Theory of surface states and of the conductivity in metal-insulator-semiconductor structures

V. A. Gergel' and R. A. Suris

(Submitted 28 July 1982)
Zh. Eksp. Teor. Fiz. **84**, 719–736 (February 1983)

It is shown that the basic properties of carriers in the subsurface layer of the semiconductor in metal-insulator-semiconductor (MIS) structures are determined by electrostatic-potential fluctuations that result from the random distribution of the charge centers contained in the insulator near the interface with the semiconductor. A quantitative theory is developed for the charge of the inversion layer and its conductivity, based on a rigorous and consistent account of the small-scale character of the fluctuating potential in the subsurface region of the semiconductor. The electron states localized at the minimum of the fluctuating potential relief are surface states. Comparison with experiment has demonstrated that it is precisely states of this type which dominate in the measured spectra of real MIS structures based on silicon. Correct allowance for the fluctuation potential has made it possible to explain qualitatively the behavior of the conductivity of the inversion layer near the threshold, find the connection between the conductivity and the charge density of the inversion layer, as well as the dependence of the conduction activation energy on the charge density at low temperatures. The theory constructed demonstrates the essentially three-dimensional character of the electron motion in inversion layers of MIS structures.

PACS numbers: 73.40.Qv, 73.20.Cw

INTRODUCTION

Ideas concerning the motion of electrons in disordered two-dimensional systems are now being intensively developed.[1] The two-dimensional state occupies a position intermediate between the one-dimensional one, where an arbitrarily weak potential leads to complete localization of the states within the confines of the entire electron spectrum,[2] and the three-dimensional one, where localized and delocalized states coexist (the demarcation energy level that separates the corresponding sections of the energy spectrum is called the mobility threshold). This is the cause of the purely cognitive interest in two-dimensional disordered system, in which a stronger manifestation of localization effects than in the three-dimensional case is expected.

No less considerable a stimulus to the study of two-dimensional disordered systems is the suggested practical significance of research of this type. We have in mind the use of the theoretical results to explain the electric characteristics of metal-insulator-semiconductor (MIS) structures that play a dominant role in modern microelectronics. Indeed, electrons located in the MIS-structure inversion layers and quantized by the transverse electric field at the interface between the semiconductor and the (in most cases) amorphous insulator appear at first glance to be an adequate physical model of a two-dimensional disordered electron system, and the circumstance that by varying the voltage on the MIS-structure electrode it is possible to vary the charge of the inversion layer, and hence the position of the Fermi level, makes this structure ideal for the study of phenomena near the mobility threshold (in the vicinity of the so-called metal-insulator junction.)

Owing to the tremendous mathematical difficulties, however, attempts at an abstract analysis of two-dimensional disordered electron systems led to no significant results that might be used for MIS structures. At the same time, there actually exist in the theory of MIS structure a number of "blank spots," principal among which is the question of the surface states on the semiconductor insulator interface. The energy spectrum of such states is determined from the degree of the deviation of the characteristics of real devices with MIS structure from the calculated relations that correspond to an ideal (without surface states) MIS structure. Surface states are, first, nonconducting, so that they cannot be related to the Tamm levels; second, the experimental spectra of the surface states of MIS structures are so continuously "smeared out" over the band gap that they do not resemble even remotely discrete levels initiated by broken bonds (Shockley states) or by foreign atoms on the interface. It is this smearing of the energy spectra that suggests the possible fluctuating nature of the surface states in real MIS structures.

The main part of the present paper is the construction of a theoretical model of fluctuating surface states in MIS structures. We shall investigate also the conductivity of MIS structure in the vicinity of the threshold voltage, and obtain the connection between the conductivity and the charge density of the inversion layer and the corresponding dependence of the conduction activation energy. The source of the disorder will be considered, as before,[3] to be static fluctuations of the density of charged centers contained in the insulator layer of the MIS structure. The crucial aspect that distinguishes the theory proposed here is a definite rejection of two-dimensional concepts as the zeroth approximation of the problem. The point is, as we shall show, that the fluctuating relief of the potential in the subsurface region of the semiconductor of the MIS structure is of such small scale that at those distances from the semiconductor-insulator interface that are significant for electron motions exceed the homogeneous electric field that tends to "two-dimensionalize" the electron

 0038-5646/83/020415-11$04.00

trajectories by crowding the electrons towards the surface. The MIS-structure theory advanced here is *ab initio* three-dimensional.

In the first section of the article we present the results of a study of the characteristics of the fluctuating potential relief in the subsurface of the semiconductor, a relief due to the random distribution of the built-in charge centers in the insulator layer near the interface with the semiconductor. The second section contains an exposition of the theory of surface states in MIS structures. In the third is investigated the conductivity of the inversion layer. The fourth and last section is devoted to a discussion of the results and to their comparison with experiment.

1. CHARACTERISTICS OF A RANDOM POTENTIAL IN THE SUBSURFACE OF THE MIS-STRUCTURE SEMICONDUCTOR. PERCOLATION LEVEL

We shall assume that the semiconductor-insulator interface is ideal in the sense that it has no bare surface states whatever. We assume furthermore that the built-in charge, whose surface density we shall designate $\Sigma(r)$, is located directly on the interface $z=0$ (positive z correspond to the semiconductor region of the MIS structure, $\mathbf{r}$ is the coordinate in the interface plane). We shall be interested in the so-called depletion and inversion regimes, when an external voltage applied to the electrode of the structure repels the majority carriers into the interior of the semiconductors. This produces at the interface a depleted layer whose average thickness is connected in a known manner with the surface potential Φ_s:

$$w=(\varepsilon_s\bar{\Phi}_s/2\pi eN)^{1/2},$$

where N is the impurity density and ε_s is the dielectric constant of the semiconductor. The fluctuations $\delta w = w(\mathbf{r}) - w$ of the depletion depth, which are due to fluctuations of the charge $\Sigma(\mathbf{r})$, are usually small, i.e., $\overline{\delta w^2}/w^2 \ll 1$, and the influence of the inhomogeneity of the depletion depth on the potential relief in the depleted layer can be taken into account with the aid of a fictitious surface charge located in the plane $z=w$ (Ref. 4). The potential in the subsurface region is then the sum of the potential of the homogeneous field of the depleted layer, $\bar{\varphi}_s - E_s z$, where $E_s = 2\bar{\varphi}_s/w$, and of the fluctuating potential $\delta\varphi(\mathbf{r},z)$. The latter is calculated as the potential initiated by the non-uniformly charged layer $\Sigma(\mathbf{r}) - \bar{\Sigma}$ located between the two capacitor electrodes $z_1 = w$ and $z_2 = -d$ (d is the thickness of the insulator layer). In this calculation it is necessary to use in the semiconductor region $z>0$ the Laplace equation rather than the Poisson equation, since the charge of the ionized acceptors (donors) is already "busy" forming the r-independent depleted-layer field E_s. Taking the foregoing into account we obtain for $z>0$

$$\delta\varphi(\mathbf{r},z)=4\pi\int d^2k\,\delta\Sigma(\mathbf{k})\frac{e^{i\mathbf{kr}}}{k}\frac{\operatorname{sh} kd\operatorname{sh} k(w-z)}{\varepsilon_s\operatorname{sh} kd\operatorname{ch} kw+\varepsilon_i\operatorname{ch} kd\operatorname{sh} kw}, \tag{1}$$

where ε_i the dielectric constant of the insulator and

$$\delta\Sigma(\mathbf{k})=(2\pi)^{-2}\int d^2r(\Sigma(\mathbf{r})-\bar{\Sigma})e^{-i\mathbf{kr}} \tag{2}$$

is the Fourier transform of the fluctuating built-in charge.

We obtain now the correlation function of the random potential, assuming completely random distribution of the charged centers on the semiconductor-insulator interface, i.e.,

$$\langle\delta\Sigma(\mathbf{k})\,\delta\Sigma(\mathbf{k}')\rangle=(e/2\pi)^2\sigma\delta(\mathbf{k}+\mathbf{k}'). \tag{3}$$

Here $\sigma=\sigma^+ + \sigma^-$ is the sum of the average densities of the positively and negatively charged centers located on the interface. It follows form (1)–(3) that

$$K(\mathbf{r},z,z')=\langle\delta\varphi(0,z)\,\delta\varphi(\mathbf{r},z')\rangle$$
$$=4e^2\sigma\int d^2k\frac{e^{i\mathbf{kr}}}{k^2}\frac{\operatorname{sh}^2 kd\operatorname{sh} k(w-z)\operatorname{sh} k(w-z')}{(\varepsilon_s\operatorname{sh} kd\operatorname{ch} kw+\varepsilon_i\operatorname{ch} kd\operatorname{sh} kw)^2}. \tag{4}$$

We note that in our problem, in contrast to three-dimensional Coulomb systems, the integrand in (4) does not diverge as $k\to 0$. This absence of the Holtsmark divergence is the consequence of the screening of the charge of the large-scale fluctuations by the image charges induced on the capacitor electrodes. In the region $0<z\ll w$ of interest to us the principal role is played in this screening by the charges induced on the metal, since usually $w\gg d$. Expression (4) can therefore be simplified by letting in it $w\to\infty$. For $z\gg d$ (but, naturally, $z\ll w$) the correlator (4) has then the following asymptotic form

$$K(\mathbf{r},z,z')=8\pi e^2\sigma\frac{d^2}{\varepsilon_i^2}\frac{z+z'}{[r^2+(z+z')^2]^{3/2}}. \tag{5}$$

So sharp a decrease of $K(0,z,z)$ with increasing z is due to the dipole character of the potential (1) at $z\gg d$. On the contrary, at small z, r, $\ll d$, the main contribution to the correlation function is made by small scale fluctuations $\delta\Sigma$ with $k>d^{-1}$, which have a pure Coulomb potential. At $z, r\ll d$ the correlator (4) tends therefore to

$$K(\mathbf{r},z,z')=2\pi\sigma\frac{e^2}{\varkappa^2}\ln\frac{2d}{z+z'+[r^2+(z+z')^2]^{1/2}}, \tag{6}$$

where $\varkappa=(\varepsilon_i+\varepsilon_s)/2$ is the effective dielectric constant and takes into account the bound charge induced on the interface between media with different ε.[1] The logarithmic divergence of the correlation function at $z, z', r\to 0$ is due to localization of the built-in charge in the plane of the interface. Despite this "lateral" location of the fluctuating charge, the potential relief in the subsurface layer of the semiconductor is essentially three-dimensional, since the character of the change of the correlation function (6) relative to r is approximately the same as relative to z. The fluctuating electric fields in the directions tangential and normal to the interface are likewise approximately equal. Indeed, differentiating (6), we get

$$\overline{\delta E_z^2}=\overline{\delta E_x^2}+\overline{\delta E_y^2}=(\pi/2)\,(e^2/\varkappa^2)\,(\sigma/z^2). \tag{7}$$

We compare now the characteristic amplitude of the fluctuating field $(\overline{\delta E_z^{(2)}})^{1/2}$ with the homogeneous space-charge E_s. We replace, for example, z in (7) with the so-called thermal thickness $l=T/eE_s$ of the inversion layer of an

ideal MIS structure (T is the temperature in energy units). It turns out that the ratio

$$(\overline{\delta E_z^2(l)})^{1/2}/E_s=\Delta/T, \quad \Delta=(e^2/\varkappa)(\pi\sigma)^{1/2}$$

does not depend at all on the homogeneous field E_s. As will be made clear later on, the fluctuating relief influences substantially the motion of the electrons when the quantity Δ, which represents in this expression the characteristic energy scale of the fluctuation, exceeds the thermal energy of the electrons. In the temperature region $T<\Delta$ of interest to us the three-dimensional fluctuating fields at a depth on the order of the thickness of the unperturbed inversion layer are therefore larger (and by many times at $T\ll\Delta$) than the homogeneous field of the depleted layer. It is clear that in this situation the use of the notion of a homogeneous inversion layer as the zeroth approximation in the problem of the influence of fluctuations on the electron states is utterly untenable. It is necessary instead to construct a physical model that would take into account initially the three-dimensional character of the fluctuating potential relief in the subsurface region of the semiconductor of the MIS structure.

Before we proceed to implement this program, we shall describe the qualitative picture of the potential relief, summarizing the relations (1)–(7) above. This three-dimensional relief can be imagined to comprise branched "ravines" separated by "ridges" against the background of a uniform variation of the average potential $\bar{\varphi}(z)=\bar{\varphi}s-E_s z$. The characteristic depths of the ravines and the heights of the ridges increase as the surface is approached [Eq. (6)]. At distances $z>d$ the amplitude of the fluctuations decreases sharply [Eq. (5)] and the potential relief becomes smoothed out. It is important that the maxima and aminima of the potential relief are located only on the semiconductor-insulator interface, while only saddle points can exist in the volume of the semiconductor. This exact statement follows from the fact that the random potential (1) is a solution of the Laplace equation.

This leads to a conclusion of importance in what follows, mainly that the so-called percolation level for the conduction along the semiconductor-insulator interface is equal to the average potential $\bar{\varphi}_s$ on this interface. Indeed, if we consider the class of two-dimensional electron motions in planes parallel to the interface $z=\text{const}$, each of them is characterized by its own percolation level, equal to the average position of the bottom of the conduction band in the given plane. The minimal among them is the level of percolation over the interface $\bar{\varphi}_s$ (see the band scheme in Fig. 1). We consider now an electron located at a certain distance z from the surface and having an energy lower than the level of percolation in the given plane, but higher the level of the percolation over the interface. The specific feature of our problem is that nothing prevents this electron from reaching the surface without change in energy and participate there in the longitudinal conduction. The point is that a level line of the potential relief (1), in accord with the known property of the Laplace equation, can be closed only on the interface. Therefore any spatial equal-energy trajectory with energy higher than the percolation surface level is infinite.

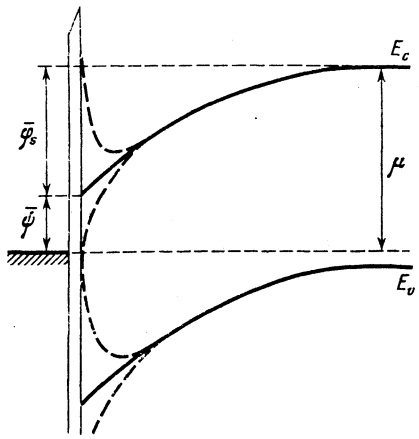

FIG. 1. Band scheme of MIS structure in the inversion regime.

So far we have not taken into account the screening effect of the electrons on the fluctuating relief; this is justified at a low electron density. Just as in Ref. 3, we take this influence into account within the framework of the concept of nonlinear screening.[5] According to Ref. 3, in the presence of electrons the region of integration in (4) must have as the lower limit the wave vector $R_c^{-1}=Q(\sigma/\pi)^{-1/2}$, inasmuch as at a given electron surface density Q their redistribution compensates for the charge of all the fluctuations with scales larger than R_c (in which the mean squared fluctuation of the density of the built-in charge is less than Q). Since, however, in an MIS structure the charge of the large-scale fluctuation is anyway screened by the charges induced on the electrodes, the electron screening becomes significant only at relatively large $Q>(\sigma/\pi d^2)^{1/2}$, when $R_c<d$. From the formal point of view, the imposition of the lower limit in (4) affects the value of the correlator (6) only at $R_c^{-1}>d^{-1}$, inasmuch as the main contribution to this integral is made at $z<d$ only by the region of large wave vectors $k>d^{-1}$. At small $R_c\ll d$, with allowance for electron screening, we have

$$\overline{\delta\varphi^2(z)}=(\pi e^2/\varkappa^2)\sigma\ln(1+R_c^2/z^2). \tag{8}$$

At $Q<(\sigma/\pi d^2)^{1/2}$, when $R_c>d$, we shall use the previous expression (6).

We note that even in the presence of inversion electrons the preceding exact statement concerning the percolation level remains in force, since the sign of the electron charge is such that the electrons can contribute only to the appearance of maxima of the potential in the interior of the subsurface layers, from which the electrons are pushed out. Therefore under conditions of substantial electron screening of the fluctuations the topology of the equal-energy conducting trajectories remains the same as before.

We note that at large Q it is also necessary to take into account the contribution of the electron charge itself to the formation of the uniform field E_s that presses the electrons towards the interface. This supplementary field is due to the purely electrostatic attraction of the charge to the metallic surface. With allowance for this circumstance we have

$$E_s=(2\pi e/\varepsilon_s)(Q+2Nw).$$

We shall need also the distribution function of the random potential in the subsurface region of the MIS structure. By definition[6]

$$P(\varphi, z)=\langle\delta[\varphi-\bar{\varphi}_s+E_s z-\int d^2r(\Sigma(\mathbf{r})-\bar{\Sigma})G(\mathbf{r}, z)]\rangle. \quad (9)$$

Here $G(\mathbf{r},z)$ is the influence function, whose Fourier transform was already used to write down (1), and the δ function is averaged over the ensemble of the realizations of the random distribution $\Sigma(\mathbf{r})$. It is easy to show that at $z\gg\sigma^{-1/2}$ the distribution function (9) tends asymptotically to the Gaussian form

$$P(\varphi, z)=(2\pi\overline{\delta\varphi^2(z)})^{-1/2}\exp\{-(\varphi-\bar{\varphi}_s+E_s z)^2/2\overline{\delta\varphi^2(z)}\}, \quad (10)$$

which we shall use in the subsequent calculations.

2. LOCALIZED CHARGE. DENSITY OF SURFACE STATES

Taking into account the three-dimensionality of the potential relief in the substrate layer, it is necessary to calculate the surface density of the inversion charge by averaging, over the area of the interface, and integrating, with respect to the normal coordinate z, the fluctuating volume concentration of the electrons

$$Q=S^{-1}\int_0^w dz\int d^2r n\{\varphi(r, z), \mu\},$$

where $S>d^2$ and w^2 is the averaging area. Assuming ergodicity, we have

$$Q=\int_0^w dz\int_{-\infty}^{\infty} d\varphi P(\varphi, z)n(\varphi, \mu). \quad (11)$$

Here $n(\varphi,\mu)$ is the known quasiclassical expression for the density

$$n(\varphi, \mu)=\frac{\sqrt{2}}{\pi^2}\left(\frac{m}{\hbar^2}\right)^{3/2}\int_{e\varphi}^{\infty}\frac{d\varepsilon(\varepsilon-e\varphi)^{1/2}}{1+\exp[(\varepsilon-\mu)/T]},$$

where ε and μ are the energy and the chemical potential reckoned from the position of the bottom of the conduction band in the volume of the crystal (μ is negative). In dimensionless energy units we have[2]

$$n(\varphi, \mu)=\frac{1}{\sqrt{2}}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}F_{1/2}\left(\frac{\mu-\varphi}{T}\right), \quad (12)$$

where $(\sigma/\pi a^2)^{3/4}$ is the characteristic density ($a=\hbar^2\varkappa/me^2$ is the Bohr radius), and

$$F_{1/2}\left(\frac{\mu-\varphi}{T}\right)=\frac{2}{\sqrt{\pi}}\int_0^{\infty}\frac{d\varepsilon\,\varepsilon^{1/2}}{1+\exp[(\varepsilon+\varphi-\mu)/T]}$$

$$=\begin{cases}\dfrac{4}{3\sqrt{\pi}}(\mu-\varphi)^{3/2} & \text{at } \mu-\varphi\gg T,\\ T^{3/2}\exp[(\mu-\varphi)/T] & \text{at } \varphi-\mu\gg T\end{cases} \quad (13)$$

is the Fermi integral. Substituting (12) and (10) in (11) we obtain

$$Q=\frac{1}{2}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}$$

$$\times\int_0^w dz\int_{-\infty}^{\infty}\frac{d\varphi}{(2\pi u(z))^{1/2}}\exp\left\{-\frac{(\varphi-\bar{\varphi}_s-\mathscr{E}_s z)^2}{4u(z)}\right\}F_{1/2}\left(\frac{\mu-\varphi}{T}\right), \quad (14)$$

where in accord with (6)

$$u(z)=\ln(d/2z), \quad \mathscr{E}_s=e|E_s|/\Delta>0.$$

To find the dependence of the inversion-layer charge on the average surface potential $\bar{\varphi}_s$, the so called band bending (in our notation the depletion and inversion correspond to negative $\bar{\varphi}_s$, see Fig. 1), we must evaluate the integral (14). At high temperature $T\ll 1$, the Gaussian function in (14) is much narrower for practically all z than the smooth Fermi integral $F^{1/2}$, and consequently plays the role of a δ-function in this expression. The integration with respect to φ yields simply the value of $F^{1/2}$ at the average value of the potential, i.e., $\exp[(\mu-\bar{\varphi}_s-\mathscr{E}_s z)/T]$, integration of which with respect to z leads to the known dependence of the inversion-layer charge on the band bending in an ideal MIS structure.[7] As expected, at $T\gg 1$ the influence of the fluctuations on the inversion layer vanishes. In the opposite $T\ll 1$ of interest to us, the main contribution to the electron density (14) is made by the occupation of the potential fluctuation minima deeper than the Fermi energy, and the Boltzmann tail of the Fermi integral can be neglected. After making the change of variables $z=\frac{1}{2}de^{-u}$ and $\varphi=\mu-\psi$ we then have

$$Q=\frac{\sqrt{2}}{3\pi}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}\frac{d}{2}$$

$$\times\int_0^{\infty}du e^{-u}\int_0^{\infty}\frac{d\psi\psi^{3/2}}{u^{1/2}}\exp\left[-\frac{(\psi+\bar{\psi}+{}^1/_2\mathscr{E}_s de^{-u})^2}{4u}\right], \quad (15)$$

where $\bar{\psi}=\bar{\varphi}_s-\mu$ is the distance from the average position of the bottom of the conduction band on the interface to the Fermi level.

We consider first not too large bends of the bands, $\bar{\psi}\gg 1$, when in the integration of (15) with respect to ψ we can neglect the quadratic term ψ^2 in the argument of the Gaussian exponential. Then

$$Q=\frac{1}{2(2\pi)^{1/2}}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}\frac{d}{2}\int_0\frac{du}{u^{1/2}}\left(\frac{2u}{\bar{\psi}}\right)^{5/2}\exp\left(-u-\frac{\bar{\psi}^2}{4u}\right). \quad (16)$$

Expression (16) was written under the assumption that $\mathscr{E}_s z\ll\bar{\psi}$ at the characteristic distances from the interface. The validity of the inequality will be verified later. The integrand in (16) has a maximum. In fact, with increasing variance of the distribution ($u\to\infty$ as $z\to 0$) the probability of the required deep fluctuations $\psi>\bar{\psi}$ increases. Then, however, the effective thickness of the layer, where the distribution function has this variance, decreases in proportion to e^{-u}. Consequently, the integral (16) can be evaluated by the saddle-point method. Calculating the corresponding derivatives, we obtain the position of the saddle point

$$u_0=1+(1+\bar{\psi}^2/4)^{1/2}\approx\bar{\psi}/2$$

(since $\bar{\psi} \gg 1$) and the characteristic width of the maximum

$$(\bar{\psi}^2/2u_0^3)^{-1/2} = \bar{\psi}^{1/2}/2 \ll \bar{\psi}/2$$

(this justified the use of the saddle-point method). As a result we get

$$Q = \left(\frac{\sigma}{4\pi a^2}\right)^{3/4} \frac{d}{2} e^{-\bar{\psi}}. \tag{17}$$

Obviously, the condition that allows us to neglect in the calculation of (17) the term with the uniform field $\mathscr{E}_s$ in the argument of the exponential is

$$\mathscr{E}_s z_0 = |\bar{\varphi}_s|\,(d/w)\, e^{-\bar{\psi}/2} \ll \bar{\psi}. \tag{18}$$

We have set here the field $\mathscr{E}_s$ equal to the field $2|\bar{\varphi}_s|/w$ of the depleted layer. This inequality is satisfied with a good margin, inasmuch as at $\bar{\psi} \gg 1$ the surface potential $|\varphi_s| \sim \bar{\psi}$. We note that (18) is a mathematical formulation of the statement made in the introduction, that at those distances from the interface which are important for the electron motions the fluctuating electric field is much stronger than the homogeneous field E_s. It is clear that Q in (17) is a localized charge, since the energy of the electrons is lower than the percolation level.

In the derivation of the quasiclassical formula (17) no account was taken of the quantum-mechanical localization energy $\hbar^2/mz_0^2$, equal to the distance between the energy of the electron state and the bottom of the potential well. This is justified when the localization energy is less than the energy scale Δ over which the probability of the potential fluctuations changes:

$$\hbar^2/mz_0^2 = \Delta\,(a/d^2\sqrt{\sigma})\, e^{\bar{\psi}} \ll \Delta. \tag{19}$$

Consequently, expression (17) does not hold and yields much too large values of Q at

$$\bar{\psi} > \ln\,(d^2\sqrt{\sigma}/a) \equiv \psi_{th} \gg 1.$$

The exposition above leads in natural fashion to the concepts of optimal fluctuations,[5,8] within the framework of which we now see clearly the physical meaning of the integrations (15) and (16) performed in the derivation of (17), as well as that of the quantum restriction (19). Taking into account the three-dimensional character of the fluctuations [Eqs. (6), (7)], the potential relief in the subsurface layer of the MIS structure should be regarded as a superposition of three-dimensional potential wells and of maxima of all possible amplitudes $\psi\Delta$ and scales R. The probability of a fluctuation of scale R and depth $\psi\Delta$ is proportional to $\exp[-\psi^2/4\ln(d/R)]$. Given the band bending ($\bar{\psi} = \text{const}$), the electrons fill quasiclassical states in wells of depth larger than $\bar{\psi}\Delta$. In each of them is concentrated a charge $(\sigma/a^2)^{3/4}(\psi - \bar{\psi})^{3/2}R^3$ proportional to the volume of the well. (Here $(\sigma/a^2)^{3/4} \approx (m\Delta/\hbar^2)^{3/2}$ is the characteristic density). Referring this quantity to a unit interface area, i.e., dividing by R^2 and multiplying by the corresponding probability, we obtain the average partial surface charge corresponding to the filling of the fluctuation minima of the given type $\{\psi, R\}$:

$$Q(\psi, R, \bar{\psi}) = (\sigma/a^2)^{3/4}(\psi - \bar{\psi})^{3/2} R \exp\,[-\psi^2/4\ln\,(d/R)]. \tag{20}$$

It is easily seen that the maxima surface density at a given $\bar{\psi}$,

$$Q_m(\bar{\psi}) = (\sigma/a^2)^{3/4}(3/2)^{3/2} d \exp\,(-\bar{\psi} - 3/2),$$

which agrees, apart from a numerical multiplier in the preexponential factor, with the relation (17), is given by the fluctuations

$$\{\psi_m = \bar{\psi} + 3/2, \quad R_m = d \exp\,(-\bar{\psi}/2 - 3/4)\},$$

whose parameters were determined by the equations

$$\frac{\partial Q}{\partial \psi} = \frac{3}{2}\frac{1}{\psi - \bar{\psi}} - \frac{\psi}{2\ln(d/R)} = 0;$$
$$\frac{\partial Q}{\partial R} = 1 - \frac{\psi^2}{4\ln^2(d/R)} = 0.$$

In these optimal quasiclassical fluctuations $\{\psi_m, R_m\}$, the best related (with respect to the average surface density), at a given $\bar{\psi}$, are the probability $\exp[-\bar{\psi}^2/4\ln(d/R)]$, of their existence, which decreases with increasing R, and the number $(\sigma/a^2)^{3/4} R_m^3 \sim e^{-3/2\psi}$, of the electrons localized in such a fluctuation. With increasing $\bar{\psi}$, the dependence of the probability of the optimal fluctuation on the scale R becomes ever stronger therefore its radius decreases. It is clear that these concepts cease to hold as soon as the employed quasiclassical expression for the number of electrons in the optimal fluctuation begins to yield values less than unity, i.e., when $(\sigma d^4/a^2)^{3/4} \exp(-\frac{3}{2}\bar{\psi}) \lesssim 1$. This is in fact the cause of the restriction (19) above.

At $\bar{\psi} > \ln(d^2\sqrt{\sigma}/a)$ the optimal fluctuations for Q are those with relatively small scale, in which the distance between the electron levels is larger than (of the order of) the characteristic energy Δ. In this case the principal role in the organization of the surface charge is played by the filling of the ground state at the bottom of a potential well of suitable depth, and the contribution of the excited states of the deeper wells is exponentially small.[8] Therefore at $\bar{\psi} > \psi$th it is necessary to write for the partial charge $Q(R, \bar{\psi})$ in place of (20)

$$Q(R, \bar{\psi}) = R^{-2} \exp\,\{-[\bar{\psi} + (a/\sqrt{\sigma})R^{-2}]^2/4\ln\,(d/R)\},$$

where the term $(a/\sqrt{\sigma})R^{-2}$ in the exponential is the electron-state localization energy. By varying this expression we obtain

$$\frac{\partial Q}{\partial R} = -2 - \frac{[\bar{\psi} + (a/\sqrt{\sigma})R^{-2}]^2}{4\ln^2(d/R)} + \frac{[\bar{\psi}(a/\sqrt{\sigma})R^{-2}]}{\ln(d/R)} \frac{a}{\sqrt{\sigma}} \frac{1}{R^2} = 0. \tag{21}$$

Hence at $1 < \bar{\psi}/\psi_{\text{th}} < 4$

$$R_m = \sqrt{a}\sigma^{-1/4}, \quad \hbar^2/mR_m^2 \approx 3/2\Delta \ll \bar{\psi}\Delta,$$

and the average surface charge is

$$Q_m(\bar{\psi}) = (\sqrt{\sigma}/a)\,\exp\,(-\bar{\psi}^2/2\psi_{th}). \tag{22}$$

Such a Gaussian expression can be obtained also from the quasiclassical integral (14), by using the lower limit of integration with respect to z the quantum dimension $\sqrt{a}/\sigma^{1/4}$. At $\bar{\psi} > 4\psi_{\text{th}}$ it follows from (21) that

$$R_m = \left(\frac{a}{\sqrt{\sigma}} \frac{2\psi_{th}}{\psi}\right)^{1/2}, \quad \frac{\hbar^2}{mR_m^2} \approx \Delta \frac{\bar{\psi}}{2\psi_{th}} \gg \Delta.$$

Such deep and small-scale fluctuations, however, are optimal only if the potential has a Gaussian distribution, an assumption valid only in the extremal (and unreaslistic) situation $\sigma a^2 \gg 1$.

In fact, let us verify now, knowing the parameters of the optimal fluctuations, whether it was correct to use in our calculation the Gaussian asymptotic form (10) for the distribution function (9). The exact relation obtained by averaging (9) over the Poisson distribution of the built-in charge

$$P(\varphi, z) = \int \frac{d\lambda}{2\pi} \exp\left\{-i\lambda\varphi + \sigma \int d^2r \,(e^{i\lambda G(r,z)} - 1)\right\},$$

expresses the distribution function in terms of the autocorrelators of the potentials of all orders. In our case

$$G \approx \frac{e^2}{\varkappa}\{[r^2+z^2]^{-1/2} - [r^2+(2d+z)^2]^{-1/2}\}$$

and all the autocorrelators are easily calculated. As a result we have

$$P(\varphi, z) = \int \frac{d\lambda}{2\pi} \exp\left\{-i\lambda\varphi - \lambda^2\Delta^2\left[\ln\frac{d}{2z} + 2\sum_{n=3}^{\infty} \frac{(i\lambda e^2/\varkappa z)^{n-2}}{n!\,(n-2)}\right]\right\}$$

The Gaussian asymptotic form (1) corresponds to neglect of the sum of the higher autocorrelators. To check on the Gaussian behavior it is necessary to compare the value of this sum at the value

$$\lambda = \frac{-i\varphi}{2\Delta^2} \ln^{-1}\frac{d}{2z},$$

which corresponds to the Gaussian saddle point, with the second autocorrelator, i.e., to check on the satisfaction of the inequality

$$\ln\frac{d}{2z} > F\left(\frac{\psi}{2\ln(d/2z)} \frac{1}{(\pi\sigma z^2)^{1/2}}\right),$$

$$F(x) = 2x \sum_{n=0}^{\infty} \frac{x^n}{(n+3)!\,(n+1)}. \qquad (23)$$

Naturally, the most dangerous to the Gaussian hypothesis is the quadratic section (22) of the function $\ln Q(\bar{\psi})$, where the dominant role is played by relatively deep, $\bar{\psi} > \psi_{\text{th}}$, and small-scale, $z \approx R_m \approx \sqrt{a}/\sigma^{1/4}$, fluctuations. Substituting these values in (23) we obtain

$$\psi_{th} > 2F[(\bar{\psi}/\psi_{th})(\pi\sigma a^2)^{-1/4}]. \qquad (24)$$

We note first of all that in our problem $\psi_{\text{th}} = \ln(d^2\sqrt{\sigma}/a)$ is a large quantity. At typical parameter values $d = 10^{-5}$, $a = 10^{-7}$, and $\sigma = 10^{12}\,\text{cm}^{-2}$ we have $\psi_{\text{th}} = (7 \text{ to } 8) \gg 1$. It is precisely to the extent of the validity of this condition that the linear section (17) of the function $\ln Q(\bar{\psi})$ is investigated. Substituting $\psi_{\text{th}} = 8$ in (24) and tabulating the function F, we obtain the following condition under which the considered optimal fluctuations are Gaussian

$$\bar{\psi} < (3-4)\,\psi_{th}(\pi\sigma a^2)^{1/4}.$$

At large $\bar{\psi}$ satisfying the opposite inequality, the main contribution to the surface density Q is made by the Poisson optimal fluctuations, which are small-scale clusters of attracting centers that form quasiatomic electron states. In this case[3]

$$Q = \frac{\bar{\Psi}}{2a^2} \exp\left\{-2(2\bar{\Psi})^{1/2} \ln\frac{\bar{\Psi}^{3/2}}{\pi\sigma^+ a^2}\right\} \quad (\sigma^+ \approx \sigma). \qquad (25)$$

We have written out this expression in terms of effective rydbergs $e^2/\kappa a = me^4/\hbar^2\kappa^2$, in which the fluctuation energy is $\Delta = (\pi\sigma a^2)^{1/2}$. In the same units, Fig. 2 shows plots of $\ln[a^2 Q(\bar{\psi})]$ calculated from Eqs. (17), (22), and (25) for three values of the dimensionless parameter $(\pi\sigma a^2)^{1/2} = e^{-1}$, e^{-2}, e^{-3} and for $d/a = 100$.

Even though in all these cases $\sigma a^2 \ll 1$, each of the curves shown has a well pronounced section with a linear $\ln Q(\bar{\varphi})$ dependence; this section corresponds to multiparticle filling of large-scale Gaussian fluctuations. The slope of these sections, which are shown by dashed lines, is proportional to $\Delta^{-1} \sim (\pi\sigma a^2)^{-1/2}$. They are followed by more or less extended sections of the Gaussian relation (22), where relatively small-scale ($R \approx \sqrt{a}/\sigma^{1/4}$), but also Gaussian fluctuations predominate. This section gives way to the Poisson dependence (25). We note that the proposed criterion (24) for a Gaussian relation indicates fairly well the position of the transition region between the relations (22) and (25). In fact, the only restriction on the validity of our theory is the condition $\sigma d^2 \gg 1$, violation of which leads to the loss of the linear section (17). It follows also from (23), if we substitute in it the parameters $z = de^{-\bar{\psi}/2}$ and $\bar{\psi}$, which are responsible for this section.

We note also that whereas the transition from the linear relation (17) to the Gaussian (22) takes place for perfectly reasonable (and observable) values $Q \approx (\pi\sigma a^2)^{1/4} da = 10^{11}$–$10^{12}\,\text{cm}^{-2}$, the transition from (22) to the Poisson relation takes place at unusually small values Q, which are simply not commensurate with the accuracy of the experimental methods of determining the value of the surface charge. A detailed investigation of such deep sections of the function $Q(\bar{\psi})$ can hardly make sense.

Of much greater importance is a generalization of Eq. (17) to the region $\bar{\psi} < \ln(d^2\sigma^{1/4}/a^{3/2})$, where the electron screening of the fluctuations is important ($R_c < d$). It is achieved by replacing in (14) the unscreened variance of the fluctuations (6) by expression (8), which takes the screening into account. If we next make z in (14) nondimensional with the aid of R_c rather than $d/2$, we are left with the same

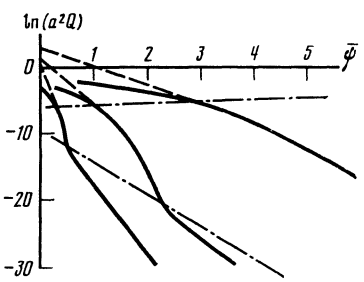

FIG. 2. Resultant theoretical dependences of the surface charge density of the inversion layer on the band bending. The dash-dot lines demarcate the linear, Gaussian, and Poisson sections.

integrals (15) and (16), but multiplied by R_c instead of $d/2$. [We note that at large $\bar{\psi}$ we can neglect the unity in the logarithm of (8).] We therefore obtain in place of (17) taking screening into account,

$$Q=\left(\frac{\sigma}{4\pi a^2}\right)^{3/4}\left(\frac{\sigma}{\pi}\right)^{1/2}\frac{1}{Q}e^{-\sqrt{\bar{\psi}}}.$$

Hence

$$Q=(2a)^{-3/4}(\sigma/\pi)^{5/8}e^{-\bar{\psi}/2}. \tag{26}$$

The self-consistent expression (26), which is valid in a rather large energy interval $2<\bar{\psi}<\ln(d^2\sigma^{1/4}/a^{3/2})\approx\psi_{\text{th}}$, describes the dependence of the density of the localized surface charge on the band bending in the inversion region—in the most interesting operating region of the MIS structure. We note that our "basic" equation (17) shows in fact no applicability region, inasmuch as $(\sigma a^2)^{1/4}\approx 1$ in the situations of practical interest. Therefore Eq. (26) simply replaces (17) over the entire interval $\bar{\psi}<\psi_{\text{th}}$, as shown in Fig. 2.

Finally, we consider the influence of the fluctuations on the majority carriers—the holes. Despite the fact that the average field E_s repels them from the interface (we consider as before the depletion and inversion regimes) the maxima of the fluctuation relief capture and localize the holes on the surface just as the minima do the electrons, inasmuch as at the vital (of the order of a $(\sigma a^2)^{-1/8}$) distances from the interface the fluctuating fields are much stronger than the average field E_s. The surface density Q_p of the captured holes is calculated in the same way as the electron density Q_n, with the obvious difference that the energy must in this case be reckoned down from the top of the valence band. As a result we obtain for the hole density Q_p expressions similar to (22) and (26).

$$Q_p=(2a_p)^{-3/4}(\sigma/\pi)^{5/8}\exp\left[-(E_g-\bar{\psi})/2\right] \tag{27}$$

at $E_g-\bar{\psi}<\ln(d^2\sqrt{\sigma}/a_p)$ and a Gaussian expression of type (22) at $E_g-\bar{\psi}>\psi_{\text{th}}$. Here a_p is the Bohr radius of the hole and $E_g=E_c-E_v$. The total bound surface charge is, of course, equal to the difference Q_p-Q_n.

As a rule experimenters cite the dependences of the derivative of the surface charge with respect to the Fermi level relative to the edges of the allowed bands on the semiconductor-insulator interface. We shall also calculate this derivative, which has the meaning of the density of the surface states. If $E_g<2\ln(d^2\sqrt{\sigma}/a)$, the regions of applicability of expressions (26) and (27) overlap and the Gaussian tails (22) can be disregarded. Then

$$N_{ss}(\bar{\psi})\equiv\frac{\partial}{\partial\bar{\psi}}(Q_p-Q_n)=\frac{1}{\Delta}\frac{(\sigma/\pi)^{5/8}}{(4a_na_p)^{3/8}}\exp\left(-\frac{E_g}{4\Delta}\right)\operatorname{ch}\frac{\bar{\psi}-\psi_0}{2\Delta}. \tag{28}$$

The position of the minimum of $N_{ss}(\bar{\psi})$ is

$$\psi_0=E_g/2+{}^3/_4\Delta\ln(m_n/m_p); \tag{29}$$

it is shifted from the center of the forbidden band towards the band with lower effective mass. The relation (28), which is cited for convenience in dimensional units, is similar to the broad U-shaped spectra N_{ss} of real silicon MIS structures. A detailed comparison with the experimental data, which will be made in the last section, will show that the function (28) is close to them not only qualitatively but also quantitatively.

3. CONDUCTIVITY

Assuming the possible energy dependence of the mobility to be weak, we refer to the conductivity as the density of electrons with energies higher than the percolation level (the average surface potential $\bar{\psi}_s$), multiplied by a certain constant, the effective mobility. In this formulation, the problem of the dependence of the surface conductivity on the temperature and the total density of the electrons reduces again to an investigation of integrals of the type (14).

We consider first the nondegenerate situation corresponding to finite temperatures and not too high a surface density of the electrons, when the Fermi level lies below the percolation level $\bar{\varphi}_s$, i.e., $\bar{\varphi}_s-\mu=\bar{\psi}>0$. The density of the localized electrons is then determined by the Fermi filling of the deep fluctuations [Eq. (26) or (22)], and the density of the delocalized electrons having an energy higher than $\bar{\varphi}_s$ is given by the Boltzmann tail of the distribution function:

$$n_d(\varphi,\mu)=\frac{\sqrt{2}}{\pi^2}\left(\frac{m}{\hbar^2}\right)^{3/2}\begin{cases}\int_{\bar{\varphi}_s}^{\infty}d\varepsilon\,(\varepsilon-\varphi)^{1/2}\exp\left(-\dfrac{\varepsilon-\mu}{T}\right), & \varphi<\bar{\varphi}_s,\\ \int_{\varphi}^{\infty}d\varepsilon\,(\varepsilon-\varphi)^{1/2}\exp\left(-\dfrac{\varepsilon-\mu}{T}\right), & \varphi>\bar{\varphi}_s.\end{cases}$$

Whence

$$n_d(\varphi,\mu)=n_c\exp\left(-\frac{\bar{\psi}}{T}\right)\begin{cases}(2/\sqrt{\pi})\left[(\bar{\varphi}_s-\varphi)/T\right]^{1/2}, & \varphi<\bar{\varphi}_s,\\ \exp\left[-(\varphi-\bar{\varphi}_s)/T\right], & \varphi>\bar{\varphi}_s,\end{cases} \tag{30}$$

where $n_c=2(mT/2\pi\hbar^2)^{3/2}$ is the effective state density in the conduction band. Averaging (30) over the distribution function (10) and integrating with respect to z, we obtain

$$Q_d=n_c\exp\left(-\frac{\bar{\psi}}{T}\right)\left\{\frac{T^{-1/2}}{\pi}\int_0^{\infty}dz\int_{-\infty}^{0}d\varphi\left[\frac{-\varphi}{u(z)}\right]^{1/2}\left[-\frac{(\varphi-\mathscr{E}_sz)^2}{4u(z)}\right]\right.$$
$$\left.+\int_0^{\infty}dz\int_0^{\infty}\frac{d\varphi}{[4\pi u(z)]^{1/2}}\exp\left[-\frac{\varphi}{T}-\frac{(\varphi-\mathscr{E}_sz)^2}{4u(z)}\right]\right\}. \tag{31}$$

The temperature dependence of the conductivity is set by $\exp(-\bar{\psi}/T)$, and the curly brackets with the integral determine the effective localization depth, at the interface, of the electrons delocalized in directions parallel to this interface. This length, multiplied by n_c, gives the effective surface density of the delocalized electron states. Just as before, φ in (32) is dimensionless, and the average normal field

$$\mathscr{E}_s=(1+\varepsilon_i/\varepsilon_s)(Q+2Nw)(\sigma/\pi)^{-1/2}$$

has the dimension of reciprocal length, while $\frac{1}{2}\ln(1+R_c^2/z^2)=u(z)$ is the dimensionless variance of the fluctuations with allowance for the screening (8). In the considered nondegenerate situation the density of the mobile carriers is much less than that of the localized ones. In the calculation of (31) there

is therefore no need for self-consistency, and it is necessary to substitute in $\mathscr{E}_s$ and R_c the density (26) of the delocalized electrons.[4] In contrast to (14)–(16) the integrals (31) depend significantly on the average field $\mathscr{E}_s$ and diverge as $\mathscr{E}_s \to 0$. At $T \ll 1$ the major role in (31) is assumed by the first integral, which takes into account electrons with energies higher than $\bar{\varphi}_s$ and located in those subsurface regions where $\varphi < \bar{\varphi}_s$, i.e., just where the localized electrons are concentrated. Making the change of variables $\varphi \to -\varphi$ and $z\mathscr{E}_s = \zeta$, this integral takes the form

$$\mathscr{E}_s^{-1} I_1 = \mathscr{E}_s^{-1} \int_0 d\zeta \int_0 \frac{d\varphi (2\varphi)^{1/2}}{\ln^{1/2}(1+R_c^2\mathscr{E}_s^2/\zeta^2)} \times \exp\left\{-\frac{(\varphi+\zeta)^2}{2\ln(1+R_c^2\mathscr{E}_s^2/\zeta^2)}\right\}. \quad (32)$$

The integral I_1 depends weakly (logarithmically) on the product

$$R_c\mathscr{E}_s = (1+\varepsilon_i/\varepsilon_s)(1+2Nw/Q_l),$$

which tends to $(1+\varepsilon_i/\varepsilon_s) \sim 1$ at $Q_l > 2Nw$ (when a noticeable conductivity can be expected). Here $I_1 = (2^{5/2}/3)\Gamma(\frac{5}{4}) \approx 2$, where $\Gamma(x)$ is a gamma function. The second integral in (31), which sums the total density of the electrons in the sublayer regions with $\varphi > \bar{\varphi}_s$, becomes the principal one at $T \gg 1$. It tends then to $T/\mathscr{E}_s$—the effective thermal thickness of the inversion layer of an ideal MIS structure, while Q_d becomes much larger than the bound charge Q_l (at $T \gg 1$ the electrons can hardly feel the fluctuations). Thus, in the low-temperature region $T \ll 1$ of interest to us we have

$$Q_d = \left(\frac{\sigma}{\pi a^2}\right)^{3/4} \frac{T}{eE_s} \exp\left(-\frac{\bar{\Psi}}{T}\right). \quad (33)$$

The character of the spatial distribution of the delocalized electrons in the subsurface region of the MIS structure is shown in Fig. 3, where the "cross section" of the mean statistical potential ravine is shown arbitrarily. The conducting electrons are concentrated in a layer having a thickness of the order of the temperature of the average surface potential—of the percolation level. Their volume density is proportional to the product of the temperature by the square root of the character depth Δ of the ravine. Since the uniform field E_s raises the average level of the potential relief with increasing distance from the interface, the effective length Δ/eE_s of such an electron "puddle" is finite. The effective surface density of the delocalized electron states is therefore proportional to $T\Delta^{3/2}/eE_s$. Their filling is determined by the Boltzmann factor $\exp(-\bar{\psi}/T)$.

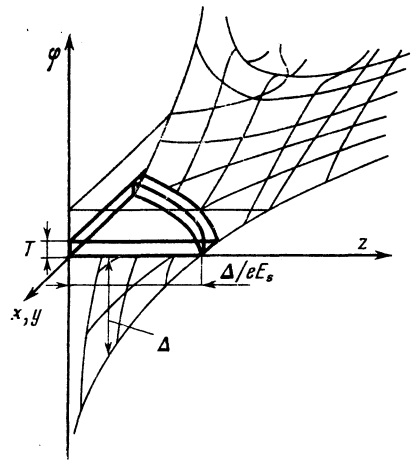

FIG. 3. Schematic picture of the potential relief in the subsurface layer of the semiconductor of an MIS structure.

Corresponding to the second integral in (31) are electron in a narrow (once $T \ll \Delta$) belt near the edge of this puddle. Its relative narrowness T/Δ is offset by its length Δ/eE_s, so that the characteristics Δ and R_c of the fluctuation relief are no longer present in the resultant surface density. This charge component (31) prevails at $T \gg \Delta$, when the thermal length T/eE_s is much larger than the fluctuation length Δ/eE_s. In this case the individual electron puddles merge also in the interior of the semiconductor, so that only crests of individual ridges project from the common electron sea near the surface. To prevent misunderstanding, we note that at low temperatures the puddles in question are interconnected by saddle regions at the interface—after all, it is precisely from this condition that we have determined the percolation level.

We consider now the degenerate situation that arises at $T \ll 1$, when the electron density Q exceeds the threshold value

$$Q_l^m = \frac{1}{a^{3/4}}\left(\frac{\sigma}{\pi}\right)^{5/8} \times \left[\frac{2}{3\pi}\int_0^\infty d\zeta \int_0^\infty \frac{d\varphi\,\varphi^{1/2}}{\ln^{1/2}(1+\zeta^{-2})} \exp\left\{-\frac{(\varphi+\zeta)^2}{2\ln(1+\zeta^{-2})}\right\}\right]^{1/2} \approx \frac{1}{(2\pi)^{1/2}}\frac{1}{a^{3/4}}\left(\frac{\sigma}{\pi}\right)^{5/8}. \quad (34)$$

This is the so-called maximum localized charge, defined by the following condition: the Fermi level is equal to the percolation level. Its value turned out to be somewhat smaller than the pre-exponential factor of (26), and the minimum screening radius corresponding to Q_l^m is

$$R_c^m = (2\pi a/\sigma^{1/2})^{1/2}(\pi\sigma a^2)^{1/8}.$$

It is clear that (34) should be approached with some skepticism in the case $\pi\sigma(R_c^m)^2 = 2(\pi^3\sigma a^2)^{3/4} \lesssim 1$, when the employed concept of nonlinear screening of the fluctuation is not valid.

At a total electron density $Q > Q_l^m$ the density of the conducting delocalized electrons is $Q_d = Q - Q_l^m$. Generally speaking, it would be possible here to confine oneself to this statement, i.e., dispense with the calculation of the dependence of Q_d on the $\bar{\varphi}_s$ band bending in the region where $\bar{\varphi}_s < \mu$ (the Fermi level is higher than the percolation level). The point is that the average surface potential $\bar{\varphi}_s$ and accordingly $\bar{\varphi}_s - \mu = \bar{\psi}$ are internal parameters of the problem, which determine the relation between the localized and delocalized charges whose total value is set by the voltage on the structure electrode. The measured quantity is the mobile

charge Q_d, and the task of the theory is to establish the connection between this mobile charge and the total charge Q. At $Q<Q_l^m$, when the Fermi level is below the percolation level, we establish this connection from the manner in which the two components of the surface charge (mobile and immobile) depend on $\bar{\psi}$ [Eqs. (26) and (33)]. In the opposite the connection between them is trivial: $Q_d = Q - Q_l^m$, and at first glance the $Q_d(\bar{\varphi}_s)$ dependence at $\bar{\varphi}_s < \mu$ is of no particular interest. We shall nevertheless find it, primarily to establish the character of the spatial distribution of the mobile electrons in the subsurface layer of the semiconductor, i.e., the structure of the inversion layer.

To this end, just as in the derivation of (31), we sum the electrons with energies higher than $\bar{\varphi}_s$, using now a Fermi distribution function with $T = 0$. As a result we obtain again the density of the delocalized electrons in the form of a sum of two integrals:

$$Q_d = \frac{2}{3\pi}\left(\frac{\sigma}{\pi a^2}\right)^{3/4} \mathscr{E}_s^{-1} \int_0^\infty \frac{d\zeta}{\ln^{1/2}(1+\zeta^{-2})} \times \left\{ \int_0^{\bar{\mu}} d\varphi [(\bar{\mu}+\varphi)^{3/2} - \varphi^{3/2}] \exp\left[-\frac{(\varphi+\zeta)^2}{2\ln(1+\zeta^{-2})} \right] + \int_0^{\bar{\mu}} d\varphi (\bar{\mu}-\varphi)^{3/2} \exp\left[-\frac{(\varphi-\zeta)^2}{2\ln(1+\zeta^{-2})} \right] \right\}, \quad (35)$$

the first of which takes into account the delocalized electrons located in those regions of the subsurface layer where $\varphi < \bar{\varphi}_s$, and the second where $\varphi > \bar{\varphi}_s$. In (35), just as in (32), $R_c \mathscr{E}_s = 1$, and the transverse coordinate ζ is made nondimensional relative to the field length $\mathscr{E}_s^{-1}$. Here $\bar{\mu} = \mu - \bar{\varphi}_s$.

Just as in (32), when calculating (35) we can neglect the change of the variance of the distribution function with changing ζ. At $\bar{\mu} \ll 1$ the main contribution to (35) is made by the first integral. Approximating in it the expression

$$[(\bar{\mu}+\varphi)^{3/2} - \varphi^{3/2}] = {}^3/_2 \bar{\mu}\varphi^{1/2} + \left[(\bar{\mu}+\varphi)^{3/2} - \varphi^{3/2} - \frac{3}{2}\bar{\mu}\varphi^{1/2} \right] \approx {}^3/_2 \bar{\mu}\varphi^{1/2} + \bar{\mu}^{3/2}$$

(in this case we overestimate it somewhat at large φ, but this is of little importance since large φ are suppressed by the Gaussian function), we obtain

$$Q_d^{\,1} = \frac{3}{2\pi}\left(\frac{\sigma}{\pi a^2}\right)^{3/4} \frac{\bar{\mu}}{\mathscr{E}_s}\left(1 + \frac{4}{9}\bar{\mu}^{1/2}\right). \quad (36)$$

The characteristic thickness of the layer occupied by this component of the moving charge is $\mathscr{E}_s^{-1}$. At $\bar{\mu} \gg 1$ the second integral in (35) is much larger than the first. Its value can be easily estimated by noting that at $\bar{\mu} \gg 1$, at the vital distances $\bar{\mu}/\mathscr{E}_s$, the variance $\ln(1+\zeta^{-2})$ is much less than the square of the characteristic potential $\varphi \sim \bar{\mu}$. This enables us to replace in it the distribution function by a δ function against the background of a smooth density of states $(\bar{\mu} - \varphi)^{3/2}$. Then

$$Q_d^{\,2} = \frac{1}{3\sqrt{\pi}}\left(\frac{4\sigma}{\pi a^2}\right)^{3/4} \mathscr{E}_s^{-1} \int_0^{\bar{\mu}} d\zeta (\bar{\mu}-\zeta)^{3/2} = \frac{4\sqrt{2}}{15\sqrt{\pi}}\left(\frac{\sigma}{\pi a^2}\right)^{3/4} \frac{\bar{\mu}^{5/2}}{\mathscr{E}_s}, \quad (37)$$

which coincides exactly with the corresponding expression for an ideal MIS structure in the case of degeneracy. Adding (36) and (37) we get

$$Q_d = Q_l^m \frac{\bar{\mu}}{\Delta} + \frac{4\sqrt{2}}{15\pi^2}\left(\frac{m\bar{\mu}}{\hbar^2}\right)^{3/2} \frac{\bar{\mu}}{eE_s}. \quad (38)$$

We have left out here the second term in the parentheses of (36), since at $\bar{\mu} \ll 1$, it is smaller than the first, and at $\bar{\mu} \gg 1$ the charge (37) is larger than both. Just as all the final expression, (38) is written in dimensional units. The spatial distribution of the delocalized electrons under degeneracy conditions is of the same character as in the nondegenerate situation considered above. The reader can visualize it qualitiatively by using Fig. 3 in which the energy thickness T is replaced by $\bar{\mu}$.

4. DISCUSSION OF RESULTS. COMPARISON WITH EXPERIMENT

Surface states. By now a tremendous number of measurements of the density of surface states in silicon, mainly MIS structures, have been reported. Despite the great variety of the experimental methods employed, the observed N_{ss} spectra are quite alike. Typical plots of $N_{ss}(\bar{\psi})$ [in $cm^{-3}\cdot eV^{-1}$] taken from Sze's monograph,[9] are shown in Fig. 4. These and other empirical laws can be naturally explained within the framework of the theory expounded here. We note that Eq. (28), which expresses the entire spectrum of the surface states in terms of a single parameter σ, the summary density of the charged centers on the semiconductor-insulator interface, is close also quantitatively to the experimental results. Indeed, if Δ is determined from the slopes of the exponential sections of the spectra of Fig. 4 at the edges of the allowed bands ($\Delta_1 = 0.05$, $\Delta_2 = 0.035$ eV, which corresponds to $\sigma_1 = 2\times 10^{12}$ and $\sigma_2 = 10^{12}$ cm^{-2}), and (28) is then used to calculate the density of states at the minimum of

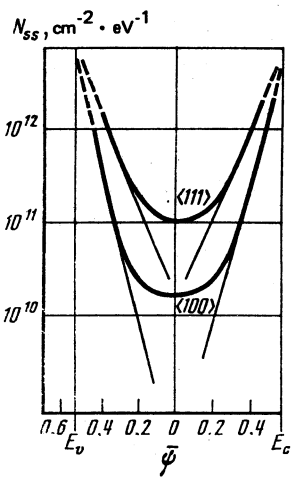

FIG. 4. Typical spectra of the surface states of silicon MIS structures.[9]

the spectrum, we obtain $N_{ss}^{\min}(1) = 3 \cdot 10^{11}$ and $N_{ss}^{\min}(2) = 2 \cdot 10^{10}\,\text{cm}^{-3}\,\text{eV}^{-1}$. In this calculation we used the values $\varkappa = (\varepsilon_i + \varepsilon_s)/2 = 8$ and $m_n = 2m_p = m_0$ inasmuch as the effective masses of the density of states are used in (28). Taking into account the exponential character of (28), such an agreement between theory and experiment should be regarded as good. We note that at these values of Δ at the center of the forbidden band of silicon Eq. (38) is at the borderline of its applicability region, since the required inequality $\frac{1}{2}E_g < \Delta \ln(d^2\sqrt{\sigma}/a)$ is not satisfied ($Eg/2 = 0.55$ and $\Delta \ln(d^2\sqrt{\sigma}/a) = 0.05 \cdot 8 = 0.4$), consequently (28) overestimates somewhat N_{ss} at the center of the band. At still lower Δ one should expect at the center of the forbidden band of silicon an abrupt dip of the density of states, corresponding to the segment (22) of the function $Q(\bar{\psi})$, as is indeed observed in high-grade MIS structures with low surface-state density. It is simply impossible to measure N_{ss} in the center of the band in such structures,[10] and usually one draws there a dashed line corresponding to the experimental accuracy (at the level of 10^{9}–$10^{10}\,\text{cm}^{-2}\cdot\text{eV}^{-1}$).

Thus, the surface states in MIS structures are mainly three-dimensional potential wells at the interface with the insulator, and are due to spatial fluctuations of the density of the built-in charged centers in the insulator near the interface with the semiconductor. The characteristic dimension of the optimal fluctuation potential wells with allowance for dispersion is

$$(1-\exp(\bar{\psi}^m/2))^{1/4} 2a(\sigma a^2)^{-1/4}.$$

It can "house" from one to $(\sigma a^2)^{3/8}\exp(3\bar{\psi}^{1/2}/2)$ electrons.

Conductivity and mobility. Having expression (26) for the localized charge, we can analyze the main features of the surface electric conductivity of MIS structures in the so-called subthreshold region, where the voltage on the transistor junction is $V < V_{\text{th}} = eQ_l^m/C_0$ ($C_0 = \varepsilon_i/4\pi d$ is the specific capacitance of the insulator). For this purpose it is necessary to solve the nonlinear algebraic equation

$$C_0 V = Q = Q_l(\bar{\psi}) + Q_d(\bar{\psi})$$

and express the density of the mobile chart Q_d in terms of the total charge Q. At $T \ll \Delta$ the mobile charge $Q_d \ll Q_l^0 e^{-\psi/2\Delta} \approx Q$, whence follows a logarithmic dependence of the surface-conduction activation energy on the total charge (voltage)

$$\bar{\psi} = 2\Delta(\ln Q_d^0 - \ln Q), \tag{39}$$

which is also in quantitative agreement with the experimental results cited, for example, in Adkin's review.[11] Substituting (39) in (33) we obtain the of the conductivity of the inversion layer on the voltage on the transistor junction;

$$j \infty Q_d = Q_d^0 (CV/Q_l^0)^{2\Delta/T},$$

which is valid at $V < V_{\text{th}}$ (the voltage V_{th} is sometimes called the threshold of the metallic conduction). This rather wide segment (up to several volts according to estimates) can be easily observed by replotting the corresponding graphs in log-log form.

In conclusion, a few qualitative ideas concerning the surface mobility. At small charges $Q < Q_l^m$ (when $\bar{\psi} > 0$) the conductivity of the inversion layer has a clearly pronounced percolation character. In this case electrons that are mobile in principle and have energies of the order of T above the percolation level $\bar{\varphi}_s$ move mainly within the confines of three-dimensional potential "bags" of depth Δ and characterize dimension R_c. Since the area of the saddle regions at the interface, that connect them with one another, are $(T/\Delta)^2$ times smaller than the total surface of the bag. Consequently, the effective mobility determined by the transitions of the electron from one bag to another will be suppressed in proportion to the ratio $(T/\Delta)^2$. This situation will remain in force until the total charge Q becomes comparable with the quantity $Q_l^m(1 + T/\Delta)$, when the Fermi level exceeds the percolation level by an amount of the order of the thermal energy. With further increase of the charge, $\bar{\mu}$ becomes larger than T and the effective area of the "necks" joining the bags increases. Therefore the effective mobility can increase until $\bar{\mu}$ reaches a value of the order of Δ ($Q \simeq 2Q_l^m$) and such topological restrictions on the conducting trajectories become insignificant. These qualitative arguments explain, in our opinion, the decrease of mobility as the threshold is approached from the high energy side, observed in MIS structures with the aid of Hall-effect[12] and other[13] methods, as well as the noticeable decrease of the mobility on this descending section when the temperature is lowered.

[1]Strictly speaking, the quantity under the logarithm sign in (6) should be some effective thickness $\tilde{d}$ that depends on the ratio $\varepsilon_i/\varepsilon_s$, but does not differ too much from d. We shall not pay attention to this difference, since even if d does enter in the final expression, it will do so only under the logarithm sign.

[2]Except for specially stipulated cases, we use below dimensionless energy parameters (normalized to Δ), and retain for them the same notation as for the dimensional quantities.

[3]This expression is derived in the same manner as the Poisson section of the density of states in ordinary statistically uniform three-dimensional systems.[5,8] We therefore present it here without a detailed discussion. Our situation differs only in that the ground state must be taken here to be the first excited state of the three-dimensional quasi-atom, since it is the latter which satisfies the boundary condition—zero of the wave function on the interface with the insulator. This circumstance accounts also for the corresponding factors 2 in the argument of the exponential in (25).

[4]The density of the localized electrons will be designated hereafter by Q_l.

[1]M. V. Sadovskii, Usp. Fiz. Nauk **133**, (1981) [Sov. Phys. Usp. **24**, (1981).
[2]V. L. Berezinskiĭ, Zh. Eksp. Teor. Fiz. **65**, 1251 (1973) [Sov. Phys. JETP **38**, 620 (1974)]. N. F. Mott and W. D. Twose, Adv. Phys. **10**, 107 (1960).
[3]V. A. Gergel' and R. A. Suris, Zh. Eksp. Teor. Fiz. **75**, 191 (1978) [Sov. Phys. JETP **48**, 95 (1978)].
[4]V. A. Gergel', Fiz. Tekh. Poluprov. **13**, 654 (1979) [Sov. Phys. Semicond. **13**, 385 (1979)].
[5]B. I. Shklovskiĭ and A. L. Éfros, Elektronnye svoĭstva legirovannyk poluprovodnikov (Electronic Properties of Doped Semiconductors), Nauka, 1979.
[6]S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[7]C. G. Garrett and W. H. Brattein, Phys. Rev. **99**, 376 (1955).
[8]I. M. Lifshitz, S. A. Gredeskul, and L. A. Pastur, Vvedenie v teoriyu neuporyadochennykh sistem (Introduction to the Theory of Disordered Systems), Nauka, 1982.
[9]C. M. Sze, Physics of Semiconductor Devices, Wiley, 1981.
[10]T. P. Ma and R. C. Barker, Solid St. Electron. **17**, 913 (1974).

[11]C. J. Adkins, J. Phys. Ser C, **11**, 952 (1978).
[12]A. V. Rzhanov, ed., Svoĭstva struktur metall-dielektrik-poluprovodnik (Properties of Metal-Insulator-Semiconductor) Structures, Nauka, 1976.
[13]E. V. Vlasenko, R. A. Suris, A. M. Filachev, and B. I. Fuks, Pis'ma Zh. Eksp. Teor. Fiz. **31**, 49 (1980) [JETP Lett. **31**, 44 (1980)].

Translated by J. G. Adashko