Thermionic emission in heterosystems with different effective electronic masses

A. A. Grinberg
Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota 55455
(Received 8 July 1985; revised manuscript received 16 December 1985)

It is shown that the usual way of calculating thermionic emission takes into account incorrectly the conservation law of the quasimomentum of the electrons on the boundary separating two regions with different effective masses of electrons. Therefore, at a thermodynamic equilibrium the opposite flows of the electrons crossing this boundary are different (because the Richardson constants are different). This paradox disappears if one takes into account that not all electrons with larger effective mass can cross the boundary even if the part of their kinetic energy perpendicular to the interface is larger than the barrier height. In fact, this condition is contained in the quantum-mechanical coefficient of electron transmission. It is shown that in the spherical and parabolic effective-mass approximation the Richardson constant is defined by the smallest effective mass. This fact is important in the calculations of the thermionic currents in semiconductor contact phenomena especially in heterostructures. It makes clear what effective mass must be used in the definition of the preexponent factor of the thermionic current.

Modern semiconductor devices often employ carrier transport across heterointerfaces—boundaries of materials with different band structures. The dominant factor determining the current in such structures is the existence of a potential barrier due to band discontinuities, which affects the thermionic or tunneling current exponentially. There exists, however, another (preexponential) factor which is not connected with the barrier but depends on the probability of electrons, crossing the interfacial region (whose thickness is of the order of the interatomic distance) where the energy dispersion law is changed from one material to the other. When both materials possess a simple electron energy spectrum, the passage of an electron through the boundary can be described in terms of the dispersion relations alone, just as in the case of light passing through the boundary of two media of different dielectric permeabilities; the knowledge of the latter completely determines the processes of light transmission and reflection.

In the parabolic mass approximation when energy spectra of the electrons are defined by relations

$$E^{(i)}(k)=\frac{\hbar^2k^2}{2m_i},\quad i=1,2\ , \tag{1}$$

where m_1 and m_2 are the particle masses in the materials "1" and "2," respectively, the densities of the opposite currents of the electrons crossing the boundary are given by

$$|J_i|=\frac{em_i\exp(\mu/k_BT)}{2\pi^2\hbar^3}\int_0^\infty \exp(-E_\parallel/k_BT)\,dE_\parallel\int_U^\infty \exp(-E_\perp/k_BT)\,dE_\perp=A_iT^2\exp(-W)\ , \tag{2}$$

where U is the height of the barrier, μ is the Fermi level, $W=(U-\mu)$ is the work function, $A_i=em_ik_B{}^2/2\pi^2\hbar^3$ is the Richardson constant containing the effective mass m_i, k_B is the Boltzmann constant, $E_\perp$ and $E_\parallel$ are parts of the kinetic energy perpendicular and parallel to the surface, and it is assumed that $W >> k_BT$.

This is the usual form of the derivation of the Richardson formula in standard texts.[1-6] We came to the paradoxical result that the currents do not balance in equilibrium if $m_1\neq m_2$. Clearly, one must have the same Richardson constant for both thermionic flows, but which electron mass should it contain? The answer derived below is that the properly defined Richardson constant should involve the smaller of the two effective electron masses.

The best known case of the electron passage across a boundary of two media with different effective masses is the case of vacuum emission from a metal. In this case the thermionic emission current density J is given by the Richardson-Dushman formula (2) (Refs. 7 and 8) where effective mass m_i must be replaced by the free-electron mass m_0. Then $A_i=A_0=em_0k_B^2/2\pi^2\hbar^3$.

As can be seen in this case, the Richardson constant involves no characteristics of the emission body—a result which is quite contrary to the experimental evidence. To explain this contradiction a number of arguments have been advanced, such as a nonuniformity of the surface, a dependence of W on temperature, etc. Though all these factors are important there is, however, a more fundamental (and trivial) reason for the constant A entering the Richardson-Dushman equation not to be universal, namely, the fact that the usually used condition $E > U$ (Refs. 1–6) may not be sufficient for transmission. Because of the conservation of the tangential momentum, electrons incident on the boundary from the heavy-mass side at a sufficiently large angle to the normal direction will suffer a complete internal reflection. In fact, this assertion was pointed out by Herring and Nichols in the review paper[9] devoted to thermionic emission. Because it was stated in general form the quantitative influence of this reflection on the Richardson constant remains not clear. This we can see from the way that thermionic emission is described in every textbook.[1-6]

The "paradox" disappears if one takes into account the conservation laws of energy and momentum at the interface. Assume that the boundary is perfectly smooth. Then from the continuity of the electron wave function along the surface it follows that wave-vector components $(k_x,k_y)=\mathbf{k}_\parallel$ parallel to the surface are conserved ($\mathbf{k}_\parallel=\mathbf{k}'_\parallel$). We thus have

$$k^2=\frac{2m_iU}{\hbar^2}+\frac{m_1}{m_2}(k_\perp'^2)+\left(\frac{m_1}{m_2}-1\right)k_\parallel^2\ . \tag{3}$$

Let m_1 be larger than m_2. Then the possible values of k are restricted by the inequality

$$k_\perp^2 \geq \frac{2m_1 U}{\hbar^2} + k_\parallel^2 \left(\frac{m_1}{m_2} - 1 \right) . \tag{4}$$

At the same time there is no restriction on the perpendicular component of the $\mathbf{k}$ vector of light particles which penetrate to the side 1 of the system.

If we plot a family of the hyperboloid surfaces defined by Eq. (3) (in the $k_\perp > 0$ of the $\mathbf{k}$ space) using $k_\perp'$ as a parameter of this family, then they will fill up an inner part of the cone that is shown in Fig. 1. The asymptotic cone of these hyperboloids is defined by the equation

$$k_\parallel^2 = k_\perp^2 \Big/ \left(\frac{m_1}{m_2} - 1 \right) . \tag{5}$$

In the case of $m_1 < m_2$ the hyperboloids are transformed to the ellipsoids that envelop the center of $\mathbf{k}$ space and therefore fill up the whole space. In other words, due to the conservation laws a right side of the $\mathbf{k}$ space is reflected on the inner part of the cone of the $\mathbf{k}$ space.

It follows from the above that the integral (2) must be taken to the right, relative to the dashed line part of the cone (see Fig. 1). To do this we have to replace the lower limit of the integration variable $E_\perp$ by $U + [(m_1/m_2) - 1]E_\parallel$. Then we obtain[10]

$$|J_1| = A_2 T^2 \exp(-W/k_B T) = |J_2| , \tag{6}$$

that is, the solution of the Richardson constant paradox.

From Eq. (6) it might be seen that in a spherical mass approximation the Richardson constant is defined by a smaller mass. As a result it follows that in the case of the vacuum emission the Richardson constant A_0 is only the upper limit of its possible values. This assertion is in a perfect agreement with numerous experimental data.[11] From the data for the 77 elements of Mendeleev's table, contained in Ref. 11 only ten of them (C, Fe, Co, Zr, Mo, Rh, Cs, W, Re, Pt) were observed to have $A > A_0$. However, the experimental data for these ten elements contradict one another and the most probable values are less than A_0. There is only one exception, for Cs ($A = 162$); the reason for this is not clear.

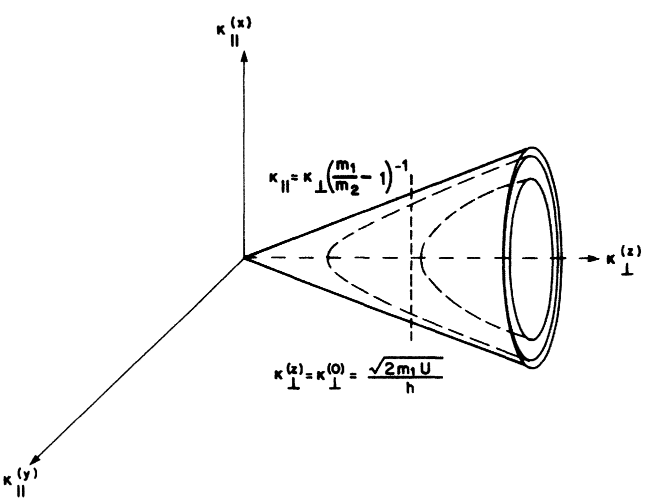

FIG. 1. A k-space cone from which particles with larger mass ($m_1 > m_2$) can cross a boundary surface.

It is necessary to underline that under "perfect agreement" we only assume the fact that observed values of the Richardson constant are not larger than its classical value A_0. From this assertion it does not follow that the average effective mass of the energy surface S [see Eq. (8)] is lower than free-electron mass in the all elements. Though such factors as the above-mentioned temperature dependence of the work function and inhomogeneity of the emitting surface certainly were taken into account in the experimentally defined values of A, it is impossible to say that there are not other factors which can decrease A. Nevertheless, it is interesting to see that an overwhelming amount of experimental data show A which are less than A_0 and we cannot exclude the possibility that the small average effective mass of the energy surface S is at least partly responsible for it.

So far we have not taken into account the wave-mechanical aspect of the electron transmission across the boundary surface. In fact, it was partly done when we restricted k space of electrons with a larger mass by the cone shown in Fig. 1. At the same time we implied that the transmission coefficient D for these electrons was equal to unity. The transmission coefficient automatically contains the restriction of the variance region of the variable $E^{(1)}$ that is equivalent to the inequality (4). It is important to emphasize that this restriction, in the case $m_1 \neq m_2$, is generally more important than the correction that is introduced by the distinction of D from unity in the region of energies E, where D is not equal to zero.

Let us now consider a more general situation with arbitrary dispersion of the particles. We will neglect the distinction of the transmission coefficient from unity except in that region where it is equal to zero. This is provided by the conservation laws from which the equation

$$E^{(1)}(k_\perp, \mathbf{k}_\parallel) = E^{(2)}(k_\perp', \mathbf{k}_\parallel) + U \tag{7}$$

follows. Then the density of the electron current from the region "1" to the region "2" can be written[9]

$$J_1 = \frac{-2e}{(2\pi)^3 \hbar} \int \int f(E) \frac{\partial E / \partial k_\perp}{|\mathrm{grad}_{\mathbf{k}} E|} dE\, dS . \tag{8}$$

The value $(\partial E/\partial k)/|\mathrm{grad}_{\mathbf{k}} E|$ is the cosine of the angle between the normal to the energy surface element dS and the normal to the interface. Therefore, the product of this value by dS is a projection of dS on a plane parallel to the interface in question. Because the region of the integration is subjected to the restriction defined by Eq. (7), an integration over S in Eq. (8) gives the area S that is the overlap of the projection surfaces $E^{(1)} = E$ and $E^{(2)} = E - U$ (see Fig. 2). We thus can write

$$J_1 = \frac{-2e}{(2\pi)^3 \hbar} \int_{E_{\min}}^{\infty} f(E) S(E)\, dE , \tag{9}$$

where $E_{\min}$ is the minimum value of $E^{(1)}(k_\perp, \mathbf{k}_\parallel)$ that is compatible with Eq. (7) and $f(E)$ is the Fermi function. Let us assume that both energies have a quadratic dependence on the wave vector so that $E^{(i)}(\alpha k) = \alpha^2 E^{(i)}$. Then $S(E)$ can be written as $S(E) = \gamma E$. As substitution of it in Eq. (8) yields

$$J = \frac{-\gamma e (k_B T)^2}{4\pi^2 \hbar} \ln\{1 + \exp[(\mu - e_{\min})/k_B T]\} . \tag{10}$$

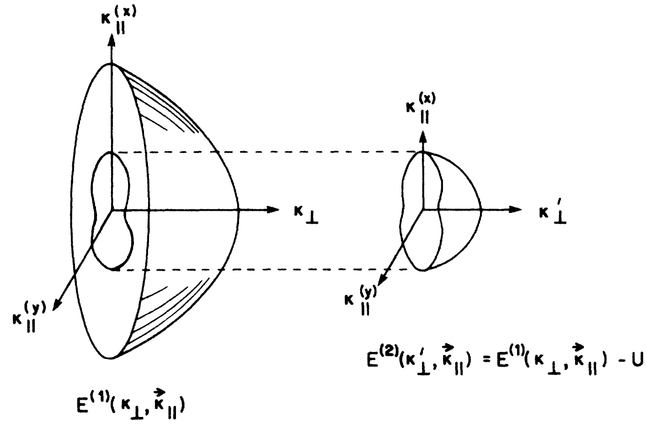

FIG. 2. Overlap of the "shadows" of the constant energy surfaces: $E^{(1)}=E$ and $E^{(2)}=E-U$.

If $W=-(\mu-E_{\min}) >> k_B T$ as is practically always the case, then we have

$$J=\frac{-\gamma e(k_B T)^2}{4\pi^2\hbar}\exp(-W/k_B T) \ . \quad (11)$$

In a parabolic mass approximation the value S is equal to the area of the main section of the constant energy sphere to which the smaller effective mass corresponds. In this case we can write $S=2\pi m_l E/\hbar^2$, where $\gamma=2\pi m_l/\hbar^2$ and we come up with the result (6).

We have considered the simplest situation when the energy minima of the electrons on both sides of the interface are located at the center of the Brillouin zones. For some systems this does not take place. In this case we can expect a dramatic decreasing of the Richardson constant because to cross the boundary surface an electron has to change wave vectors by a value of the order of the reciprocal lattice vector. A large difference of the electron wave vectors can be compensated by their interaction with some inhomogeneities of the surface. (It is not surprising, for instance, that the observed value of the Richardson constant of the Si is less than 10 $A/cm^2 K^2$.[11]) The thermionic current calculation then presents a much more complicated problem. It should be noted, also, that if at least one of the materials contains overlapping bands so that an incident electron can enter either one of the bands, then even the complete knowledge of the energy spectrum is insufficient to describe electron transmission. In this case one must also know the form of electronic interaction within the interface layer in order to determine the relative probabilities of electron injection into each of the bands. This represents a complicated problem which, as far as we are aware, has not been considered.

The author is indebted to S. Luryi for very useful discussions. This work was supported by Microelectronics and Information Center at the University of Minnesota and AT&T Bell Laboratories.

[1]K. Seeger, *Semiconductor Physics* (Springer-Verlag, New York, 1973).
[2]G. Busch and H. Schade, in *Lectures on Solid State Physics*, International Series on Natural Philosophy, Vol. 79, edited by D. Ter Haar (Pergamon, New York, 1976).
[3]J. S. Blakemore, *Solid State Physics* (Saunders, Philadelphia, 1974).
[4]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
[5]A. Van der Ziel, *Solid State Physical Electronics*, 3rd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1976).
[6]L. Solymar and D. Walsh, *Lectures on the Electrical Properties of Materials* (Oxford Univ. Press, Oxford, 1984).
[7]O. W. Richardson, *Emission of Electricity from Hot Bodies* (Longmans, New York, 1921).
[8]S. Dushman, Rev. Mod. Phys. **2**, 381 (1930).
[9]C. Herring and M. H. Nichols, Rev. Mod. Phys. **21**, 185 (1949).
[10]A. Van der Ziel (private communication) came to the same conclusion independently.
[11]V. S. Fomenko, *Handbook of Thermionic Properties* (Plenum, New York, 1966).