Fluctuation surface states and conductivity of inversion layers in MIS structures

V. A. Gergel' and G. V. Shpatakovskaya

F. V. Lukin Scientific-Research Institute of Physical Problems
(Submitted 20 February 1992)
Zh. Eksp. Teor. Fiz. **102**, 640–647 (August 1992)

The effect of statistical fluctuations of the built-in charge density on the surface conductivity of inversion layers in MIS structures is investigated. The charge fluctuations considered induce in the boundary layer of the semiconductor a random potential relief, near whose minima bound (localized) electronic states form. Some electrons in the inversion layer are trapped in the states and form localized charge. The mobile, delocalized charge consists of electrons whose energies lie above the percolation level—the average surface potential. The dependences of the bound and free charges on the temperature and bending of the bands were calculated in the quasiclassical approximation by averaging over fluctuations of the potential. It is shown that, for temperatures $T \ll \Delta$ (the characteristic energy of the fluctuations), as the total charge in the inversion layer increases, first the bound states and then the free states are filled. This gives rise to a corresponding shift in the threshold of the dependence of the surface conductivity on the controlling voltage on the electrode of the MIS structure. Conversely, at high temperatures almost all of the inversion charge is mobile. It was determined that at intermediate temperatures $2\Delta > T > \Delta/2$ (this range can constitute several hundreds of degrees) the mobile and bound components of the total charge of the inversion layer change in proportion to one another. This is manifested experimentally as an appreciable decrease in the effective surface mobility of electrons as compared with the values in the bulk.

1. It is well known that a semiconductor–insulator interface always contains a certain number of electrically charged points defects—so-called built-in (or fixed) charge.[1,2] It was previously believed that the effect of the built-in charge on the basic electrical characteristics of MIS structures [voltage dependence of the differential capacitance (C-V characteristic) and the dependence of the conductivity of the inversion layer on the same voltage on the controlling electrode (gate of the MIS structure)] reduces to an additive shift in these characteristics along the voltage axis by an amount proportional to the average surface density of this charge.[3] It turns out, however, that the effective influence of the built-in charge on the mobile carriers (electrons or holes) appearing in the part of the semiconductor adjoining the insulator with voltages on the electrode which correspond to the inversion regime is associated not so much with the average surface density of the built-in charge as with fluctuations of this density in the plane of the charged centers, i.e., at the semiconductor–insulator interface. The point is that the statistical fluctuations of the built-in charge density induce in the surface part of the semiconductor of the MIS-structure a random potential relief, near whose minima localized (bound) states with energy less than the so-called percolation level are formed. In accordance with the theory of electronic states in disordered systems,[4,5] the electrons filling these localized states can undergo only finite motion and therefore they do not contribute to the surface electric conductivity of the inversion layer. The latter quantity is determined by delocalized electrons in states whose energies lie above the percolation level. Thus the total electronic surface charge Q_t at the interface with the corresponding voltage on the electrodes of the MIS-structure [$Q_t = C_0(V - V_t)$, where $C_0 = \varepsilon_i/4\pi d$ is the capacitance of an insulator layer of thickness d and ε_i is the permittivity of the insulator] consists of two parts: the mobile part Q_d and the immobile (bound) part Q_l.

In order to find the relation between the delocalized Q_d and localized Q_l charge (or the total charge $Q_t = Q_d + Q_l$) each component considered must be expressed in terms of the quantity $\psi = \varphi_s - \mu$, which determines how closely the percolation level—the average position of the edge of the conduction band of the semiconductor at the boundary with the insulator—approaches the Fermi level μ [φ_s is the average surface potential or bending of the bands (see Fig. 1)]. This problem has already been studied in Ref. 6, where the following simple expression was obtained for the charge as a function of the average surface potential in the limit of low temperatures, taking into account the nonlinear electronic screening of the fluctuation potential:

$$Q_l = Q_l^{max} \exp\{-\psi/2\Delta\}. \tag{1}$$

Here $Q_l^{\max} = (2a)^{-3/4}(\sigma/\pi)^{5/8}$, $a = \varkappa\hbar^2/me^2$ is the Bohr radius, $\sigma^+ + \sigma^-$ is the sum of the average surface densities of the positively and negatively charged components of the built-in charge, $\Delta = e^2(\pi\sigma)^{1/2}/\varkappa$ is the characteristic energy scale of the fluctuations of the potential, $\varkappa = (\varepsilon_i + \varepsilon_s)/2$ is the effective permittivity, and ε_s is the permittivity of the semiconductor. This expression is valid when the bending of the bands $|\varphi_s|$ is not too large, so long as $\psi = \varphi_s - \mu > 0$ holds, i.e., so long as the average position of the conduction-band edge at the boundary with the insulator has still not crossed the Fermi level.[1)] Then, correspondingly, $Q_l = Q_t < Q_l^{\max}$, where $Q_l^{\max}$ is the so-called maximum localized charge. For $Q_t > Q_l^{\max}$, we have $Q_l = Q_l^{\max}$ and $Q_d = Q_t - Q_l^{\max}$, and therefore the desired dependence has a pronounced threshold character.

By analogy to the corresponding concept in semicon-

0038-5646/92/080342-05$05.00

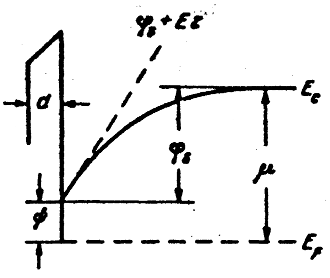

FIG. 1. Portion of the band scheme of an MOS structure in the inversion regime.

ductor physics, the situation described by the formula (1) can be called freezing-out of the electrons of the inversion layer onto the fluctuation surface states. It is obvious that as the temperature increases an increasingly large fraction of the electrons of the inversion layer will leave the bound, localized states and occupy delocalized, mobile states, thereby giving rise to an increase in the surface conductivity. It is of great interest and important to investigate in detail the corresponding temperature dependence in order to determine the following: a) the temperature limit T_1 of the freezing-out region (where the formula (1) holds); b) the characteristic temperature T_2 at which virtually all inversion electrons become delocalized; and, most importantly, c) the change in the ratio of the localized and delocalized charge components of the inversion layer at intermediate temperatures $T_1 < T < T_2$, which, as will be shown below, is found to be quite extended and corresponds to many experimental results and working characteristics of real MOS transistors.

2. In order to evaluate the total charge Q_t and its delocalized part Q_d we start from the general considerations presented in Ref. 6:

$$Q_{t,d}=\frac{2^{1/2}}{\pi^2}\left(\frac{m}{\hbar^2}\right)^{3/2}\int_{-\infty}^{\infty}dz\int_0^{\infty}d\varphi\, P(\varphi,z)\int_{\varphi_{t,d}}^{\infty}d\varepsilon\,\frac{(\varepsilon-\varphi)^{1/2}}{1+e^{(\varepsilon-\mu)/T}}. \tag{2}$$

Here the total charge satisfies $Q_t - \varphi_t = \varphi$, the delocalized charge satisfies $Q_d - \varphi_d = \max(\varphi,\varphi_s)$; and

$$P(\varphi,z)=\frac{1}{(2\pi\langle\delta\varphi^2\rangle)^{1/2}}\exp\left\{-\frac{(\varphi-\varphi_s-Ez)^2}{2\langle\delta\varphi^2\rangle}\right\}$$

is the distribution function of the random potential. The formula (2) corresponds to averaging of the quasiclassical volume density of the electrons over fluctuations of the potential followed by integration of this density over the depth of the inversion layer z ($z>0$ is the semiconductor region of the MIS structure). The order of integration employed in Eq. 2 is necessary, since both the average value of the potential $\varphi=\varphi_s+Ez$ (E is the average electric field pressing electrons to the interface) and the variance of its distribution

$$\langle\delta\varphi^2\rangle=\Delta^2\ln\left(1+\frac{R^2}{z^2}\right)\equiv 2\Delta^2 u(z), \tag{3}$$

where $R=(\sigma/\pi)^{1/2}/Q_t$ is the nonlinear electron screening radius,[5,6] are functions of z.

We normalize all energy characteristics of the problem to the characteristic fluctuation energy Δ. Then Eq. (2) can be rewritten as follows:

$$Q_t=\frac{2}{3\pi^{1/2}T}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}\int_0^{\infty}\frac{dz}{[2\pi u(z)]^{1/2}}\int_{-\infty}^{\infty}d\varphi\exp\left[-\frac{(\varphi-\varphi_s-Ez)^2}{4u(z)}\right]\times\int_0^{\infty}d\varepsilon\,\frac{\varepsilon^{3/2}\exp\{(\varepsilon+\varphi-\mu)/T\}}{(1+\exp\{(\varepsilon+\varphi-\mu)/T\})^2}, \tag{4}$$

$$Q_d=\frac{1}{\pi^{1/2}}\left(\frac{\sigma}{\pi a^2}\right)^{3/4}\int_0^{\infty}\frac{dz}{[2\pi u(z)]^{1/2}}\int_0^{\infty}d\varepsilon'(\varepsilon')^{1/2}\int_0^{\infty}d\varphi'\,\frac{\exp\{-(\varphi'-\varepsilon'-Ez)^2/4u(z)\}}{1+\exp\{(\psi+\varphi')/T\}}. \tag{5}$$

In deriving the formula (4) for the total charge the integral over the energy ε was rewritten in a more convenient form with the help of integration by parts. As a result, the integrand is represented in the (φ,ε)-plane as a product of two functions with delta-function-like behavior: the function $P(\varphi,z)$ with an extremum at the point $\varphi=\varphi_s+Ez$ and half-width $2u^{1/2}$ and the function

$$\exp\left\{\frac{\varepsilon+\varphi-\mu}{T}\right\}\Big/\left(1+\exp\left\{\frac{\varepsilon+\varphi-\mu}{T}\right\}\right)^2$$

with extremum at the point $\varphi=\mu-\varepsilon$ and half-width T. The value of the integral over φ depends on the ratio of these widths (as will be shown below, it depends on the parameter $\beta=2u^{1/2}/T$).

The expression (5) was obtained from Eq. (2) by the substitution of variables $\varepsilon'=\varepsilon-\varphi$ and $\varphi'=\varphi-\varphi_s$ and by changing the order of integration over φ' and ε'.

3. In order to calculate the total charge Q_t as a function of the temperature T and other parameters (ψ,E,R) it is convenient to rewrite (4), by changing the order of integration, as follows:

$$Q_t=C\int_0^{\infty}d\varepsilon\,\varepsilon^{3/2}\int_0^{\infty}dz\,J(\varepsilon,z), \tag{6}$$

where

$$J(\varepsilon,z)=\int_{-\infty}^{\infty}dt\exp(-f(t)),$$

$$f(t)=t^2-\alpha-\beta t+2\ln(1+\exp(\alpha+\beta t)),$$

$$C=\frac{4}{3\pi}\left(\frac{\sigma}{4\pi a^2}\right)^{3/4}\frac{1}{T},\quad t=\frac{\varphi-\varphi_s-Ez}{2u^{1/2}},$$

$$\alpha=\frac{\varepsilon+\psi+Ez}{T},\quad \beta=\frac{2u^{1/2}}{T}.$$

We evaluate the integral over t by the saddle-point method. The result is

$$J(\varepsilon,z)\approx[2\pi/f_{tt}''(t_0)]^{1/2}\exp(-f(t_0)). \tag{7}$$

Here the saddle point t_0 is found from the equation

$$\tanh\left(\frac{\alpha+\beta t_0}{2}\right)=-\frac{2t_0}{\beta}, \tag{8}$$

and the second derivative $f''_{tt}(t_0)$ has the form

$$f_{tt}''(t_0)=2+\beta^2/2\cosh^2\left(\frac{\alpha+\beta t_0}{2}\right)\geqslant 2.$$

At sufficiently low temperatures ($T\ll 1$, $\beta\gg 1$) the hyperbolic tangent in Eq. (8) can be approximated by the first term of its series expansion. Here $t_0 = (-\alpha/\beta)(1-T^2/u+T^4/u^2)$, and the subsequent integration can be performed analytically, which gives the expression

$$Q_t\approx Q_t=Q_t^{max}\exp\left(-\frac{\psi}{2}+\frac{T^2}{2}\right), \tag{9}$$

which differs by the factor $\exp(T^2/2)$, demonstrating the tendency for the charge Q_t to increase with temperature for fixed ψ, from Eq. (1) obtained in Ref. 6.

In the general case of arbitrary temperatures the transcendental equation (8) must be solved numerically. Substituting Eq. (7) into Eq. (6) and transforming from z to $u=\ln(R/z)$, we evaluate the resulting integral over u by the saddle-point method. As a result, only the integration over ε remains:

$$Q_t=2\pi CR\int\limits_0^\infty \frac{d\varepsilon\varepsilon^{1/2}\exp\{-f(u_0,t_0)\}}{\{f_{tt}''(u_0,t_0)f_{uu}''(u_0,t_0)\}^{1/2}}\frac{\exp(-u_0)}{\{1-e^{-2u_0}\}^{1/2}}, \tag{10}$$

where the saddle point u_0 satisfies the equation

$$1+\frac{t_0}{u_0^{1/2}}\left(\frac{ERe^{-u_0}}{(1-\exp(-2u_0))^{1/2}}-\frac{t_0}{u_0^{1/2}}\right)=0. \tag{11}$$

In Eqs. (10) and (11) the dependence $u(z)$ in the form (3) was employed. Since for $\psi\gg 1$ small values of ε make the main contribution to the integral over ε, we obtain finally for the normalized charge density

$$\tilde{Q}_t=\frac{Q_t}{{}^1/_2\pi^{1/2}R(\sigma/4\pi a^2)^{3/4}}=\frac{2^{1/2}}{ER}$$
$$\times\frac{(1-x^2)(1-x^2T^2)\exp(-u_0x^2)}{x^{1/2}\left\{(1-x^2)T^2+(1-x^2T^2)\left[(1-x^2)u_0+\frac{1}{2}\left(x+\frac{1}{x}\right)^2\right]\right\}^{1/2}}. \tag{12}$$

Here $u_0=\frac{1}{2}\ln(1+(ERx/(1-x^2))^2)$, and the variable $x=t_0/u_0^{1/2}$ satisfies the equation

$$\tanh\left\{\frac{\psi+(1-x^2)/x-x\ln[1+\{ERx/(1-x^2)\}^2]}{2T}\right\}=xT,$$

obtained from the system of equations (8) and (11) for t_0 and u_0. In Eq. (12) Q_t is normalized so that at low temperatures $\tilde{Q}_t\to\exp(-\psi)$.

The results of a series of calculations, performed using the formula (12), are presented on a semilogarithmic scale in Fig. 2. We set $ER=1$, since for not too small values of Q_t the normalized electric field is equal to[6]

$$E=\frac{\varepsilon_i+\varepsilon_s}{\varepsilon_i}\frac{Q_t}{(\sigma/\pi)^{1/2}}$$

and therefore $ER=1+\varepsilon_i/\varepsilon_s\approx 1$. The solid lines in Fig. 2 correspond to different values of the parameter ψ. We note that for $T=1$ ($T=\Delta$ in dimensional units) the quantity Q_t is three to five times greater than its asymptotic value at zero temperature.

4. We now consider the delocalized charge (5), i.e., we determine the average number of electrons which for fixed T and ψ have energies lying above the percolation level. For simplicity we assume $\psi>T$, in order to be able to approximate the distribution function with a Boltzmann distribution. This immediately enables us to separate the characteristic exponential factor $\exp(-\psi/T)$ in the formula for Q_d. Evaluating the integral over φ exactly (this integral can be expressed in terms of the error function), we evaluate the integral over the energy ε and the variable u by the saddle-point method. The result for the normalized delocalized charge is

$$\tilde{Q}_d=\varphi(\beta_0)T^{1/2}(T/ER)\exp(-\psi/T), \tag{13}$$

where the function $\varphi(\beta_0)$ is determined in terms of the error function $\Phi(x)$ and elementary functions:

$$\varphi(\beta_0)=\frac{2^{3/2}e^{\beta_0^2/4+1/2}}{(3-2e^{-\beta_0^2})^{1/2}}$$
$$\times\begin{cases} e^{-1/2}\left[1-\Phi\left(\frac{\beta_0}{2}-\frac{3}{2\beta_0}\right)\right], \\ \frac{ERe^{-u_0}}{(1-e^{-2u_0})^{1/2}}=T, \quad \beta_0=\frac{2u_0^{1/2}}{T}<1 \\ \beta_0^{5/2}e^{-1/2\beta_0}\left[1-\Phi\left(\frac{\beta_0}{2}-\frac{3}{2}\right)\right], \\ \frac{ERe^{-u_0}}{(1-e^{-2u_0})^{1/2}}=2u_0^{1/2}, \quad \beta_0>1. \end{cases}$$

In the limit of high temperatures we have $\varphi(\beta_0)=2^{5/2}/e$ and $Q_d=Q_t$, calculated according to the formula (12). At low temperatures $\varphi(\beta_0)\sim(2u_0^{1/2}/T)^{3/2}$ and the temperature dependence of Q_d is similar to that obtained in Ref. 6:

$$\tilde{Q}_d\sim\frac{T}{ER}\exp\left(-\frac{\psi}{T}\right).$$

The computational results obtained using the formula (13) correspond to the dashed lines in Fig. 2, which correspond to the same values of the parameter ψ.

5. Now the values of the charges Q_t and Q_d calculated according to Eqs. (12) and (13) must be grouped together so as to construct the desired dependence $Q_d(Q_t)$, or equivalently, $Q_d(V-V_t)$ at fixed temperature. In so doing, the dependence $R\sim Q_t^{-1}$ must also be taken into account. The corresponding curves are shown in Figs. 3 and 4.

At low temperatures the delocalized charge Q_d, remaining much less than Q_t in virtually the entire interval $0<Q_t<Q_t^{\max}$, increases in a power-law fashion with increasing total charge Q_t: $\ln Q_d\propto T^{-1}\ln Q_t$, which repro-

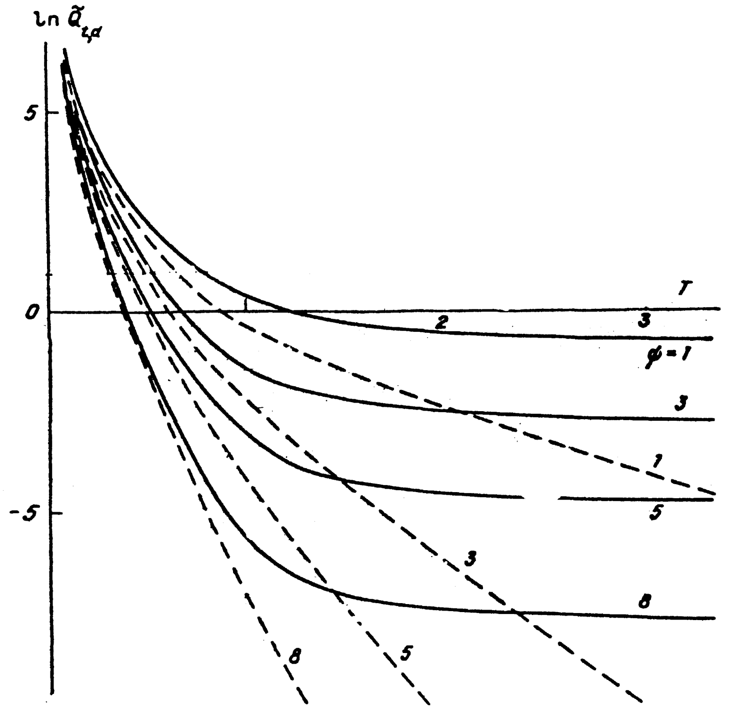

FIG. 2. Temperature dependence of the total charge Q_t(—) and the delocalized charge Q_d(- - -) for fixed $\psi = \varphi_s - \mu$ (φ_s is the band bending).

duces to a significant degree the threshold character of the dependence $Q_d(Q_t)$ at $T=0$.

It is interesting that for intermediate temperatures ($0.5 \approx T_1 < T < T_2 \approx 2$) the charges under consideration are virtually proportional to one another, similar to the high-temperature asymptotic behavior, but with a smaller slope ($Q_d/Q_t \approx 0.4T < 1$). This means that in this temperature interval, as the gate voltage increases, in addition to an increase in charge mobility, the stationary charges further charge the fluctuation surface states at the same time. It is obvious that because of the direct proportionality which has been shown to exist, this will be perceived experimentally as a corresponding decrease in the effective electron mobility.

Thus the extension, elaborated here, of the theory of fluctuation surface states in MIS structures[6] to the case of finite temperatures explains the basic features of the surface conductivity of inversion layers: the decrease in the surface mobility as compared with the volume mobility, the weakening of the temperature dependence of the surface mobility, and the low-temperature "jump" in the threshold voltage of an MOS transistor. In particular, the well-known difference between the surface mobility of electrons in silicon MOS transistors and the bulk value of the mobility (by approximately a factor of three at room temperature) can be easily interpreted on the basis of the theoretical model developed here. For this, it is sufficient to assume that the density of charged defects $\sigma = \sigma^+ + \sigma^-$ at the Si–SiO_2 interface is equal to $4 \cdot 10^{11}$ cm^{-2}, which falls into the range of built-in charge densities typical for Si–SiO_2.

It is also natural to assume that the characteristic established here—the existence of bound charge at intermediate temperatures—should also be strongly manifested in the

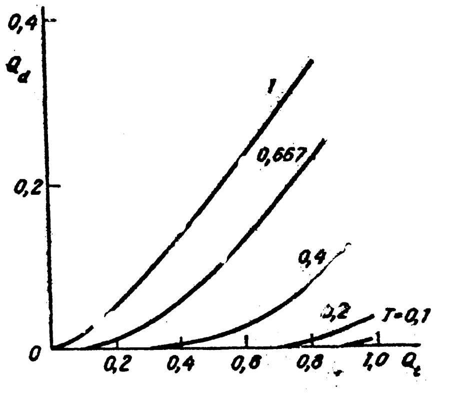

FIG. 3. Delocalized charge versus the total charge $Q_d(Q_t)$ at low temperature.

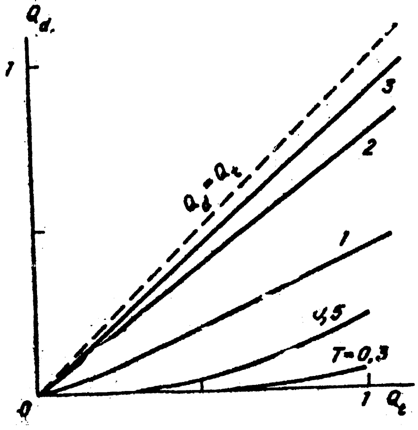

FIG. 4. Delocalized charge versus the total charge $Q_d(Q_t)$ at fixed temperature.

transient dynamics of MIS transistors and it should also have a determining effect on the noise characteristics of devices with surface transport of charge.

1) In our notatation φ_s and $\mu < 0$ (see Fig. 1), since the energy is measured from the bottom of the conduction band in the volume of the semiconductor.

[1] B. E. Deal, J. Electrochem. Soc. **121**, 198 (1974).
[2] O. V. Romanov, Fiz. Tekh. Poluprovodn. **16**, 419 (1982) [Sov. Phys. Semicond. **16**(3), 266 (1982)].
[3] V. G. Litovchenko and A. P. Gorban', [title omitted], [in Russian], Naukova dumka, Kiev, 1978.
[4] I. M. Lifshitz, S. A. Gredeskul, and L. A. Pastur, [title omitted], [in Russian], Nauka, Moscow, 1982.
[5] B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors*, Springer, New York, 1984.
[6] V. A. Gergel' and R. A. Suris, Zh. Eksp. Teor. Fiz. **84**, 719 (1983) [Sov. Phys. JETP **57**(2), 415 (1983).

Translated by M. E. Alferieff