PHYSICAL REVIEW VOLUME 151, NUMBER 2 11 NOVEMBER 1966

Role of Screening in Surface Ion Neutralization*

VOLKER HEINE†

Department of Physics and Institute for the Study of Metals, The University of Chicago, Chicago, Illinois

(Received 12 May 1966; revised manuscript received 13 June 1966)

Two points in the theory of surface ion neutralization are discussed, following the lines established by Hagstrum and others. Firstly, because of the large energy $\hbar\omega$ transferred to the ejected election, the Coulomb interaction which is responsible for the process is effectively unscreened, but it is shown that the long range of the unscreened potential does not lead to any divergences in calculating the emergent current. The ejected electrons originate largely in the first one or two atomic layers of the solid,[1] and some even originate outside it. Secondly, we summarize previous work relating the observed directional dependence to variations of the matrix element and of the transmission coefficient through the surface, and add some new ideas.

I. INTRODUCTION

THE potential energy seen by an electron when some free ion such as Ne^+ is near a metallic surface is shown in Fig. 1. Electron 1 can tunnel from a state ϕ_a in the conduction band of the metal to the ground state E_g in the ion, thus neutralizing it. The excess energy $\hbar\omega = E_a - E_g$ is communicated via the Coulomb interaction to electron 2 in state ϕ_b of the conduction band, which is excited to state ϕ_e and may emerge from the metal with kinetic energy T_e. Such is the general picture of the neutralization process, established by Hagstrum and others[1–5] through experiment, interpretation, and theory.

However, some questions have remained, in particular, how far below the surface do the emergent electrons originate? One would suppose that the Coulomb interaction is effectively unscreened because of the large energy transfer $\hbar\omega$. However, if one then calculated with it naively, the total number of electrons excited, one would obtain a *divergent* contribution from processes deep inside the metal, arising from the long-range nature of the Coulomb interaction. Other questions concern the directional variation of the emergent current.

The rate of the neutralization process is given by[1–3]

$$(2\pi/\hbar) \sum_{e,a,b} |M_{eab}|^2 \delta(E_e + E_g - E_a - E_b), \tag{1}$$

where M_{eab} is the matrix element for a particular transition

$$\int\int \phi_g^*(\mathbf{r}_1)\phi_a(\mathbf{r}_1) V_{\text{int}}(\mathbf{r}_1,\mathbf{r}_2)\phi_b^*(\mathbf{r}_2)\phi_e(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \tag{2}$$

Strictly, we should subtract the corresponding exchange contribution,[2,3] which we shall ignore since it does not affect the following argument. V_{int} is the Coulomb interaction e^2/r_{12}, screened as appropriate. It is a very complicated function since the space of $\mathbf{r}_1$, $\mathbf{r}_2$ includes the region outside the metal around the gas ion, as well as the interior of the metal. In (1), transitions from different sets of band states a, b simply add, and it is therefore convenient to consider those involving a particular pair ϕ_a, ϕ_b. The energy of the emergent beam is then fixed at

$$E_e = E_a + E_b - E_g = E_b + \hbar\omega. \tag{3}$$

With ϕ_a fixed we may also perform in principle the $\mathbf{r}_1$ integration in (2). We can treat

$$v(\mathbf{r}_2) = \int \phi_g^*(\mathbf{r}_1)\phi_a(\mathbf{r}_1) V_{\text{int}}(\mathbf{r}_1,\mathbf{r}_2) d\mathbf{r}_1 \tag{4a}$$

$$= \int v(\mathbf{q})(\exp i\mathbf{q}\cdot\mathbf{r}) d\mathbf{q} \tag{4b}$$

as a perturbation in the one-electron Hamiltonian of electron 2. At large $\mathbf{r}_2$ it is a Coulomb potential, appropriately screened. It is centered at the point of maximum overlap of ϕ_g and ϕ_a, which is somewhere near the center of the surface ion because ϕ_g is so highly localized.

II. THE COULOMB INTERACTION

The main problem is the long range of $v(\mathbf{r}_2)$, i.e., the divergence of $v(\mathbf{q})$ at small $\mathbf{q}$, because the screening is very small for large energy transfer $\hbar\omega$ greater than the plasma frequency. We can see that the divergence disappears even in the simplest model by taking plane

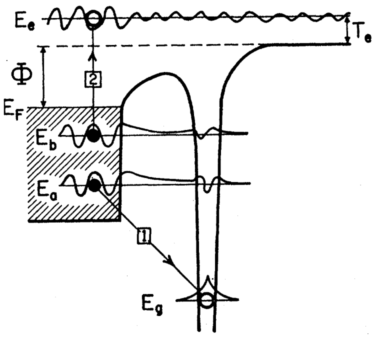

FIG. 1. Surface ion neutralization process.

---

[*] Supported in part by the National Aeronautics and Space Administration.

[†] Permanent address: Cavendish Laboratory, Cambridge, England.

[1] See, for example, H. D. Hagstrum, Phys. Rev. **96**, 336 (1954).

[2] H. D. Hagstrum, Phys. Rev. **122**, 83 (1961).

[3] H. D. Hagstrum, Phys. Rev. **150**, 495 (1966).

[4] H. D. Hagstrum and G. E. Becker, Phys. Rev. Letters **16**, 230 (1966).

[5] H. D. Hagstrum, Y. Takeishi, and D. D. Pretzer, Phys. Rev. **139**, A526 (1965).

waves for ϕ_b, ϕ_e, and expressing v in terms of a free-electron screening constant $\epsilon(q,\omega)$:

$$v(q)=4\pi e^2/q^2\epsilon(q,\omega). \qquad (5)$$

In any case the free-electron model is a suitable zero-order approximation for many simple metals and semiconductors. The matrix element is simply $v(Q)$, where $\mathbf{Q}=\mathbf{k}_e-\mathbf{k}_b$. Here $\mathbf{k}_b$ is assumed to be some fixed state and $\mathbf{k}_e$ can lie on a sphere of radius k_e determined by (3). The point is that Q can vary from k_e-k_b to k_e+k_b, but *components* $v(q)$ *with* $q\approx 0$ *do not contribute to the excitation of electrons at the high energies studied.* The minimum value of $\hbar\omega$ for an electron to be observed in this type of experiment is the work function Φ. Of course other electrons with lower $\hbar\omega$ are produced but cannot escape through the surface. But for these also, Q cannot tend to zero unless $\hbar\omega$ tends to zero as well, and in that limit the screening $\epsilon(q,\omega)$ removes the $1/q^2$ divergence in $v(q)$.

The conclusions remain valid when we take Bloch waves $\phi(\mathbf{k}_b,n_b;\,\mathbf{r}_2)$ and $\phi(\mathbf{k}_e,n_e;\,\mathbf{r}_2)$, where $\mathbf{k}$ is now the reduced wave vector and n the band index. It is now perfectly possible to have $\mathbf{k}_e=\mathbf{k}_b$ and for small q terms in (5) to contribute. Let us focus attention this time on transitions from all ϕ_b in a filled band n_b to a particular final state ϕ_e in the higher band n_e. The number of emergent electrons in ϕ_e is proportional to

$$\int[v(q)M(\mathbf{q})]^2d\mathbf{q}, \qquad (6)$$

where

$$M(\mathbf{q})=\int\phi^*(\mathbf{k}_e+\mathbf{q},n_b)\exp(i\mathbf{q}\cdot\mathbf{r})\phi(\mathbf{k}_e,n_e)d\mathbf{r}. \qquad (7)$$

$M(\mathbf{q})$ must vanish by orthogonality for $q=0$, and from $\mathbf{k}\cdot\mathbf{p}$ perturbation theory we have $M(\mathbf{q})\propto q$, which kills one of the factors of q in (5). The other, when squared, disappears with the volume element $d\mathbf{q}=4\pi q^2dq$. Thus the small q region gives a finite contribution: there is no divergence.

Finally, it remains to consider the effect in the real situation of terminating the wave function ϕ_b at the surface of the metal, with some tail overlapping the surface ion. ϕ_e exists both inside and outside the metal. If we choose cylindrical coordinates ρ, θ, z with axis perpendicular to the surface, we see that the $\rho d\rho$ in the volume element is not strong enough to give a divergence with $1/r$ in the matrix element

$$\int\phi_e^*(\mathbf{r})v(\mathbf{r})\phi_b(\mathbf{r})d\mathbf{r} \qquad (8)$$

when we integrate over the volume just outside the surface of the metal.

Since the center of $v(\mathbf{r})$ lies near the surface ion, an appreciable fraction of (8) may come from the region between the metal surface and the ion.[1–3] Thus, some of the emergent electrons may be regarded as generated outside the metal proper, for the complete emergent wave packet originating from a particular ϕ_b may be written[6] (treating the ϕ_e as free waves for simplicity)

$$\phi(\mathbf{r})=\int\frac{4\pi\exp(ik_e|\mathbf{r}-\mathbf{r}_2|)}{|\mathbf{r}-\mathbf{r}_2|}v(\mathbf{r}_2)\phi_b(\mathbf{r}_2)d\mathbf{r}_2, \qquad (9)$$

where in classical terms $v(\mathbf{r}_2)\phi_b(\mathbf{r}_2)$ is the source of the wave. As regards the inside of the metal, we have already seen that the source function has finite Fourier components as $q\rightarrow 0$, so that its scale in real space is given by the inverse of its extent in q space. The main contributions come from q comparable with the first reciprocal lattice vectors as shown by the free-electron model; and the same will be even more true for d bands in transition metals. The emergent electrons originate largely in the first one or two atomic layers.

We can illustrate this by two model calculations for the total source function[7]

$$S(\mathbf{r}_2)=\sum{}_b r_2^{-1}\phi_b(\mathbf{r}_2), \qquad (10)$$

summed over all initial electrons with $E_b<E_F$. We have included no screening of the Coulomb potential so that (10) applies for high-energy electrons where ϵ is effectively unity. Also we have approximated $v(\mathbf{r}_2)$ by a pure r_2^{-1} potential which is only valid for r_2 greater than, say, one atomic radius R, but this does not matter since we are interested in asking how quickly S falls off at distances $r_2>R$. For a free-electron gas, (10) gives

$$S(\mathbf{r}_2)\propto\left[\frac{\sin x}{x^4}-\frac{\cos x}{x^3}\right]_{x=k_Fr_2}, \qquad (11)$$

which falls off rapidly as r_2^{-3}. This should be accurate to about 20% for nontransition metals and group-IV semiconductors. For the latter the situation at long range is even more favorable because in (10) the summation is over full bands, so that $\sum_b\phi_b$ may be replaced by the Wannier function, which at large distances has an exponential decay,[8] though for small band gaps the exponent is small and probably not of great significance. Since (11) oscillates at large r_2 with wave number k_F which in general does not match that of ϕ_e, the wavelets generated by S over a volume encompassing one wavelength $2\pi/k_F$ largely cancel, so that the mean effectiveness of (11) falls off as r_2^{-4}. S must not be interpreted too literally; since the experiment measures the momentum (i.e., energy and direction) of the emerging electrons, it is meaningless to ask where in space they come from. What is physically relevant is how far the emergent beam is degraded by inelastic electron-electron collisions, etc., which of course are proportional to the depth of solid the electrons have to traverse. Analogously to (9), the total emergent beam at energy E may

---

[6] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., p. 165.

[7] Since the whole reaction is localized around the ion on the surface, the processes involving different hole pairs ϕ_a, ϕ_b, but the same ϕ_e have to be added *coherently*.

[8] W. Kohn, Phys. Rev. **115**, 809 (1959).

be written

$$\phi(\mathbf{r})=\int G(\mathbf{r},\mathbf{r}_2,E_e)S(\mathbf{r}_2)d\mathbf{r}_2, \tag{12}$$

where the Green's function G is a unit wavelet centered at $\mathbf{r}_2$ which includes all collision effects. [The $\mathbf{r}$ in $\phi(\mathbf{r})$ and $G(\mathbf{r})$ must now be interpreted in a generalized sense to include information about the electron-hole pairs generated by the inelastic collisions.] The collision terms in G increase with $|\mathbf{r}-\mathbf{r}_2|$. As already noted, in the outer regions of S there is considerable cancellation between wavelets because of oscillations in the amplitude factor $S(\mathbf{r}_2)$ and in G, and this interference applies to the whole of the wavelets *including* the collision part. The effective amplitude of a beam which we have to think of (in a semiclassical sense) as a collision modified from a point $\mathbf{r}_2$ therefore falls off as r_2^{-4}. So much for the nearly-free-electron solids. At the opposite extreme lies excitation from the d bands of copper, ignoring hybridization with the nearly-free-electron "conduction" band of the type already discussed. Then the next-nearest-neighbor contributions in a tight-binding expansion of the band structure are negligible,[9] and hence the Wannier function and $S(\mathbf{r}_2)$ extend no further than nearest neighbors.

In some experiments the surface has been covered with a well-formed monolayer of oxide. In inhomogeneous systems it is most useful to consider what has been termed the "local density of states,"

$$n(E,\mathbf{r})=\langle\phi_E{}^*(\mathbf{r})\phi_E(\mathbf{r})\rangle n(E), \tag{13}$$

where $\langle\ \rangle$ denotes the average value for wave functions ϕ of energy E. If we combine the density-of-states factors[4] inherent in (1) with the ϕ from (2), we see that it is approximately the quantity (13) which the neutralization process involves at E_a, E_b, and E_e. Now the $v(\mathbf{r}_2)$ will penetrate through the oxide into the underlying metal so that some of the emergent electrons will originate there; their distribution would reflect the density of states in the metal as usual. For the electrons generated in the oxide layer, $n(E_b,\mathbf{r}_2)$ is the sum of two parts. One comes from the filled valence band of the oxide. The other is from the tails of the metal wave functions at energies in the band gap of the oxide, which decay exponentially in the oxide as in tunneling experiments. The $n(E_a,\mathbf{r}_1)$ for electron 1 is the same as $n(E_b,\mathbf{r}_2)$ in the oxide. We, therefore, expect[3] the unfolded density of states[3,4] to contain the attenuated ghost of the metal density of states superposed on the oxide density of states, and this is not an unreasonable description of the observations as already noted by Hagstrum.[3]

III. INTENSITY VARIATIONS

In this section we summarize previous work, sharpen up one or two points, and add some new ones. Two orientational effects arise from the surface. In the free-electron model the total number of electrons excited to $\mathbf{k}_e$ is independent of its direction. However, states with $\mathbf{k}_e$ perpendicular to the surface have the largest transmission coefficient through it, as discussed in detail by Hagstrum.[1-3] Similarly, the magnitude of the tail of ϕ_b (and of ϕ_a) outside the metal is largest when $\mathbf{k}_b$ ($\mathbf{k}_a$) is perpendicular to the surface[1-3]; the decay constant for the tail of negative-energy states is analogous to the transmission coefficient for positive energy. Further, the contributions to the matrix element from both inside and outside the metal are largest when $\mathbf{k}_e$ and $\mathbf{k}_b$ are parallel, for then q is smallest, and $v(q)$ largest. For free electrons, both effects give an orientational dependence relative to the direction of the surface, the former probably being dominant. With a real-band structure both effects will also give a variation with surface orientation because the states of a given E_b (and E_a, E_e) will be weighted differently.[3]

The d bands in transitional metals do not contribute as strongly, relative to the plane-wave bands, as in optical excitation.[3,4] Firstly, the amplitude of the tail of ϕ_a overlapping the ion is smaller,[3] and secondly, the matrix element (7) does not have the enhancement provided in the optical case by the momentum operator, i.e., by differentiating once. So much for the over-all intensity of the d band. Within it there will be marked variations of the matrix element depending on the precise form of ϕ_a, ϕ_b, and the orientation of the surface, as seen for example in the field-ion microscope.[10] We may think of the d states in terms of tight-binding wave functions. The amplitude of the tail of the wave function depends on the amplitude of the ϕ at the surface, and will be larger for bonding- than for antibonding-type functions, as well as for Bloch functions made out of orbitals which have a lobe pointing directly out of the surface. The decay constant of the wave functions outside the surface depends on the wave vector $k_{\parallel}$ of ϕ parallel to the surface. For present purposes this is not just the parallel component of the $\mathbf{k}$ labeling the Bloch state, but is a combination of that and the angular variation of the atomic orbitals. We must take $k_{\parallel}$ from the variation of the complete Bloch function on the surface plane.

ACKNOWLEDGMENTS

I am greatly indebted to Dr. H. D. Hagstrum for formulating the questions which prompted the present paper in reply. The work was largely carried out during a visit to the Bell Telephone Laboratories.

---

[9] F. M. Mueller, Phys. Rev. (to be published).

[10] E. M. Mueller, S. Nakamura, O. Nishikawa, and S. B. McLane, J. Appl. Phys. 36, 2496 (1965).