{2}\,\xi_{kl} + \dots]\\ \mu_{kl} &= 2^{-1/2}\,[\delta_{kl} + T_{kk}/(\varepsilon_k - \varepsilon_l) + \tfrac{1}{2}\,\xi_{kl} + \dots].\end{aligned} \tag{9}$$

We have introduced the single-particle matrix element of the current-across-the-barrier operator of Bardeen. Since we will always be interested in situations for which ε_k does not differ too much from ε_l, we can just keep the diagonal element. For these situations, ξ_{kl} is given by the overlap

$$\xi_{kl} = (\varphi_k, \chi_l)$$

and vanishes for states differing largely in energy (even though the inner product does not).

Perhaps for completeness it can be recorded that

$$T_{kl} = \delta_{k_y l_y}\,\delta_{k_z l_z}\,\frac{2\varkappa \exp(-2\varkappa b)}{ma[1 + (\varkappa/k\ )^2]}$$

where $\varkappa$ was defined before, and k_x is the wave number in the x direction. In this model, T_{kl} is diagonal, except for the variable k_x.

Neglecting, for now, the interaction, we have

$$\begin{aligned}H = \sum_{klm} &(\varepsilon_k^{(\text{even})})\,\lambda_{kl}\,(c_l^\dagger + d_l^\dagger)\,\lambda_{km}\,(c_m + d_m)\\ &+ (\varepsilon_k^{(\text{odd})})\,\mu_{kl}\,(c_l^\dagger - d_l^\dagger)\,\mu_{km}\,(c_m - d_m)\end{aligned}$$

where

$$\begin{aligned}\varepsilon_k^{(\text{even})} &\cong \varepsilon_k - T_{kk}\\ \varepsilon_k^{(\text{odd})} &\cong \varepsilon_k + T_{kk}.\end{aligned}$$

The algebra can be done easily, with the result

$$\begin{aligned}H = &\textstyle\sum_k \varepsilon_k c_k^\dagger c_k + \sum_k \varepsilon_k d_k^\dagger d_k\\ &- T_{kl}\,(c_k^\dagger d_l + d_l^\dagger c_k)\\ &- \tfrac{1}{2}\,\xi_{kl}\,(\varepsilon_k + \varepsilon_l)\,(c_k^\dagger d_l + d_l^\dagger c_k) + \dots\end{aligned} \tag{10}$$

This is what we expected, except for the last term (in ξ_{kl}). However, we should not forget that the $c_k^\dagger$ do not commute with the d_l, but that

$$\{c_k^\dagger, d_l\} = \xi_{kl}\,.$$

Therefore, $H_R = \Sigma_k\, \varepsilon_k c_k^\dagger c_k$ does not commute with $H_L = \Sigma_k \varepsilon_k d_k^\dagger d_k$. Consider, however, the effect of $H_R + H_L$ on some product state,

$$\Phi_R \Phi_L |0\rangle$$

where Φ_R contains only $c^\dagger$'s, and Φ_L contains only $d^\dagger$'s. Then

$$\begin{aligned}(H_R + H_L)\,\Phi_R\Phi_L|0\rangle &\simeq \Phi_L H_R \Phi_R|0\rangle \\ &\quad + \Phi_R H_L \Phi_L|0\rangle \\ &\quad + \sum_{kl} \varepsilon_k \xi_{kl}\,(c_k^\dagger d_l + d_k^\dagger c_l)\Phi_R\Phi_L|0\rangle.\end{aligned} \tag{11}$$

Thus, for processes such that $\varepsilon_k \sim \varepsilon_l$, the H_R may be regarded as commuting with H_L, provided we at the same time drop the last term of Eq. (10), which we wanted to do anyway.

The remaining thing is to consider the effects of interactions on this formulation. Suppose, for example, we have an electron-phonon interaction. We can write this as

$$H_{lp} = \sum_q v_q{}^i Q_q \varrho_q$$

as discussed by Professor Pines in this volume. Q_q is the canonical coordinate of the phonon, and ϱ_q is the density operator,

$$\varrho_q = \int \psi^\dagger(x)\psi(x) \exp(iq\cdot x)\, d^3x\,. \tag{12}$$

We want to express the $\psi(x)$ in terms of the localized operators c_n and d_n as before.

This may be done, with the result

$$\varrho_q \approx \sum_{mn} \varrho_{mn}{}^{rq} c_m{}^\dagger c_n + \varrho_{mn}{}^{lq} d_m{}^\dagger d_n + \varrho_{mnq}{}^T (c_m{}^\dagger d_n + d_n{}^\dagger c_m). \tag{13}$$

We have written

$$\varrho_{mn}{}^{rq} = \int \varphi_m \varphi_n \exp\,(iq\cdot xd^3x)$$

$$\varrho_{mn}{}^{lq} = \int \chi_m \chi_n \exp\,(iq\cdot xd^3x)$$

which just represent the usual electron-phonon interaction in the main part of the metal. (We have tacitly ignored the effect of the barrier on the phonons which, if taken into account, would complicate the formula.)

The amplitude $\varrho_{mnq}{}^T$ is expressed by

$$\varrho_{mnq}{}^T = \int \varphi_m \chi_n \exp\,(iq\cdot xd^3x) - \sum_s \xi_{ms}\varrho_{sn}{}^{lq} - \xi_{ns}\varrho_{ms}{}^{rq}\,. \tag{14}$$

We should not forget, however, that the c's and d's do not quite commute. We can pretend that they do, however, provided we drop the last terms (in $\xi_{\nu s}$) in Eq. (14).

Thus, the effect of the interaction on the tunneling is expressed by the addition of the small term

$$\sum v_q{}^i Q_q \varrho_{mn}{}^{qc}\,(c_m{}^\dagger d_n + d_n{}^\dagger c_m) \tag{15}$$

to the Hamiltonian, where

$$\varrho_{mn}{}^{qc} = \int \varphi_m \chi_n \exp\,(iq\cdot x)\, d^3x\,. \tag{16}$$

This is rather small, not only because it involves the overlap of the φ, χ functions, but because their overlap is rather smooth (in the x direction) over a region of (say) 30 Å. Thus, if q_x is not too small, we can expect this factor to be small. However, I don't believe it has been taken into account in the discussion of Kadanoff and Schrieffer in this volume.

In summary, it can be said that insofar as the simple model we have adopted has not led us seriously astray, it is possible to write the Hamiltonian in the form of Eq. (3) as long as we only want to use it to calculate the wave function to first order. Higher order corrections seem to be model dependent, but have never been calculated. Still, it is probably not misleading to use Eq. (3) to higher order, as Josephson has done, provided we are interested mainly in phase dependent terms, as explained by Professor Anderson in this volume.

References

1. J. Bardeen, *Phys. Rev. Letters* **6**, 57 (1961).
2. M. H. Cohen, L. M. Falicov, and J. C. Phillips, *Phys. Rev. Letters* **8**, 316 (1962).
3. J. Bardeen, *Phys. Rev. Letters* **9**, 147 (1962).
4· B. D. Josephson, *Phys. Letters* **1**, 251 (1962).
5. R. E. Prange, *Phys. Rev.* **31**, 1083 (1963).