

Tunneling of electrons from a two-dimensional channel into the bulk

S. V. Meshkov

Institute of Solid State Physics, Academy of Sciences of the USSR, Moscow

(Submitted 23 June 1986)

Zh. Eksp. Teor. Fiz. **91**, 2252–2262 (December 1986)

It is shown that the interaction of charge carriers quantized in a two-dimensional channel, both with each other and with inhomogeneities, has a fundamental influence on their probability of tunneling into the bulk, and as a result there is a slowing down in the decrease of the electron density with distance for carriers with a nonzero kinetic energy of motion along the surface. In the many-valley case an additional increase in the tunneling probability is observed. This increase is connected with the decrease of the mass in the tunneling direction on account of virtual intervalley transitions.

We shall consider a planar two-dimensional channel $V(z)$, situated on the surface of a semiconductor crystal. We shall assume that the channel is formed by a one-dimensional potential well that has (for simplicity) a single energy level ε_0 and is smooth enough for the effective-mass approximation to be applicable. It is known that for a nondegenerate carrier-dispersion law, in the absence of electron-electron interaction and inhomogeneities, the free motion along the channel can be separated completely from the quantized transverse motion. For this reason, the wave function Ψ of the charge carrier situated in the channel, irrespective of their energy ε , decay with distance from the surface at the same rate:

$$\Psi \propto \exp\left\{-\hbar^{-1} \int^z [2m(V(z) - \varepsilon_0)]^{1/2} dz\right\}.$$

The carriers tunnel into the bulk as if they had not the total energy ε but the energy ε_0 corresponding to the bottom of the surface band, and the kinetic energy of the motion along the channel remains irrelevant.

In this paper we ascertain to what extent this idealized picture is applicable to real two-dimensional channels—MIS structures and heterojunctions. The above-indicated separation of variables by such factors as scattering of carriers by inhomogeneities of the structure and by each other, and this qualitatively changes the situation. However weak the scattering processes, at sufficiently large distances they lead to the electron-density-decay law that would apply if the carriers had tunneled in the one-dimensional potential with their total energy:

$$|\Psi|^2 \propto \exp\left\{-2\hbar^{-1} \int^z [2m(V(z) - \varepsilon)]^{1/2} dz\right\}.$$

In addition, if the dispersion law is complicated, the carriers tunnel with that mass value which ensures the slowest decay of the tunneling exponential.

The statement formulated is fairly obvious and can be proved in an extremely general form. However, to estimate the distances at which the exponential effects under discussion become appreciable, it is necessary to know also the pre-exponential coefficients, which depend on the concrete scat-

tering mechanism. Below, for illustration, we consider model examples in which the exponential dependences are trivial, and the scattering processes are taken into account by perturbation theory. The simple expressions obtained for the pre-exponential factors are valid for practical estimates of the strength of the effects under consideration in real structures.

In Sec. 1 we obtain the zeroth approximation to the one-particle Green function for a crystal with a short-range channel on the surface. Using this Green function, in Sec. 2 we consider the short-range electron-electron interaction, and in Sec. 3 we consider scattering by inhomogeneities in the bulk and on the surface. In Sec. 4 we investigate the influence of bulk scatterers on the tunneling in the many-valley case. In the Conclusion we discuss a general rule for the calculation of the tunneling exponential far from the channel, and possible experimental consequences.

1. THE GREEN FUNCTION

The spatial distribution of electrons with a specified energy ε is completely described by the one-particle density matrix $\rho(\mathbf{r}, \mathbf{r}'|\varepsilon)$, which, in turn, is determined by the imaginary part of the causal Green function:

$$\rho(\mathbf{r}, \mathbf{r}'|\varepsilon) = \pi^{-1} \theta(\varepsilon_F - \varepsilon) \text{Im} G(\mathbf{r}, \mathbf{r}'|\varepsilon),$$

$$G(\mathbf{r}, \mathbf{r}'|\varepsilon) = \hbar^{-1} \int_{-\infty}^{+\infty} G(\mathbf{r}, \mathbf{r}'|t-t') \exp[i\varepsilon(t-t')/\hbar] dt. \quad (1)$$

To study the tunneling of carriers from a surface channel into the bulk it is necessary to construct the corresponding density matrix and to consider its asymptotic form at large distances from the surface. The most important quantity describing the electron-density distribution is the concentration $n(\mathbf{r}, \varepsilon) = \rho(\mathbf{r}, \mathbf{r}|\varepsilon)$ of electrons of a specified energy. We shall not confine ourselves to the diagonal elements of the density matrix, since in the study of the various processes involving tunneling carriers, e.g., in the calculation of recombination matrix elements, more-complete information is required.

We shall consider the simplest formulation of the problem, modeling the channel by a very narrow and deep poten-

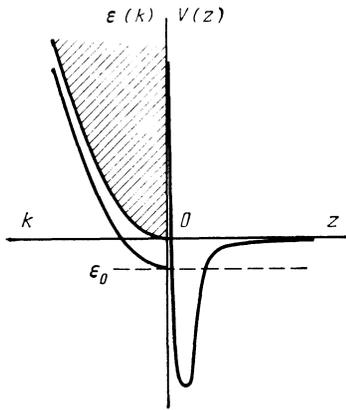


FIG. 1. The potential $V(z)$ of the surface channel. In the left half the surface band formed from the shallow level ε_0 in the well $V(z)$ is shown. The projection of the bulk band on the surface is indicated by the shading.

tial well $V(z)$, bounded on the side $z < 0$ by an infinitely high wall. On the crystal side $z > 0$ we shall neglect the bending of the bands (see Fig. 1). We shall ignore the spin-orbit interaction, and, accordingly, omit the spin indices from the Green functions. For the moment we shall assume the effective mass m of the electrons to be isotropic. The energy ε_0 of the single (for simplicity) bound state in the well $V(z)$ will be assumed to be considerably smaller than the depth of the well. This model is convenient in that the surface wave functions are localized on scales large in comparison with the size of the well, and the well can be replaced by the effective boundary condition

$$\left. \frac{\partial \Psi}{\partial z} \right|_{z=0} = -(-2m\hbar^{-2}\varepsilon_0)^{1/2} \Psi \quad (2)$$

on the electron wave function.

In the absence of inhomogeneities the one-electron wave functions of the surface states have the form

$$\Psi_{\mathbf{k}}(\mathbf{r}) \propto \exp[-(-2m\hbar^{-2}\varepsilon_0)^{1/2}z + i\mathbf{k}\boldsymbol{\rho}] \quad (3)$$

($\boldsymbol{\rho}$ and \mathbf{k} are the two-dimensional position vector and wave vector in the xy plane), and correspond to the energies

$$\varepsilon(\mathbf{k}) = \hbar^2 k^2 / 2m + \varepsilon_0.$$

The continuum states are described by the wave functions

$$\Psi_{\mathbf{k}, k_z}(\mathbf{r}) \propto e^{i\mathbf{k}\boldsymbol{\rho}} \cos \{k_z z + \arctg [(-2m\varepsilon_0)^{1/2} / \hbar k_z]\} \quad (4)$$

with energies $\varepsilon(\mathbf{k}, k_z) = \hbar^2 (k^2 + k_z^2) / 2m$ corresponding to the bulk dispersion law. The zeroth-approximation Green function $G^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon)$ can be obtained as the summation

$$G^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) = \sum_n \frac{\Psi_n^*(\mathbf{r}') \Psi_n(\mathbf{r})}{\varepsilon - \varepsilon_n + i0 \cdot \text{sign}(\varepsilon_n - \varepsilon_F)} \quad (5)$$

over all the stationary states with the appropriately normalized wave functions (3), (4). The surface band is assumed to be filled up to the Fermi level $\varepsilon_F = \varepsilon_0 + \hbar^2 k_F^2 / 2m < 0$. It is simpler, however, to solve directly the Schrödinger equation for the function $G^{(0)}$,

$$\left[\varepsilon + \frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] G^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) = -\delta(\mathbf{r} - \mathbf{r}'), \quad (6)$$

with a boundary condition analogous to (2). It is convenient to go over to the momentum representation in the plane of the channel, while remaining in the coordinate representation in the z direction. Solving Eq. (6) in this representation, we obtain the function $G^{(0)}$ in the form

$$\begin{aligned} G^{(0)}(\mathbf{k}, z, z' | \varepsilon) &= \int G^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) e^{i\mathbf{k}(\boldsymbol{\rho} - \boldsymbol{\rho}')} d^2\rho \\ &= \frac{m}{\hbar^2 (k^2 - 2m\hbar^{-2}\varepsilon)^{1/2}} \left\{ \exp[-(k^2 - 2m\hbar^{-2}\varepsilon)^{1/2} |z - z'|] \right. \\ &\quad - \frac{[(\hbar^2 k^2 / 2m - \varepsilon)^{1/2} + (-\varepsilon_0)^{1/2}]^2}{\varepsilon - \varepsilon_0 - \hbar^2 k^2 / 2m + i0 \cdot \text{sign}(k - k_F)} \\ &\quad \left. \times \exp[-(k^2 - 2m\hbar^{-2}\varepsilon)^{1/2} (z + z')] \right\}. \end{aligned} \quad (7)$$

The physical meaning of the two terms in the curly brackets is clear: The first term, which does not contain ε_0 , describes the purely bulk tunneling of the electron, while the second corresponds to tunneling with reflection from the surface. The contribution of the bulk term to the function $G^{(0)}$ in the coordinate representation is equal to

$$G^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) = \frac{m}{2\pi\hbar^2} \frac{\exp[-(-2m\hbar^{-2}\varepsilon)^{1/2} |\mathbf{r} - \mathbf{r}'|]}{|\mathbf{r} - \mathbf{r}'|}. \quad (8)$$

This contribution dominates in $G^{(0)}$ when the points \mathbf{r} and \mathbf{r}' are far from the surface, but (for $\varepsilon < 0$) does not have an imaginary part, which appears in the function $G^{(0)}$ only as a result of the participation of the surface states in the sum (5).

The electron density corresponding to the Green function (7) is

$$\begin{aligned} \rho^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) &= \theta(\varepsilon_F - \varepsilon) \theta(\varepsilon - \varepsilon_0) \frac{(-2m^3\varepsilon_0)^{1/2}}{\pi\hbar^3} J_0(k|\boldsymbol{\rho} - \boldsymbol{\rho}'|) \\ &\quad \times \exp[-(-2m\hbar^{-2}\varepsilon_0)^{1/2} (z + z')] \end{aligned} \quad (9)$$

[$J_0(x)$ is a Bessel function] and is associated exclusively with the surface states. It falls off exponentially with the distance of \mathbf{r} and \mathbf{r}' from the channel, and does not depend on ε in the interval $\varepsilon_0 \leq \varepsilon \leq \varepsilon_F$. Below we shall study the change of this dependence on account of interactions of the electrons with inhomogeneities and with each other, the effect of which can be interpreted as virtual transitions into the continuous spectrum. These interactions will be assumed to be short-range and will be taken into account in the Born approximation. The latter restriction can be lifted easily by replacing the Born matrix elements by the corresponding scattering amplitudes (see Ref. 1, Sec. 6).

2. ELECTRON-ELECTRON INTERACTION

In the perturbation-theory series for the one-particle Green function it is necessary to select the diagrams whose imaginary part falls off most slowly with distance from the surface. It can be seen from the expression (7) that diagrams in which it is required to take into account the imaginary

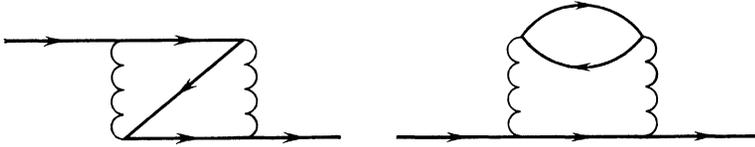


FIG. 2. Principal diagrams describing the electron-electron interaction. The solid lines are the electron G -functions, and the wavy lines are matrix elements of the interactions.

part in the external $G^{(0)}$ -lines decay in accordance with the same law $\exp[-(-2m\hbar^{-2}\varepsilon_0)^{1/2}(z+z')]$, as does the imaginary part of the function $G^{(0)}$ itself. Thus, we must take into account diagrams whose imaginary part does not vanish when, in the external $G^{(0)}$ -lines, only the real parts are retained. The first terms of this kind in the perturbation-theory series are the two second-order diagrams depicted in Fig. 2. In the calculation of the contribution of these diagrams to the imaginary part of the Green function, only the imaginary parts of the internal $G^{(0)}$ -functions are important. The contribution of the second diagram of Fig. 2 is opposite in sign to that of the first diagram, and, because of the summation over the omitted spin indices, has twice the magnitude. The calculation differs from that of the damping of the quasiparticles in a weakly nonideal Fermi gas (see, e.g., Ref. 1, Sec. 21) only in the two-dimensionality of the corresponding integrals over the two internal momenta. Instead of the integration over the third components of the momenta, an integration over the z -coordinates of the two points of collision of the electrons appears. As a result we arrive at the expression

$$\text{Im } G(\mathbf{k}, z, z' | \varepsilon) = \frac{16\pi U_0^2 m \varepsilon_0^2 \varphi(\varepsilon, k)}{\hbar^2 (\hbar^2 k^2 / 2m - \varepsilon) (-9\varepsilon_0 + \varepsilon - \hbar^2 k^2 / 2m)}$$

$$\times \exp[-(k^2 - 2m\hbar^{-2}\varepsilon)^{1/2}(z+z')], \quad (10)$$

where $U_0 = \int U(r) d^3r$ is the Born matrix element of the short-range potential $U(r)$ of the electron-electron interaction. The function

$$\begin{aligned} \varphi(\varepsilon, k) = & \int \{n(\mathbf{k}_1) [1 - n(\mathbf{k}_2)] [1 - n(\mathbf{k} + \mathbf{k}_1 - \mathbf{k}_2)] \\ & - [1 - n(\mathbf{k}_1)] n(\mathbf{k}_2) n(\mathbf{k} + \mathbf{k}_1 - \mathbf{k}_2)\} \delta\left(\varepsilon - \varepsilon_0 + \frac{\hbar^2}{2m} [\mathbf{k}_1^2 + \mathbf{k}_2^2 \right. \\ & \left. - (\mathbf{k} + \mathbf{k}_1 - \mathbf{k}_2)^2\right] (2\pi)^{-4} d^2k_1 d^2k_2, \\ & n(\mathbf{k}) \equiv \theta(\varepsilon_F - \hbar^2 k^2 / 2m - \varepsilon_0), \end{aligned} \quad (11)$$

coincides with the integral appearing in the expression for the damping of quasiparticles in a two-dimensional Fermi gas (compare with the three-dimensional expression [formula (21.17) in Ref. 1]). Since we are interested in energies below the Fermi level, the first term in the curly brackets does not give a contribution to the integral (11). Going over in (10) to the coordinate representation, for a large distance of the points \mathbf{r} and \mathbf{r}' from the surface we can use the method of steepest descent, setting $\mathbf{k} = 0$ everywhere except in the exponents. As a result, we obtain the density matrix in the form

$$\begin{aligned} \rho(\mathbf{r}, \mathbf{r}' | \varepsilon) = & 8 \left(\frac{m}{\pi\hbar^2}\right)^4 U_0^2 \frac{(\varepsilon_F - \varepsilon_0) \Phi[(\varepsilon - \varepsilon_0) / (\varepsilon_F - \varepsilon_0)] \exp[-(-2m\hbar^{-2}\varepsilon)^{1/2}R]}{(9 + \varepsilon/\varepsilon_0)^2 (-2m\hbar^{-2}\varepsilon)^{1/2} R} \\ & R = [(\boldsymbol{\rho} - \boldsymbol{\rho}')^2 + (z + z')^2]^{1/2}, \\ \Phi(\nu) \equiv & \int_{\nu^{1/2}}^1 \{ [1 - (\nu/2\eta)^2]^{1/2} - [1 - (\eta - \nu/2\eta)^2]^{1/2} \} \frac{d\eta}{\eta}, \quad 0 < \nu < 1, \end{aligned} \quad (12)$$

where $\Phi(\nu)$ is the function obtained from the integral (11) for $\mathbf{k} = 0$. As the argument ν changes from zero to unity (i.e., the energy ε changes from the bottom ε_0 to ε_F), $\Phi(\nu)$ decreases from $1 - \ln 2 \approx 0.39$ to zero ($\Phi(\nu) \approx (1 - \nu)^2 / 3^{1/2} \cdot 4$ for $1 - \nu \ll 1$).

3. SCATTERING BY INHOMOGENEITIES

The tunneling of electrons in the presence of sub-barrier elastic scattering by randomly distributed identical point defects with a small concentration was investigated by Lifshitz and Kirpichenkov.² These authors considered normal incidence of electrons on a potential barrier of infinite extent in two dimensions and of rectangular form in the third dimension. Two characteristic ranges of energy of the tunneling

electrons were distinguished: a resonance region near the impurity levels, and a nonresonance region far from the impurity levels.

In application to the problem considered here, the theory of Ref. 2 describes tunneling of electrons from the bottom of the surface band. In nonresonance conditions the effect of the impurities is that multiple scattering causes a linear (in the small concentration of scatterers) change in the decay constant describing the decay of the electron density in the interior of the forbidden region. However, for electrons with a nonzero kinetic energy of motion along the surface,¹⁾ for the reasons outlined in the Introduction the decay constant should experience a decrease of zeroth order in the concentration. This change is already discernible in

the first (linear in the concentration of scatterers) correction to the density matrix; this correction will be calculated here. Allowance for multiple scattering leads only to a small additional change in the decay constant—a change that is linear in the concentration and will not be taken into account here. The situation is analogous to a certain extent to that considered by Shklovskii and Éfros,³ who showed that in a strong magnetic field scattering by impurities qualitatively changes the character of electron tunneling.

For the Green function we shall consider the perturbation-theory series in the scattering amplitude for scattering by defects. Upon averaging over the positions of the scatterers, we obtain a technique similar to the usual “cross” technique (see Ref. 4, Sec. 39). The integration over the z -coordinates of the impurities should be carried out explicitly. It can be verified that the contributions of the diagrams in which each scatterer is encountered once decay in accordance with the same law as $G^{(0)}(\mathbf{k}, z, z' | \varepsilon)$. The first diagram giving a slowly decaying contribution to the averaged density matrix $\langle \rho(\mathbf{r}, \mathbf{r}' | \varepsilon) \rangle$ is the second-order diagram in which both scatterings are by the same impurity (Fig. 3). In the external $G^{(0)}$ -lines we need take into account only the real part, and in the middle $G^{(0)}$ -line, only the imaginary part. For the density matrix far from the surface we obtain

$$\langle \rho(\mathbf{r}, \mathbf{r}' | \varepsilon) \rangle = \frac{n}{4\pi} \left(\frac{mU_0}{\hbar^2} \right)^2 \frac{(-2m\hbar^{-2}\varepsilon)^{1/2}}{[(-\varepsilon_0)^{1/2} - (-\varepsilon)^{1/2}]^2} \frac{\exp[-(-2m\hbar^{-2}\varepsilon)^{1/2}R]}{R}, \quad (13)$$

where n is the concentration of impurities, $U_0 = \int U(\mathbf{r}) d^3r$ is the integral of the potential $U(\mathbf{r})$ of a single impurity, and R is the same as in (12). But this expression describes scattering in the bulk by a random Gaussian potential $\tilde{U}(\mathbf{r})$ of the “white noise” type with correlator

$$\langle \tilde{U}(\mathbf{r}) \tilde{U}(\mathbf{r}') \rangle = nU_0^2 \delta(\mathbf{r} - \mathbf{r}').$$

Besides scattering in the bulk of the semiconductor, scattering by inhomogeneities of the two-dimensional channel itself can be important for the tunneling. We shall treat the latter inhomogeneities as small fluctuations of the channel depth that lead to a dependence of the energy ε_0 of the bottom of the surface band on the coordinates ρ in the plane of the channel:

$$\varepsilon_0(\rho) = \bar{\varepsilon}_0 + \bar{\varepsilon}_0(\rho), \quad \langle \bar{\varepsilon}_0(\rho) \rangle = 0.$$

Instead of taking this dependence into account in the boundary condition (2), it is more convenient, from the point of



FIG. 3. Cross-technique diagram describing the linear (in the impurity concentration) contribution to the electron Green function. The crosses denote matrix elements of the impurity potential. The dashed line shows that the two crosses correspond to the same impurity.

view of application of perturbation theory, to replace it by a fluctuating surface potential

$$\tilde{U}(\mathbf{r}) = \frac{\bar{\varepsilon}_0(\rho)}{2(-2m\hbar^{-2}\varepsilon_0)^{1/2}} \delta(z+0).$$

The main contribution to the density matrix is given, as in the case of scattering by impurities, by the second order of perturbation theory in the potential $\tilde{U}(\mathbf{r})$. Calculating this contribution in analogy with the calculation of the diagram of Fig. 3, we obtain

$$\langle \rho(\mathbf{r}, \mathbf{r}' | \varepsilon) \rangle = \frac{m^2}{4\pi\hbar^4} \frac{Q([2m\hbar^{-2}(\varepsilon - \varepsilon_0)]^{1/2})}{[(-\varepsilon_0)^{1/2} - (-\varepsilon)^{1/2}]^2} \frac{\exp[-(-2m\hbar^{-2}\varepsilon)^{1/2}R]}{R}, \quad (14)$$

where

$$Q(k) = \int \langle \bar{\varepsilon}_0(\rho) \bar{\varepsilon}_0(\rho') \rangle \exp[-ik(\rho - \rho')] d^2\rho$$

is a Fourier component of the correlator of the binding energy, and R is the same as in (12).

4. INTERVALLEY SCATTERING

We now consider a semiconductor whose conduction band has several identical valleys with an anisotropic effective mass. To be specific, we shall have in mind silicon, the six valleys of which in k -space have a larger longitudinal effective mass m_{\parallel} and a smaller transverse effective mass m_{\perp} and are oriented in pairs along mutually perpendicular axes of a cubic lattice.

In itself, the anisotropy of the effective mass, without allowance for intervalley transitions, does not introduce anything fundamentally new into the effects considered above. This is fairly obvious, because of the possibility of making the effective mass in any of the valleys isotropic by means of a linear change of variables; the surface remains planar, and the problem essentially reduces to those already considered. It is also obvious that intervalley transitions are unimportant when the orientations of the valleys are equivalent with respect to the surface. For example, in silicon all the valleys make the same angle with the $\langle 111 \rangle$ surface, and transitions from valley to valley cannot facilitate tunneling.

When the orientations of the valleys are not equivalent with respect to the surface, the deepest surface band is formed on account of the valleys with the largest mass in the direction perpendicular to the surface, while for tunneling it is favorable to have the smallest mass in this direction. Therefore, virtual transitions from valleys forming the surface band to other valleys should exponentially increase the probability of tunneling. We note that in contrast to the one-valley case this increase should occur even for electrons tunneling from the bottom of the surface band.

As is well known, intervalley transitions are made difficult because of the large magnitude of the wave vector separating the valleys. Intervalley scattering by inhomogeneities is quantitatively small in comparison with intravalley scattering, and, therefore, in the treatment of intervalley scattering fundamental complications do not arise. They do appear,

however, when we consider the electron-electron interaction, which conserves the total momentum of the electrons. For real values of the surface Fermi wave vector k_F , an interval transition is possible only as a result of many-particle processes in which an electron, crossing over into a favorable valley, gathers momentum from several electrons remaining in the surface band.

We shall consider the model of the n -channel on the (001) surface of silicon. We shall give the valleys oriented along the x axis the labels 1 and 1', those along the y axis the labels 2 and 2', and those along z axis the labels 3 and 3', and provide the Green functions²⁾ with the corresponding index α . The surface band (a double band, when the valley-orbital splitting is neglected) is formed from the valleys 3 and 3' with their axes perpendicular to the surface. The Green functions of the zeroth approximation have for these valleys the form (compare with the expression (7) with an isotropic effective mass)

$$G_{\alpha}^{(0)}(\mathbf{k}, z, z' | \varepsilon) = \frac{m_{\parallel}^{1/2}}{\hbar^2 (k^2/m_{\perp} - 2\varepsilon/\hbar^2)^{1/2}} \left\{ \exp \left[- \left(m_{\parallel} \left(\frac{k^2}{m_{\perp}} - \frac{2\varepsilon}{\hbar^2} \right) \right)^{1/2} |z-z'| \right] - \frac{[(\hbar^2 k^2/2m_{\perp} - \varepsilon) + (-\varepsilon_0)^{1/2}]^2}{\varepsilon - \varepsilon_0 - \hbar^2 k^2/2m_{\perp} + i0 \operatorname{sign}(k - k_F)} \right. \\ \left. \times \exp \left[- \left(m_{\parallel} \left(\frac{k^2}{m_{\perp}} - \frac{2\varepsilon}{\hbar^2} \right) \right)^{1/2} (z+z') \right] \right\}, \quad (15)$$

$$\alpha = 3, 3',$$

$$k_F = [2m_{\perp} \hbar^{-2} (\varepsilon_F - \varepsilon_0)]^{1/2}.$$

The narrow and deep well $D(z)$ (see Fig. 1), which has for a particle of mass m_{\parallel} a shallow (in comparison with the depth of the well) level ε_0 , practically does not act on a particle of the appreciably lower mass m_{\perp} ; this corresponds to the boundary condition $\Psi = 0$ on the surface. For simplicity, this is the condition we shall adopt for the valleys parallel to the surface, neglecting the possible difference in the effective surface potential for these valleys. With this assumption, for the valleys parallel to the surface we obtain purely real $G^{(0)}$ -functions:

$$G_{\alpha}^{(0)}(\mathbf{k}, z, z' | \varepsilon) = \frac{m_{\perp}^{1/2}}{\hbar^2 (k_x^2/m_{\parallel} + k_y^2/m_{\perp} - 2\varepsilon/\hbar^2)^{1/2}} \times \left\{ \exp \left[- \left(m_{\perp} \left(\frac{k_x^2}{m_{\parallel}} + \frac{k_y^2}{m_{\perp}} - \frac{2\varepsilon}{\hbar^2} \right) \right)^{1/2} |z-z'| \right] - \exp \left[- \left(m_{\perp} \left(\frac{k_x^2}{m_{\parallel}} + \frac{k_y^2}{m_{\perp}} - \frac{2\varepsilon}{\hbar^2} \right) \right)^{1/2} (z+z') \right] \right\}, \quad \alpha = 1, 1';$$

$$(16)$$

for $\alpha = 2, 2'$ we must interchange k_x and k_y . The two-dimensional wave vector \mathbf{k} for each α is measured from the bottom of the corresponding valley.

The electron density matrix in the absence of interaction is obtained from the Green function (15) of the valleys perpendicular to the surface. Because of the anisotropy of the effective mass, the expression (9) is somewhat modified and takes the form

$$\rho_{\alpha}^{(0)}(\mathbf{r}, \mathbf{r}' | \varepsilon) = \theta(\varepsilon_F - \varepsilon) \theta(\varepsilon - \varepsilon_0) (-2m_{\parallel} \varepsilon_0)^{1/2} / \pi \hbar^3 \\ \times J_0 \left((2m_{\perp} \hbar^{-2} (\varepsilon - \varepsilon_0))^{1/2} |\mathbf{r} - \mathbf{r}'| \right) \\ \times \exp \left[- (2m_{\parallel} \hbar^{-2} \varepsilon_0)^{1/2} (z+z') \right], \quad \alpha = 3, 3'. \quad (17)$$

The components ρ_{α} corresponding to the valleys parallel to the surface ($\alpha = 1, 1', 2, 2'$) are equal to zero.

It is now necessary to calculate that correction to the Green function which is linear in the concentration and is described by the same diagram (Fig. 3) as in the one-valley case, by taking for the internal line the $G^{(0)}$ -function (15), and for the external lines the $G^{(0)}$ -functions (16) corresponding to the valleys parallel to the surface. For the Born matrix element U_0 we must now use the matrix element of the full Hamiltonian of the crystal plus impurity between the Bloch functions of the bottoms of different valleys; here we are interested only in transitions from valleys perpendicular to the surface to valleys parallel to the surface. The density matrix averaged over the distribution of impurities has the form

$$\langle \rho_{\alpha}(\mathbf{r}, \mathbf{r}' | \varepsilon) \rangle = \frac{n}{2\pi} \left(\frac{m_{\perp} U_0}{\hbar^2} \right)^2 \frac{m_{\perp} (-2m_{\parallel} \hbar^{-2} \varepsilon)^{1/2} \exp \left[- (-2m_{\perp} \hbar^{-2} \varepsilon)^{1/2} R \right]}{m_{\perp} \varepsilon - m_{\parallel} \varepsilon_0 R} \\ R = \left[\frac{m_{\parallel}}{m_{\perp}} (x-x')^2 + (y-y')^2 + (z+z')^2 \right]^{1/2} \quad (18)$$

for $\alpha = 1, 1'$; for $\alpha = 2, 2'$ we must interchange x and y . In the expression (18) we have taken into account a factor 2, which arises from the double character of the surface band. Comparison of (18) with (17) shows that besides the tunneling-probability increase on account of the replacement of ε_0 by ε in the exponent, as in the one-valley case, there is an additional increase on account of the replacement of the larger mass m_{\parallel} by the smaller mass m_{\perp} . At large distances from the surface the electrons from the valleys parallel to the surface should dominate, and this can be manifested, e.g., in the polarization of recombination emission at holes far from the surface.

We shall say a few words about the hole channel for a degenerate valence band. Despite the presence of two masses in the bulk dispersion law in this case, a strong exponential effect analogous to that in the many-valley case is impossible here. In the wave functions of the surface states the components corresponding to light and heavy holes are mixed together (an exception is the point $\mathbf{k} = 0$), and the decay constant of the hole density now contains the mass of the light hole. We may expect only a weaker effect, similar to that which arises in the nondegenerate one-valley case. Observation of this effect is complicated by the fact that, unlike the electron density in the n -channel, the hole density depends on the energy even in the absence of scattering.

CONCLUSION

The illustrative examples considered above confirm the qualitative considerations, outlined in the Introduction, concerning the influence of different scattering processes on the tunneling of carriers from the surface band. It is possible, apparently, to formulate the following general rule, valid for

an arbitrary (degenerate, or taking account of spin-orbit interaction, or nonparabolic) bulk dispersion law $\varepsilon(\mathbf{k}, k_z)$ for the carriers, and for an arbitrary form of the potential $V(z)$.

At large distances from the channel the density matrix should have the form

$$\rho(\mathbf{r}, \mathbf{r}' | \varepsilon) \propto \exp [i\hbar^{-1}S(\varepsilon)],$$

where $S(\varepsilon)$ is the minimum (in magnitude) value of the purely imaginary truncated action $\int \mathbf{p} \cdot d\mathbf{r}$ on a trajectory going from the point \mathbf{r}' to a certain point on the surface, and from the latter point to the point \mathbf{r} , with a specified value of the total energy ε . Tunneling corresponds to imaginary values of the momentum and of the time of motion, and this is why the action is imaginary. In the case of a degenerate dispersion law, we should take as the kinetic energy its light branch. For a quadratic dispersion law (in the case of a degenerate band the law may be "corrugated"), by a change of variables we can reduce the problem to a more visualizable form by considering the motion of a particle in real time but with a reversed sign of the energy ε and potential $V(z)$. In explicit form, the action can be written for a nondegenerate quadratic dispersion law (see Ref. 5, Sec. 44)

$$S(\varepsilon) = \int \{2[V(z) - \varepsilon] m_{\alpha\beta} dr_\alpha dr_\beta\}^{1/2}, \quad (19)$$

where $m_{\alpha\beta}$ is the effective-mass tensor. The examples considered above correspond to the absence of band bending ($V(z) \equiv 0$) and to a diagonal effective-mass tensor, when minimization of the expression (19) leads to a rectilinear trajectory that is specularly reflected from the surface. In the expressions (12)–(14) the length of this path appears directly in the exponent, while the expression (18) contains the path length corrected for the mass anisotropy.

The rule formulated gives a complete description of the principal, exponential effect, which does not depend on the form or strength of factors violating the separation of variables in the channel, and is entirely general. The weakness of the scattering mechanisms leads only to pre-exponential smallness of the corrections to the unperturbed density matrix and essentially effects only the value of the distance from the surface at which these corrections begin to dominate on account of the more slowly decaying exponential. This distance depends on the energy of the tunneling carriers: It is smaller the closer is the energy to the bottom of the bulk band, and increases (in the one-valley case, to infinity) as the energy approaches the bottom of the two-dimensional surface band.

The energy dependence of the pre-exponential factor is determined by the concrete scattering mechanism. In all the examples considered above, the qualitative character of this dependence is the same: The pre-exponential factor decreases monotonically with increase of the energy from the bottom of the surface band to ε_F . This character of the dependence is evidently a property of the models used. In the general case, when one uses the complete energy-dependent scattering amplitude instead of the Born matrix elements, the energy dependence of the pre-exponential factor in the density matrix can vary greatly. In the presence in the semi-

conductor of scatterers having bound states with levels in the region of the surface-band energies, the exponential dependences can change even as a result of the resonance-tunneling effect.²

We shall say a few words about surface states with energies above the bottom of the bulk band (see Fig. 1). On an ideal surface of a pure crystal in the one-electron approximation these states do not differ in any essential way from states lying in the forbidden band of the semiconductor. But the effect of inhomogeneities and electron-electron interaction on the different types of surface states is qualitatively different. As the energy of a surface state moves downward toward the bottom of the bulk band the state becomes more and more delocalized, and its wave function decays more and more slowly in the interior of the crystal. When the state passes through the bottom of the bulk band the virtual transitions responsible for the tunneling are transformed into real Auger processes and elastic scattering, and the surface state becomes a decay (quasi-stationary) state.

The effects considered can be manifested experimentally in different processes involving the tunneling of carriers, and primarily in the recombination of carriers with carriers of the opposite sign localized outside the channel. At greater distances of the recombination centers from the surface the rectangular shape of the luminescence line should be deformed on account of the slower decay of its blue edge. The one-particle density matrix (1) can be used in the case when the carriers of the opposite sign are localized at sufficiently deep centers, when the tunneling of only the surface carriers need be taken into account. In the simplest case of direct, zero-phonon recombination at deep centers in an allowed optical transition, the shape of the recombination-emission line is determined by the energy dependence of the electron concentration $n(\mathbf{r}, \varepsilon) = \rho(\mathbf{r}, \mathbf{r} | \varepsilon)$.

Very recently, experimental observations^{6,7} of recombination radiation of electrons from two-dimensional channels have appeared; in the MIS system of Ref. 6 a rectangular line shape is observed, while in the heterostructure of Ref. 7 a deformed line shape is observed. Evidently, this can be understood in the light of the results of the present work, since the concentration of impurities in the heterostructure of Ref. 7 is higher, and the tunneling length in this experiment greater, than in the MIS system of Ref. 6.

The idea of the present work arose in a discussion of the results of the experiment of Ref. 7 with V. D. Kulakovskii, to whom the author expresses his gratitude.

¹ In the formulation in Ref. 2, this would correspond to oblique incidence of electrons on the barrier.

² Scattering by inhomogeneities destroys the diagonal character of the one-particle Green function in the valley indices, but the contribution of interest to us is diagonal after averaging over the positions of the scatterers.

¹E. M. Lifshitz and L. P. Pitaevskii, *Statisticheskaya fizika* (Statistical Physics), Part 2, Nauka, Moscow (1978) [English translation published by Pergamon Press, Oxford (1980)].

²I. M. Lifshitz and V. Ya. Kirpichenkov, *Zh. Eksp. Teor. Fiz.* 77, 989

(1979) [Sov. Phys. JETP **50**, 499 (1979)].

³B. I. Shklovskii and A. L. Éfros, Zh. Eksp. Teor. Fiz. **84**, 811 (1983) [Sov. Phys. JETP **57**, 470 (1983)].

⁴A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Metody kvantovoi teorii polya v statisticheskoi fizike* (Quantum Field Theoretical Methods in Statistical Physics), Fizmatgiz, Moscow (1962) [English translation published by Pergamon Press, Oxford (1965)].

⁵L. D. Landau and E. M. Lifshitz, *Mekhanika* (Mechanics), Nauka, Mos-

cow (1973) [English translation published by Pergamon Press, Oxford (1976)].

⁶I. V. Kukushkin and V. B. Timofeev, Pis'ma Zh. Eksp. Teor. Fiz. **40**, 413 (1984) [JETP Lett. **40**, 1231 (1984)].

⁷V. D. Kulakovskii, B. N. Shepel', A. A. Denisov, and A. P. Senichkin, Fiz. Tekh. Poluprovodn. **21**, (1987) [sic].

Translated by P. J. Shepherd