Conductivity of inhomogeneous surface channels

A. Ya. Shik

A. F. Ioffe Physico-technical Institute, USSR Academy of Sciences
(Submitted November 13, 1975)
Zh. Eksp. Teor. Fiz. **70**, 2211–2217 (June 1976)

The carrier concentration and conductivity in an inhomogeneous inversion channel near the surface of a semiconductor are calculated. It is assumed that the inhomogeneities are due to random distribution of the surface charges and are sufficiently smooth to permit one to apply the percolation-theory concepts and to neglect tunneling. It is shown that allowance for the inhomogeneities can explain some of the singularities of the dependences of the kinetic coefficients in the channel on temperature and band bending. Various methods for determining mobilities in the channels are considered and it is shown that in the inhomogeneous case they lead to significantly different results.

PACS numbers: 72.20.Dp

1. INTRODUCTION

In an experimental investigation of kinetic phenomena in semiconductor surface channels one frequently observes two singularities that are explained neither by the classical[1] nor by the quantum[2] theory of surface conductivity. The first recalls the Mott transition. The surface conductivity σ_s and the Hall mobility in the channel μ_H, which depend activationally on the temperature in the case of small band bending, decrease the activation energy with increase of the bending and ultimately acquire a "metallic" temperature-independent character.[3–7] In 1973, Mott[8] had proposed that the aforementioned transition is attributed to localization of the carriers in potential wells that are connected with inhomogeneities of the surface channel. The second singularity consists in the presence of a sharp maximum on the dependence of μ_H on the band bending.[3,7,9–13] Hypotheses have also been advanced in the literature concerning the connection of this effect with inhomogeneities of the channel,[7,10–12] but no consistent theory has been constructed.

In this paper we construct a theory of the conductivity of an inhomogeneous surface channel; this theory explains, in particular, the effects indicated above.

2. CALCULATION OF THE EQUILIBRIUM FORM OF THE CHANNEL. THE "METAL–INSULATOR" TRANSITION

We consider, for the sake of argument, a surface inversion channel in a p-type semiconductor. The calculations of the equilibrium form of such a channel have been the subject of a considerable number of studies. In most of them, however, the answer is either a very complicated expression or the result of a numerical calculation. We carry out an approximate self-consistent calculation, which leads to somewhat less accurate but on the other hand simple formulas that can be generalized to include the inhomogeneous case.

Assume that the bottom of the conduction band on the surface of the semiconductor has an energy $V_0<0$ (the origin is taken to be the position of the bottom of the band in the interior of the sample). The electric field on the surface is then

$$E=-\frac{4\pi}{\varkappa}Q+2\left(\frac{2\pi N|V_0|}{\varkappa}\right)^{1/2}+\frac{4\pi e}{\varkappa}n_s. \tag{1}$$

Here Q is the charge density on the surface states, N is the acceptor concentration, $\varkappa$ is the dielectric constant of the semiconductor, and n_S is the surface concentration of the electrons in the channel.[1)]

We assume that $Q=\text{const}(V_0)$. In the study of the properties of the channel in the vicinity of the "metal–semiconductor" transition, this assumption is valid, for it will be shown below that the level of the chemical potential passes in this case near the bottom of the conduction band on the surface, and the greater part of the surface states remains lower than this level, and their occupation does not change, even V_0 is varied in a relatively wide range.

To calculate n_S we assume that the bulk of the electrons is concentrated in that surface region in which the potential energy can be regarded as linear: $V(z)=V_0+eEz$. For a degenerate electron gas we then have (the nondegenerate case will be considered in Sec. 3)

$$n_s=\frac{2^{1/2}m^{3/2}}{3\pi^2\hbar^3}\int_0^{z_0}(\zeta-V(z))^{3/2}dz=\frac{2^{5/2}}{15\pi^2}\frac{m^{3/2}}{\hbar^3eE}(\zeta-V_0)^{5/2}\theta(\zeta-V_0), \tag{2}$$
$$z_0=(\zeta-V_0)/eE.$$

if the effects of quantization in a channel are insignificant and

$$n_S=\frac{m}{\pi\hbar^2}(\zeta-V_0-\varepsilon_1)\theta(\zeta-V_0-\varepsilon_1) \tag{3}$$

for a quantum channel with one filled level. Here ζ is the chemical-potential level ($\zeta<0$),

$$\varepsilon_1=1.86(e^2E^2\hbar^2/m)^{1/3}$$

is the energy of the first quantum level, and $\theta(x)$ is the unit step function. We shall henceforth neglect ε_1, since in inversion channels ($|\zeta|\sim\varepsilon_g$, where ε_g is the width of the forbidden band) it is small in comparison with $|\zeta|$ and its dependence on V_0 is weaker than linear.[14]

Formulas (1) and (2) [or (1) and (3)] give the connection between the quantities V_0 and E. It must be recognized, however, that in the experiment the given quantity

is not V_0 and not E, but the voltage φ on the gap, which is connected with V_0 and E by the expression

$$e\varphi = -V_0 + \varkappa eEd/\varkappa_D, \tag{4}$$

where d and $\varkappa_D$ are the thickness and dielectric constant of the insulator layer. All the characteristics of the channel will henceforth be expressed in terms of φ.

We proceed now to the inhomogeneous channel. As already noted,[15-18] the inhomogeneity of the channel usually gives rise to an inhomogeneously distributed surface charge. The quantity Q in (1) will therefore be regarded as a random function of the coordinates x and y. Since φ is the same at all points of the surface, the quantities V_0, E, and n_S, just as Q, fluctuate from point to point. If the correlation radius R_Q of Q greatly exceeds the insulator thickness d, the mean free path l, and the screening radius r_{scr} in the channel, then the local values of Q, V_0, E, and n_S are related as before by means of formulas (1)–(4).

At zero temperature, the electron concentration in the channel is equal to zero in regions where $V_0 > \zeta$, i.e.,

$$Q < \left(\frac{\varkappa}{2\pi} N|\zeta|\right)^{1/2} + \frac{\varkappa_D}{4\pi d}\left(\frac{1}{e}\zeta + \varphi\right)$$

("semiconducting" regions) and differs from zero at larger Q ("metallic" regions). In accordance with the conclusions of two-dimensional percolation theory,[19] the channel is conducting if the relative area of the metallic regions exceeds 50% (at least for a symmetrical distribution function Q). This occurs if the voltage on the gap is larger than critical

$$\varphi_{cr} = \frac{|\zeta|}{e} + \frac{d}{\varkappa_D}\{-4\pi\bar{Q} + 2(2\pi\varkappa N|\zeta|)^{1/2}\}. \tag{5}$$

Here $\bar{Q}$ is the average value of the random function Q.

3. CARRIER DENSITY IN INHOMOGENEOUS CHANNEL

We investigate the variation of the electron density in the channel near the "metal–semiconductor" transition. To this end we calculate first the distribution function of the random quantity V_0. We confine ourselves to the ultraquantum limit, which is realized in an appreciable fraction of the experiments on inversion channels.[14,20] Since the electric field of the depletion layer [the second term of (1)] depends on V_0 less than the other terms of (4), we shall assume the approximation $V_0 \approx \zeta$ in the indicated term. In this case V_0 and Q are connected by the linear relation

$$Q = \frac{em}{\pi\hbar^2}(\zeta - V_0)\theta(\zeta - V_0) + \left(\frac{\varkappa N|\zeta|}{2\pi}\right)^{1/2} - \frac{\varkappa_D(V_0 + e\varphi)}{4\pi ed}. \tag{6}$$

If $Q(x, y)$ is a Gaussian random function with mean value $\bar{Q} = 0$ and variance ΔQ, then $V_0(x, y)$ is a combination of two Gaussian functions: at $V_0 < \zeta$—with mean value

$$\bar{V}_0^- = \zeta - \frac{e(\varphi - \varphi_{cr})}{1 + 4e^2md/\varkappa_D\hbar^2}$$

and variance

$$\Delta V_0^- = \frac{4\pi ed\Delta Q}{\varkappa_D(1 + 4e^2md/\varkappa_D\hbar^2)},$$

and at $V_0 > \zeta$—with

$$\bar{V}_0^+ = \zeta - e(\varphi - \varphi_{cr}), \quad \Delta V_0^+ = 4\pi ed\Delta Q/\varkappa_D.$$

It is seen that the random function V_0 is strongly asymmetrical: the minima of $V_0(x, y)$, which contain the electrons, have a lower amplitude than the maxima, owing to the screening by the electrons.

With the aid of (3) and the described distribution function for V_0 we can calculate the surface electron density in the channel at $T = 0$:

$$\bar{n}_S = \frac{m}{\pi\hbar^2}\left\{\frac{\Delta V_0^-}{(2\pi)^{1/2}}\exp\left[-\frac{(\zeta - \bar{V}_0^-)^2}{2(\Delta V_0^-)^2}\right] + \frac{\zeta - \bar{V}_0^-}{2}\left[1 - \Phi\left(-\frac{\zeta - \bar{V}_0^-}{\sqrt{2}\,\Delta V_0^-}\right)\right]\right\} \tag{7}$$

($\Phi(x)$ is the probablity integral). It is seen the n_S has no singularities at $\varphi = \varphi_{cr}$ and remains different from zero at lower values of φ.

However, the conductivity of the inhomogeneous channel is determined not by the value of n_S but by the carrier density (which we denote by n_S^*), with an energy exceeding the "percolation level" ε_{per}.[19] The latter, as already mentioned, should divide the potential relief on the $V_0(x, y)$ surface into two equally probable parts, from which it follows that

$$\varepsilon_{per} = \begin{cases} \bar{V}_0^- & \text{if} \quad \varphi > \varphi_{cr} \\ \bar{V}_0^+ & \text{if} \quad \varphi < \varphi_{cr} \end{cases}. \tag{8}$$

It is easily seen from (8) that in the semiconducting region $\varphi < \varphi_{cr}$ the temperature dependence of n_S^* has an activation character with an activation energy $e(\varphi_{cr} - \varphi)$. This conclusion is valid for $T \ll \Delta V_0^-$ and not too small φ (the limitation on φ is given in Sec. 4 below), when the electron gas can be regarded as degenerate.

With further decrease of φ, the electrons in the channel become nondegenerate. Their density at small φ is low enough, and in (1) we can neglect the term with n_S, after which (4) turns into an equation quadratic in $|V_0|^{1/2}$ and can be easily solved:

$$V_0 = -\left[-\frac{(2\pi\varkappa N)^{1/2}ed}{\varkappa_D} + \left(\frac{2\pi\varkappa Ne^2d^2}{\varkappa_D^2} + \frac{4\pi edQ}{\varkappa_D} + e\varphi\right)^{1/2}\right]^2. \tag{9}$$

Now, knowing the function $V_0(Q)$, we can calculate the average carrier density, which is defined in the nondegerate case by the formula

$$\bar{n}_S = \frac{mT}{\pi(2\pi)^{1/2}\hbar^2\Delta Q}\int_{-\infty}^{\infty} dQ \exp\left[-\frac{Q^2}{2(\Delta Q)^2} + \frac{\zeta - V_0}{T}\right]. \tag{10}$$

The integral in (10) can be calculated with the aid of the asymptotic Laplace formula. As a result we obtain

$$\bar{n}_S \sim \exp\left[\frac{\zeta + e\varphi}{T} + \frac{8\pi^2e^2d^2(\Delta Q)^2}{\varkappa_D^2T^2}\right] \tag{11}$$

at $\varkappa_D^2\varphi/2\pi\varkappa ed^2N \equiv \varphi/\varphi_0 \gg 1$;

$$\bar{n}_S \sim \exp\left[\frac{\zeta}{T} + \frac{8\pi^2e^2d^2(\Delta Q)^2}{\varkappa_D^2T^2}\right] \tag{12}$$

at $\varphi/\varphi_0 \ll 1$, and $8\pi(\Delta Q)^2/\varkappa TN \equiv \delta \gg 1$;

$$\bar{n}_s \sim \exp\left[\frac{\zeta}{T} + \frac{1}{16\pi}\frac{\varkappa_D^2\varphi^2}{\varkappa N d^2 T}\left(1+\frac{1}{4}\delta\right)\right] \tag{13}$$

at $\varphi/\varphi_0 \ll 1$, and $\delta \ll 1$.

Since $\varepsilon_{per} = V_0(\bar{Q})$, the effective density, which determines the conductivity is

$$n_s^* = \frac{mT}{\pi\hbar^2}\exp\left[\frac{\zeta - V_0(\bar{Q})}{T}\right] \sim \begin{cases} \exp\left(\dfrac{\zeta+e\varphi}{T}\right) & \text{if} \quad \varphi \gg \varphi_0; \\ \exp\left(\dfrac{\zeta}{T} + \dfrac{\varkappa_D^2\varphi^2}{16\pi\varkappa N d^2 T}\right) & \text{if} \quad \varphi \ll \varphi_0. \end{cases} \tag{14}$$

4. SURFACE MOBILITY

There are many methods of determining the surface mobility. In the homogeneous case all lead to results that differ by not more than a small numerical factor. In the inhomogeneous channel, to the contrary, the "mobilities" determined by different methods can, as we shall see, not only differ strongly quantitatively, but also have qualitatively different dependences on the temperature and on the band bending. The question of the mobility in the inhomogeneous channel must therefore be considered separately for each of the methods used for its determination.

A. True (drift) mobility μ_0. The surface mobility in homogeneous channels was investigated theoretically by many workers. It is lower than the mobility in a bulky samples because of the additional scattering from the surface. The role of this scattering increases with decreasing width, so that μ_0 decreases monotonically with increasing bending of the bands. The temperature dependence of μ_0 is determined by the dominant scattering mechanism, but is always given by a power-law function. In an inhomogeneous channel with smooth inhomogeneities, μ_0 depends on the coordinates but its local values have similar properties.

B. Hall mobility μ_H. The calculations for certain simpler models and a number of experimental results (see[21]) show that in samples with inhomogeneous carrier density the Hall constant is determined by the concentration averaged over the sample. It follows therefore that the effective Hall mobility in an inhomogeneous channel is

$$\mu_H \sim \bar{\mu}_0 n_s^*/\bar{n}_s, \tag{15}$$

where $\bar{\mu}_0$ is a certain mean value of μ_0. Consequently, at $\varphi > \varphi_{cr}$ the value of μ_H differs little from the mobility in the homogeneous channel and, just as in the latter, is a power-law function of the temperature, while at $\varphi < \varphi_{cr}$ (by virtue of the appreciable difference between n_S^* and n_S) μ_H decreases sharply and acquires an activation temperature dependence, as observed in the studies cited in the Introduction.

Systematizing the foregoing results, we can trace the dependence of μ_H on the temperature and on the gap voltage in the entire region $\varphi < \varphi_{cr}$. In the case of strong inhomogeneities ($\delta \gg 1$) we have for $0 < \varphi_{cr} - \varphi < 8\pi^2 ed^2 (\Delta Q)^2/\varkappa_D^2 T$

$$\mu_H \sim \bar{\mu}_0 \exp[-e(\varphi_{cr}-\varphi)/T], \tag{16}$$

and for smaller values of φ the mobility μ_H no longer depends on the gap voltage:

$$\mu_H \sim \bar{\mu}_0 \exp[-2(2\pi ed\Delta Q/\varkappa_D T)^2]. \tag{17}$$

For weak inhomogeneities ($\delta \ll 1$), formulas (16) and (17) remain valid only in the region $\varphi > \varphi_0$, while at smaller the value of μ_H again starts to increase, tending to $\bar{\mu}_0$ like

$$\mu_H \sim \bar{\mu}_0 \exp[-2(\varkappa_D \Delta Q\varphi/4dNT)^2]. \tag{18}$$

Indeed, from the condition $\delta \ll 1$ it follows that as $\varphi \to 0$ the fluctuations of Q lead to fluctuations of the surface bending of the bands; the amplitudes of these fluctuations are much lower than the thermal energy, and they therefore exert no influence on the properties of the channel. To the contrary, at $\delta \gg 1$ the screening action of the depletion layer turns out to be insufficient and the potential relief on the surface has an appreciable value, thereby decreasing μ_H significantly.

It is obvious that if $\varphi_0 \gtrsim \varphi_{cr}$ then, for weak inhomogeneities, Eq. (17) is not valid anywhere, and (16) goes over directly into (18). The function $\mu_H(\varphi)$ in the region $0 < \varphi < \varphi_{cr}$ has then an asymmetric minimum near φ_{cr}. At the minimum we have

$$(\mu_H)_{min} \sim \bar{\mu}_0 \exp[-2(\varkappa_D \Delta Q\varphi_{cr}/4dNT)^2]. \tag{19}$$

We recall that all the expressions given above for μ_H were obtained with the aid of the asymptotic Laplace formula, and are therefore valid only when the argument of the exponential is larger than unity in absolute value.

To use the obtained formulas for the reduction of the experimental results it is necessary to determine which of the considered cases (strong or weak inhomogeneity) is realized in the given actual experiment. This is easy to do. A measure of the inhomogeneity parameter ΔQ is the surface carrier density $\bar{n}_{cr}$ (determined from the Hall effect or from capacitive measurements) at the "metal–semiconductor" transition point at $T = 0$. From (7) it follows that

$$\bar{n}_{cr} = \frac{\varkappa_D}{(2\pi)^{1/2}}\left(\frac{\varkappa_D\hbar^2}{4e^2dm}+1\right)^{-1}\frac{\Delta Q}{e}. \tag{20}$$

As a rule, in experiments we have $\varkappa_D\hbar^2/4e^2dm \ll 1$, therefore

$$\delta = 16\pi^2 e^2 \bar{n}_{cr}^2/\varkappa_D^2 TN. \tag{21}$$

In most of the cited papers[3–7] $\bar{n}_{cr} \sim 10^{11}$–$10^{12}$ cm^{-2} and $N \sim 10^{15}$–10^{16} cm^{-3}. The strong-inhomogeneity condition $\delta \gg 1$ is satisfied all the way to room temperatures. Indeed, measurements[7] confirm the presence of saturation of μ_H at small band bending, a saturation predicted by formula (7) but not occurring if $\delta \ll 1$.

We note that the conditions that the fluctuations be smooth, $R_Q \gg l$ and $R_Q \gg r_{scr}$, which were used in the

derivation of the formulas of this paper, were satisfied in the cited experiments. Indeed, according to estimates by Chen and Muller[7], $R_Q=(2-7)\times 10^{-5}$ cm, whereas the mean free path in all the experiments did not exceed 10^{-6} cm, and the screening radius, estimated from the formulas of [22] remained lower than R_Q all the way to $\bar{n}_s \sim 10^9-10^{10}$ cm^{-2}, which is much lower than $\bar{n}_{cr}$.

C. Field-effect mobility μ_{FE}. By definition,

$$\mu_{FE}=\frac{1}{C}\frac{d\sigma_s}{d\varphi}, \quad C=\frac{d}{d\varphi}\left(e\bar{n}_s+\left(\frac{\varkappa N|V_0|}{2}\right)^{1/2}\right), \tag{22}$$

where C is the specific capacitance of the system. At not too small φ, when the principal role in the capacitance is played by the charge of the inversion layer, we have

$$\mu_{FE}\approx\bar{\mu}_0\frac{dn_s^*}{d\varphi}\left(\frac{d\bar{n}_s}{d\varphi}\right)^{-1}. \tag{23}$$

Inasmuch as the functions $n_s^*(\varphi)$ and $\bar{n}_s(\varphi)$ are an exponential[11-14] at $\varphi<\varphi_{cr}$ and $\bar{n}_s \gg n_s$, we can deduce from a comparison of (15) and (23) that μ_{FE} has the same dependence on the temperature and on the band bending as μ_H, but is somewhat larger than the latter. This conclusion is confirmed by the experimental results.[3,9] With further decrease of φ, the value of μ_{FE} decreases exponentially in accordance with the same law (14) as n_s^*.

D. High-frequency mobility $\mu_\sim$. One other possibility of determining the mobility is by using contactless microwave methods (the Faraday effect, cyclotron resonance, etc.). It is obvious that, in contrast to the stationary conductivity, these effects receive contributions from all the carriers in the band, so that $\mu_\sim$ should be close the the true mobility μ_0. Cyclotron resonance experiments indeed show that, in contrast to μ_H, which has a steep rise in the region of small φ, the high-frequency mobility $\mu_\sim$ decreases monotonically with increasing φ. This difference between $\mu_\sim$ and μ_H (observed also in bulky inhomogeneous semiconductors[24]) confirms the conclusion that the anomalies described above are connected not with the behavior of the true mobility, but with the specifics of the conductivity and of the Hall effect in inhomogeneous semiconductors.

The author thanks V. N. Dobrovol'skiĭ, Yu. S. Zharkikh, and O. S. Frolov for discussions.

[1)]In the second term, which is the field of the depletion layer, we have neglected the potential drop across the thickness of the inversion layer.

[1]J. R. Schrieffer, Phys. Rev. **97**, 641 (1955).
[2]F. Stern and W. E. Howard, Phys. Rev. **163**, 816 (1967).
[3]F. F. Fang and A. B. Bowler, Phys. Rev. **169**, 619 (1968).
[4]M. Pepper, S. Pollitt, C. J. Adkins, and R. E. Oakley, Phys. Lett. **47A**, 71 (1974); **48A**, 113 (1974).
[5]D. C. Tsui and S. J. Allen, Phys. Rev. Lett. **32**, 1200 (1974).
[6]E. Arnold, Appl. Phys. Lett. **25**, 705 (1974).
[7]J. T. C. Chen and R. S. Muller, J. Appl. Phys. **45**, 828 (1974).
[8]N. Mott, Electr. and Power **19**, 321 (1973).
[9]N. S. J. Murphy, F. Berz, and I. Flinn, Solid-State Electron. **12**, 775 (1969).
[10]V. N. Dobrovol'skiĭ, Yu. S. Zharkikh, and L. N. Abessonova, Fiz. Tekh. Poluprovodn. **5**, 723 (1971) [Sov. Phys. Semicond. **5**, 633 (1971)].
[11]S. A. Ribeiro and J. C. Pfister, Solid State Commun. **10**, 63 (1972).
[12]A. A. Guzev, G. L. Kurishev, and S. P. Sinitsa, Phys. Status Solidi **a14**, 41 (1972).
[13]V. I. Zolotarev, V. P. Pospelov, E. A. Fetisov, and R. Z. Khafizov, Mikroélektronika **3**, 413 (1974).
[14]J. A. Pals, Philips Res. Rep. Suppl. 7 (1972).
[15]E. H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. **46**, 1055 (1967).
[16]H. Katto and C. T. Sah, Phys. Status Solidi **A13**, 417 (1972).
[17]D. H. Eaton and C. T. Sah, Phys. Status Solidi **A12**, 95 (1972).
[18]T. H. Ning and C. T. Sah, Phys. Rev. **B9**, 527 (1974).
[19]B. I. Shklovskiĭ and A. L. Éfros, Usp. Fiz. Nauk **117**, 401 (1975) [Sov. Phys. Usp. **18**, 845 (1975)].
[20]F. Stern, Phys. Rev. **B5**, 4891 (1972).
[21]L. N. Abessonova, V. N. Dobrovol'skiĭ, Yu. S. Zharkikh, O. S. Frolov, and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **10**, 406 (1976) [Sov. Phys. Semicond. **10**, 244 (1976)].
[22]A. V. Chaplik, Zh. Eksp. Teor. Fiz. **60**, 1845 (1971) [Sov. Phys. JETP **33**, 997 (1971)].
[23]S. J. Allen, D. C. Tsui, and J. V. Dalton, Phys. Rev. Lett. **32**, 107 (1974).
[24]V. T. Potapov, V. I. Trifonov, I. I. Chusov, and N. G. Yaremenko, Fiz. Tekh. Poluprovodn. **6**, 1227 (1972) [Sov. Phys. Semicond. **6**, 1076 (1973)].

Translated by J. G. Adashko