Fluctuation slow surface traps

B. I. Fuks

Radio Engineering and Electronics Institute, Russian Academy of Sciences
(Submitted 12 December 1991; revised 7 April 1992)
Zh. Eksp. Teor. Fiz. **102**, 555–576 (August 1992)

It is shown that in a semiconductor-insulator interface there are traps with exceptionally low free-carrier capture cross sections σ. They form because of random fluctuations of the number density of charged centers built into the insulator along this interface. A slow trap is created by a large-scale repulsive fluctuation of centers, which forms the barrier, and a cluster of attractive centers, the "nucleus," that finds itself in the fluctuation. A relation between the parameters of these traps and the probability of their appearance is established for different temperatures T and surface carrier densities $\overline{Q}$. It is shown that many mesoscopic MOSFETs (metal–oxide–semiconductor field effect transistors) may contain, at certain values of T and $\overline{Q}$, a single trap in which charge exchange generates a noise current of the "random telegraphic signal" (RTS) type with characteristic times ranging from $\sim 10^{-3}$ to 10^2 s. The temperature dependence of σ for this trap is specified by the equation $\sigma(T) = \sigma_0 \exp(-\Delta E_B/T)$, with $\Delta E_B \gg T$. The most probable values of σ_0 and ΔE_B that follow from the theory, their variations, and the variations of the values of $\overline{Q}$ optimal for the emergence of RTS when the temperature range changes resemble those observed in RTS studies. The basis of the theory of RTS generated in a mesoscopic MOSFET by a single slow fluctuation surface trap is constructed. The properties of such a trap are found to differ considerably from those of other electron traps owing to the special structure of the nucleus and barrier and the great difference in size (a typical barrier radius l_0 is several dozen nanometers, while that of a nucleus ranges from one to two nanometers). Also determined are the RTS amplitude and its dependence on the temperature and the voltage across the gate for different ratios of l_0 and the oxide layer thickness, similar to the dependence of the carrier capture and emission times at the trap level. Analysis of RTS data suggests that this dependence be explained only by assuming that these are fluctuation traps. Finally, a number of important consequences of the fluctuation origin of slow surface traps and the exceptionally high information capacity of RTS are noted.

1. INTRODUCTION

Advances in manufacturing solid-state nanostructures have stimulated great interest in studying the properties of such small objects. They are created not only artificially, by technological means, but also by nature itself. This paper discusses the physical consequences of the existence in the semiconductor-insulator heterojunction plane of random nanostructures comprised of charged centers, structures that emerge because of the random distribution of such centers and generate surface traps in semiconductors with extraordinarily small free-carrier capture cross sections.

To be specific, we examine the Si:SiO_2 system, which is characterized by the high quality of the heterojunction. Despite this, SiO_2 always contains a sizable built-in space charge, which, as is well-known (e.g., Refs. 1 and 2), is concentrated near the interface with Si (no farther than 30 Å from the interface, this value being only the resolution limit of the measurements) and the charge can be expected to be in the junction region in Si:SiO_2 whose thickness is about one to two atomic layers[3]). The value of this space charge, measured by the variation in the flat-band potential,[1,2] is the result of a fairly exact balance between the relatively large positive and negative built-in charges. This can be assumed from experiments in the localization of carriers at low temperatures.[4] The proximity to the semiconductor and the high total number density of the charged centers lead to a situation in which density fluctuations along the interface create a random potential pattern with a high amplitude, which localizes electrons and holes at the semiconductor's surface.[5] The surface states, or traps, considered in Ref. 5 are "fast." But fluctuations of the built-in charge also form "slow" surface traps. These generate small-scale clusters of attractive centers, or "nuclei," that find themselves inside the large-scale repulsive fluctuation. The multiply charged nucleus localizes the captured particle inside a very small volume surrounded by a high and extended potential barrier. The capture time exponentially increases with the height of the barrier, and for strong barriers created by large, and therefore infrequent, fluctuations can be very long.

There are noticeably fewer slow surface traps than fast traps. But since the probabilities of the formation of nuclei with a high binding energy of the captured particle and of strong repulsive barriers are not very low,[6] there are enough such traps for a sizable flicker surface noise in ordinary metal–oxide–semiconductor field-emission transistors (MOSFETs) and for the possibility of observing at least one slow trap in transistors of submicron dimensions.

In their properties slow fluctuation traps most closely resemble repulsive centers, such as Au in Ge or Zn in Si, whose capture cross section is determined by the tunneling of high-energy free charge carriers (with an energy $\varepsilon \gg T$) under a fairly low barrier.[7] But there are considerable differences here. For instance, capturing by fluctuation traps occurs because of the direct tunneling of a particle to one of the excited (resonance) states created by the attractive nucleus. Only after it has given away its energy does the particle go to the ground state. Resonance states have no repulsive centers, and owing to indirect tunneling the particle immediately finds itself in the ground state.[8] The considerable diversity in

 0038-5646/92/080294-12$05.00

the properties of fluctuation traps introduces a very great difference between the radii of the nucleus and barrier. The typical value of the first ranges from one to two nanometers, and that of the second is on the order of the thickness of the insulator in the MOS structure, that is, several dozen nanometers, and differs considerably from trap to trap owing to the randomness of their formation. Since the nucleus and barrier, so different in scale and the effect they have on the particle, together form a slow trap, the properties of this trap must contain various information about the heterojunction and the adjacent layers of the semiconductor and insulator.

Investigations in the conduction of mesoscopic MOSFETs stimulated progress in the study of the properties of slow surface traps. Ralls *et al.*[9] discovered random current switching corresponding to jumps in the conductivity in this extremely small sample under charge exchange in the only slow surface trap in the sample with times of capture and emission of electrons, τ_c and τ_e, of the order of 1 s and longer, and that had the shape of a random telegraphic signal (RTS). Many experiments (see, e.g., the review article by Kirton and Uren[10]) showed that $1/f$ noise in a MOSFET of ordinary dimensions constitutes a combination of RTSs emerging as a result of charge exchange in a large number of such traps. But it may be more important that the study of the RTS produced by a single trap proved to be exceptionally informative and the properties of traps found from RTS studies have been quite unexpected and yield with difficulty to explanations based on the ordinary ideas of electron traps (see Ref. 10 and Secs. 7 and 8 of this paper).

Further analysis shows that fluctuation traps possess just such properties. Below we determine for such traps the optimal structure of the nucleus and shape of the energy barrier and then the times τ_c and τ_e for a trap with a given nucleus and barrier. Since the most complete information about slow traps is provided by RTS studies, subsequent determination of the probability for these traps to appear, of the most probable properties of traps, and of the optimal conditions for observing traps with τ_c and τ_e lying in the given range have been done for MOSFETs of submicron dimensions. The basic laws governing the behavior of the RTS generated by the charge exchange in the fluctuation trap are also determined; namely, the RTS amplitude and its dependence on temperature and voltage on the field electrode of the MOS structure for different ratios of barrier radius to insulator thickness and similar behavior for τ_c and τ_e. The experimental RTS data are then discussed and it is shown that interpreting the data correctly requires using the concept of fluctuation traps. Finally, we discuss the new possibilities opened up by the remarkable richness of the RTS data provided by the charge exchange in the slow fluctuation surface traps, and of other important corollaries of the agreement between theory and experiment.

2. STRUCTURE OF NUCLEUS AND SHAPE OF BARRIER OF A SLOW SURFACE TRAP

Consider the localization of particles (electrons from the inversion n-type channel of the MOS structure, for the sake of definiteness) in the potential pattern created by the charges built into the interface between the semiconductor and the insulator and distributed at random along the interface. The average surface densities of the positively and negatively charged centers are denoted by Σ_+ and Σ_-. On the basis of this statement of the problem, Gergel' and Suris[5] have shown that the fluctuations of the built-in charge generate localized surface states for electrons and holes. The expression for the density of surface states at the Fermi level, $N_{SS}(\varepsilon_F)$, derived in the Gaussian approximation and therefore depending only on the total center density, $\Sigma = \Sigma_+ + \Sigma_-$, provides at $\Sigma \sim 10^{12}$ cm^{-2} a good description of the $N_{SS}(\varepsilon_F)$ dependence for Si:SiO$_2$ structures, with the exception of values of ε_F close to the middle of the forbidden band. Gergel' and Suris's theory, which allows only for quasiclassical localization of carriers in large-scale potential fluctuations, states that for these values of ε_F the $N_{SS}(\varepsilon_F)$ curve acquires a sharp dip.

Actually, no such dip should exist owing to the large density $\rho_{\text{quantum}}(\varepsilon')$ of the quantum states generated by small-scale, Poisson (non-Gaussian) clusters of attractive centers. Such a "nucleus," with an electron binding energy ε', consists of a small number $Z_{\varepsilon'}$ of centers inside a disk whose radius is close to that of the wave function, $a_{\varepsilon'}$. For estimates we can assume $Z_{\varepsilon'} \sim (\varepsilon'/\varepsilon_1)^{1/2}$ and $a_{\varepsilon'} \sim a_1(\varepsilon_1/\varepsilon')^{1/2}$ for $\varepsilon' \gg \varepsilon_1$. Here $\varepsilon_1 = me^4/8\varkappa^2\hbar^2$ and $a_1 = 4\varkappa\hbar^2/me^2$ are the binding energy and the scale of the wave function the ground state of an electron of mass m on an isolated charge $+e$ built into the interface (see Refs. 3 and 5), and $\varkappa = (\varepsilon_s + \varepsilon_i)/2$, where ε_s and ε_i the dielectric constants of the semiconductor and insulator (we assume in numerical estimates that $\varepsilon_s = 12$, $\varepsilon_i = 4$, $m = 4\times 10^{-28}$ g, $\varepsilon_1 = 0.025$ eV, and $a_1 = 35$ Å). An estimate for $Z_{\varepsilon'}$ and $a_{\varepsilon'}$ follows from reasoning similar to that used in Refs. 6 and 11 in determining the optimal nuclei in the bulk of a doped semiconductor. The same reasoning was used in Ref. 5, but the conclusion reached there that $\rho_{\text{quantum}}(\varepsilon')$ is negligible was based on the asymptotic behavior of quasipoint nuclei with a radius small compared to $a_{\varepsilon'}$, suggested in Ref. 11. As shown in Ref. 6, this asymptotic behavior can be employed only for unrealistically large values of ε'. For practical values of ε' the radius of an optimal nucleus considerably exceeds $a_{\varepsilon'}$. Hence, the value of $\rho_{\text{quantum}}(\varepsilon')$ proves to be much higher than the one calculated in the asymptotic approximation. For instance, Ref. 6 shows that if the bulk donor concentration N satisfies $\widetilde{N} = (4\pi/3)Na_B^3 \sim 1$, with a_B the Bohr radius in the bulk of the semiconductor, the asymptotic formula understates $\rho_{\text{quantum}}(\varepsilon')$ by approximately ten orders of magnitude. It is then natural to expect that for a two-dimensional distribution of centers with $\widetilde{\Sigma}_+ = \pi\Sigma_+ a_1^2 \sim 1$, that is, with $\Sigma_+ \sim 10^{12}$ cm^{-2}, the asymptotic formula of Ref. 5 understates $\rho_{\text{quantum}}(\varepsilon')$ by $(2/3)10 \approx 6$–7 orders of magnitude. Besides, for high ε' (small $a_{\varepsilon'}$) the spatial dispersion of ε_s and ε_i must provide a considerable contribution to ε', since highly localized electrons are coupled very effectively by the potential of the cell containing a charged center. As is known,[12] in silicon this contribution is considerable even for single charged centers and very high (~ 1 eV) for double charged centers. Clearly, the contribution of spatial dispersion, diminishing as the charges from a common central cell are dispersed, is easily restored by an increase in the number of charges in a region of radius $\sim a_{\varepsilon'}$. Because of this effect the rate at which $\rho_{\text{quantum}}(\varepsilon')$ decreases for large values of ε' must drop drastically. These facts and the results of preliminary calculations carried out by the present author together with N. M. Storonskiĭ via the refined formulas of Ref. 13 seriously suggest that in Si:SiO$_2$ with $\Sigma_+ \sim \Sigma_- \sim 10^{12}$ cm^{-2}, the value of $\rho_{\text{quantum}}(\varepsilon')$ amounts to roughly 10^9–10^{10} eV^{-1} cm^{-2} for both electrons and holes even at $\varepsilon' \sim 0.5$ eV and slowly decreases as ε' increases and Σ_+ and Σ_- decrease.

When measured from the middle level of the bottom of the conduction band at the interface (in what follows all energies are measured from this level), the energy of the ground state of an electron bound by the nucleus, ε_t, consists of two terms, $-\varepsilon'$ and the shift (primarily classical) in the field of the charges separated from the nucleus by a distance much greater than $a_{\varepsilon'}$. Infrequent, slow surface traps with multiply charged nuclei surrounded by strong large-scale repulsive fluctuations, have a moderate (or even negative) binding energy because of the balance of these two large terms. For this reason their contributions to $N_{SS}(\varepsilon_F)$ is small. But such traps (as well as similar bulk traps discussed in Ref. 6) manifest themselves vividly in nonequilibrium phenomena owing to very slow charge exchange.

Now let us find the large-scale fluctuation $\xi(\mathbf{l})$, where $\mathbf{l}$ is the vector in the interface plane, that generates with the highest probability a barrier with a given tunneling transmission coefficient for free electrons with energy ε_T captured by the given nucleus. Near the nucleus but for $r \gg a_{\varepsilon'}$ the height of the barrier is

$$U(\mathbf{r}) = -\frac{Z_{\varepsilon'}e^2}{\varkappa r} + \frac{e^2}{\varkappa}\int^{\Omega_L} d\mathbf{l}\,\frac{\xi(\mathbf{l})-\bar{Q}}{|\mathbf{r}-\mathbf{l}|}, \tag{1}$$

where Ω_L is the area of the barrier area and $\bar{Q} = \bar{Q}_l + \bar{Q}_d$, with $\bar{Q}_l$ and $\bar{Q}_d$ the average surface densities of localized and delocalized (free) electrons. For $\bar{Q}_l \gg \bar{Q}_d$ Eq. (1) holds for $\varepsilon_T \gg T$ and $\varepsilon_T \gg \Delta$, where $\Delta = (e^2/\varkappa)(\pi\Sigma)^{1/2}$ is the scale of the fluctuation pattern. For $2\Delta < -\varepsilon_F < \Delta \ln(8d^2\Sigma^{1/4}/a_1^{3/2})$, where d is the insulator thickness, $d \ll \varepsilon_i w/\varepsilon_s$, and w is the thickness of the space-charge region separating the n-channel from the p-substrate, the following formulas for $\bar{Q}_l$ and $\bar{Q}_d$ hold[5]

$$\bar{Q}_l = (2/a_1)^{1/4}(\Sigma/\pi)^{5/8}\exp(\varepsilon_F/2\Delta), \tag{2}$$

$$\bar{Q}_d = (16\Sigma/\pi a_1^2)^{1/4}(T/eE_s)\exp(\varepsilon_F/T), \tag{3}$$

where E_s is the homogeneous field pressing the electrons to the surface of the interface,

$$E_s = (2\pi e/\varepsilon_s)(2N_a w + \bar{Q}_l), \tag{4}$$

and N_a is the concentration of acceptors in the substrate.

Another condition for the applicability of Eq. (1) is the inequality $l_0 \ll d$, where l_0 is the radius of the barrier. It is met when $\bar{Q}$ is not very small and allows for ignoring image charges in Eq. (1). Here Ω_L and $\xi(\mathbf{l})$ are linked by the condition that the total charge on Ω_L be zero:

$$\int^{\Omega_L} d\mathbf{l}[\xi(\mathbf{l}) - \bar{Q}] = 0, \tag{5}$$

where we have ignored the charge $Z_{\varepsilon'}$ (Eq. 5). The exponent in the tunneling transmission coefficient for electrons with an energy ε, $\Lambda(\Sigma)$, moving in the direction of fastest tunneling to the nucleus (along the x axis at right angles to the interface; see below) is

$$\Lambda(\varepsilon) = 2^{3/2}\frac{m^{1/2}}{\hbar}\int_{x_1(\varepsilon)}^{x_2(\varepsilon)} dx\left[-\frac{Z_{\varepsilon}e^2}{\varkappa x} + \frac{e^2}{\varkappa}\int^{\Omega_L} dl\,\frac{\xi(\mathbf{l})-\bar{Q}}{(x^2+l^2)^{1/2}} - \varepsilon\right]^{1/2}, \tag{6}$$

where we have allowed for Eq. (1), and $x_1(\varepsilon)$ and $x_2(\varepsilon)$ are the roots of the equation $U(x,\mathbf{l}=0) = \varepsilon$. By employing formulas (5) and (6) and the requirement that entropy be minimal, we can obtain, via a variational procedure, an equation for the optimal fluctuation $\hat{\xi}(\mathbf{l})$ for given $\varepsilon = \varepsilon_T$, $\Lambda(\varepsilon) = \Lambda_T$, and $Z_{\varepsilon'}$. This nonlinear integral equation can be solved only numerically, but it allows us to clarify a number of general laws governing the behavior of $\hat{\xi}(\mathbf{l})$, namely, that $\hat{\xi}(\mathbf{l})$ possesses radial symmetry; for $x_2(\varepsilon_T) \ll l \ll l_0$ we have $\hat{\xi}(l) \propto [u(l) - u(l_0)]$, with the potential $u(\mathbf{r})$ of a separate center being approximately $e^2/\varkappa r$, and for $l \ll x_1(\varepsilon_T)$ the value of $\hat{\xi}(l)$ is practically independent of l. Hence we seek the optimal fluctuation below in the form

$$\hat{\xi}(l) = B[(y^2+l^2)^{-1/2} - (y^2+l_0^2)^{-1/2}]. \tag{7}$$

In this way the problem reduces to finding the optimal values of B, y, and l_0. Naturally, the probability of slow-trap formation is underestimated only slightly because the form of $\hat{\xi}(l)$ is not known precisely. Assuming that y, $x_1(\varepsilon_T)$, and $x_2(\varepsilon_T)$ are much smaller than l_0 (this assumption is verified below), we find from (5) and (7) that $B = \bar{Q}l_0$. The entropy of a large-scale fluctuation calculated in the Gaussian approximation is

$$S = \pi\int_0^{l_0} dl\,l[\hat{\xi}^2(l)/\Sigma] = (\pi\bar{Q}^2 l_0^2/\Sigma)[\ln(l_0/y) - (3/2)], \tag{8}$$

and Eq. (6) acquires the form

$$\Lambda_T = 2^{3/2}\frac{m^{1/2}}{\hbar}\int_{x_1(\varepsilon_T)}^{x_2(\varepsilon_T)} dx\left[\frac{2\pi e^2\bar{Q}l_0}{\varkappa}\left(\ln\frac{2l_0}{x+y} - 2\right) - \varepsilon_T - \frac{Z_{\varepsilon'}e^2}{\varkappa x}\right]^{1/2}. \tag{9}$$

Expressing y in terms of l_0 and S via Eq. (8) and substituting the result into (9), we get the function $\Lambda_T(S,l_0)$, which increases with S, as can easily be shown. Hence, the minimum in the entropy, S_m, for given Λ_T, ε_T, and $Z_{\varepsilon'}$ corresponds to the maximum in $\Lambda_T(S,l_0)$ as a function of l_0, that is, S_m can be found by solving the equations $\Lambda_T(S,l_0) = \Lambda_T$ and $\partial\Lambda_T(S,l_0)/\partial l_0 = 0$ simultaneously. Without going into the details of the calculations involved, we give the final result:

$$S_m \approx (\tilde{Q}^2\tilde{l}_0^2/\tilde{\Sigma})[1 + (\tilde{\varepsilon}_T/4\tilde{Q}\tilde{l}_0)],$$
$$y/l_0 \approx \exp[-(5/2) - (\tilde{\varepsilon}_T/4\tilde{Q}\tilde{l}_0)], \tag{10}$$

where $\tilde{l}_0 \equiv l_0/a_1$ is the solution to the equation

$$\Lambda_T = 4Z_{\varepsilon'}\tilde{Q}^{-1/2}\tilde{l}_0^{-1/2}b^{-1}$$
$$\times(\tilde{l}_0)\int_{\tilde{x}_1}^{\tilde{x}_2} d\tilde{x}\,[1/2 + \ln 2 - \ln(1+\tilde{x}) - b(\tilde{l}_0)/\tilde{x}]^{1/2}. \tag{11}$$

Here the following dimensionless quantities have been introduced: $\tilde{\varepsilon}_T = \varepsilon_T/\varepsilon_1$, $\tilde{\Sigma} = \pi\Sigma a_1^2$, $\tilde{Q} = \pi\bar{Q}a_1^2$, $\tilde{x} = x/y$, and

$$b(\tilde{l}_0) = (Z_{\varepsilon'}/2\tilde{Q}\tilde{l}_0^2)\exp[(\tilde{\varepsilon}_T/4\tilde{Q}\tilde{l}_0) + 5/2]),$$

and $\tilde{x}_1$ and $\tilde{x}_2$ are the roots of the equation

$$\ln(1+\tilde{x}) + b(\tilde{l}_0)/\tilde{x} = 1/2 + \ln 2$$

with $b(\tilde{l}_0) < 1/2$ and $\hat{x}_1 < 1 < \tilde{x}_2$. From Eq. (1) we find that the optimal barrier $\hat{U}(\mathbf{r})$ for $r \ll l_0$ consists of two terms, of which the spherically symmetric term

$$U_1(r) = -Z_{\varepsilon'} e^2/\varkappa r + \varepsilon_T + 4\tilde{Q}\tilde{l}_0\varepsilon_1[{}^1/_2 + \ln 2 - \ln(1 + r/y)], \quad (12)$$

is the principal one, while the other term, $U_2(\mathbf{r})$, depends on the angle θ between $\mathbf{r}$ and the x axis and for $r \ll y$ is given by the following formula:

$$U_2(\mathbf{r}) = 4\tilde{Q}\tilde{l}_0\varepsilon_1(r/y)(1 - \cos\theta). \quad (13)$$

For small values of $\overline{Q}$ this approach is inapplicable, since as the screening weakens the radius of an optimal fluctuation grows and exceeds the insulator's thickness. Then the strength of the barrier is limited by screening due to the image charges induced on the field electrode (gate) of the MOSFET. Note that $\hat{\xi}(l)$ does not change for $l \ll x_1(\varepsilon)$, and for $x_2(\varepsilon) \ll l \ll l_0$ we have $\hat{\xi}(l) \propto u(l)$. But since the potential $u(\mathbf{r})$ of the center, which for $r \ll d$ is proportional to r^{-1}, decreases much faster for $r \gg d$ because of screening, instead of (7) we seek $\xi(l)$ in the form $\hat{\xi}(l) = B(y^2 + l^2)^{-1/2}$ for $l < \eta d$ and $\hat{\xi}(l) = 0$ for $l > \eta d$, with $\eta \sim 1$. [In subsequent numerical estimates we assume that $\eta = 2$, since at $\varepsilon_s = 12$ and $\varepsilon_i = 4$ the difference between $u(\mathbf{r})$ and $e^2/\varkappa r$ is moderate as long as $r \leq 2d$ holds.] The problem now is to find the optimum values of B and y, and its solution in many respects is the same as in the previous case. The value of B is determined by an equation resembling (11):

$$\Lambda_T = 4Z_{\varepsilon'}\tilde{B}^{-1/2}b_1^{-1}(\tilde{B})\int_{\tilde{x}_1}^{\tilde{x}_2} dx\,[{}^1/_2 + \ln 2 - \ln(1 + \tilde{x}) - b_1(\tilde{B})/\tilde{x}]^{1/2}, \quad (14)$$

where $\tilde{B} = \pi B a_1$, and

$$b_1(\tilde{B}) = (Z_{\varepsilon'}/2\eta\tilde{d}\tilde{B})\exp[{}^1/_2 + (\tilde{\varepsilon}_T/4\tilde{B})],$$

with $\tilde{d} = d/a_1$. The following equation links y with B:

$$y/d = \eta\exp[-{}^1/_2 - (\tilde{\varepsilon}_T/4\tilde{B})].$$

The expression for S_m has the form

$$S_m = \tilde{B}^2/2\tilde{\Sigma} + \tilde{\varepsilon}_T\tilde{B}/4\tilde{\Sigma}, \quad (15)$$

and the potentials $U_1(\mathbf{r})$ and $U_2(\mathbf{r})$ are specified as follows:

$$U_1(r) = -Z_{\varepsilon'}e^2/\varkappa r + \varepsilon_T + 4\varepsilon_1\tilde{B}[{}^1/_2 + \ln 2 - \ln(1 + r/y)] \quad \text{for} \quad r \ll \eta d, \quad (12a)$$

$$U_2(\mathbf{r}) = 4\varepsilon_1\tilde{B}(1 - \cos\theta)r/y \quad \text{for} \quad r \ll y. \quad (13a)$$

Here are the main aspects of determining the optimal fluctuations in the case of large $\overline{Q}$, for $\overline{Q}_d \gg \overline{Q}_l$, which for situations with $\tilde{\Sigma} \sim 1$ is equivalent to the condition $\overline{Q} \gg \Sigma$. the large-scale fluctuation that generated the barrier cannot be considered in the Gaussian approximation in this case. The capturing process here is characterized by the fact that the electrons are pressed to the interface by the strong electric field and, consequently, are quasi-two-dimensional and tunnel to the trap in the l-plane.[1)] Also, as an electron moves toward the nucleus, the barrier becomes still higher because of the weakening of the interelectron interaction, which is considerable for large $\overline{Q}$ (see Ref. 3).

3. CAPTURE AND EJECTION OF CARRIERS BY A SLOW TRAP

A potential well that finds itself inside a potential barrier can generate excited, quasidiscrete, states in addition to the ground state. Assuming the nucleus to be a point and ignoring the variation in barrier height within the well, we find that the energies of these excited states are

$$\varepsilon_{\hat{n}} = \varepsilon_T + 4\varepsilon_1 K({}^1/_2 + \ln 2) - 4\varepsilon_1 Z_{\varepsilon'}^2/(n + L + 1)^2, \quad (16)$$

where $\hat{n}$ stands for a set of quantum numbers including the radial quantum number n, the azimuthal quantum number L, and the magnetic quantum number M, and $K = \tilde{Q}\tilde{l}_0$ if $l_0 \ll d$ and $K = \tilde{B}$ if $l_0 \gg d$. Of all the $\hat{n}$ only those correspond to solutions to our problem at which the wave functions vanish at the interface with the insulator (at $x = 0$), that is, the difference $L - M$ must be an odd number, and the values of n and L are limited by the condition that the energies $\varepsilon_{\hat{n}}$ lie below the top of the barrier, $U_1(r) + \hbar^2(L + 1/2)^2/2mr^2$. In the quasiclassical approximation this yields

$$n + L + 1 \leq Z_{\varepsilon'}(2Kb^{1/2})^{-1/2}, \quad (17)$$

where $b = \tilde{b}(l_0)$ if $l_0 \ll d$ and $b = b_1(\tilde{B})$ if $l_0 \gg d$. For the optimal fluctuations studied in this paper, the right-hand side of (17) at high temperatures is large compared to unity. Here the electric field of the barrier produces a fairly moderate shift (decrease) in the energies $\varepsilon_{\hat{n}}$ satisfying (17). Understandably, the substitution of a point charge $Z_{\varepsilon'}$ for the optimal nucleus leads to an error in determining $\varepsilon_{\hat{n}}$, and the greater the quantity by which the radius of the wave function of state $\hat{n}$ exceeds the radius of the optimal nucleus, that is, the greater $n + L$ is, the smaller the error. For the ground state (with $n = 0$, $L = 1$, and $M = 0$) the last term in (16) should be replaced with $-\varepsilon'$:

$$\varepsilon_t \equiv \varepsilon_{0,1,0} = \varepsilon_T + 4\varepsilon_1 K({}^1/_2 + \ln 2) - \varepsilon'. \quad (18)$$

As the temperature is lowered, the important values of $\overline{Q}$ grow and those of $Z_{\varepsilon'}$ drop (Sec. 4). In the process the right-hand side of Eq. (17) and the number of resonance states decrease and the shift in $\varepsilon_{\hat{n}}$ induced by the barrier field and the average field that presses the electrons to the surface increases. On the Si[100] surface, however, even at high values of $\overline{Q}$ the slow electron traps have resonance states that have split off from the ground state because of the lifting of many-valley degeneracy by the requirement that the wave function at the insulator surface vanish.[3]

Now let us calculate the electron capture and ejection timer. First we determine the time that it takes an electron to leave the well from a resonance level $\hat{n}$. If we employ the quasiclassical approximation and ignore the potential $U_2(\mathbf{r})$, this time $\tau_{\hat{n}}$ can easily be found:[14]

$$\tau_{\hat{n}} = T_{n,L}\exp\left\{(2^{3/2}m^{1/2}/\hbar)\int_{r_1}^{r_2} dr'\,[U_1(r') + \hbar^2(L + {}^1/_2)^2/2mr'^2 - \varepsilon_{\hat{n}}]^{1/2}\right\}, \quad (19)$$

where $T_{n,L}$ is the period of the classical motion of an electron in state $\hat{n}$,

$$T_{n,L} \approx \pi\hbar(n + L + 1)^3/4\varepsilon_1 Z_{\varepsilon'}^2. \quad (20)$$

The "tunneling" exponent in (19) differs from (6) in that it is augmented by the centrifugal energy $\hbar^2(L + 1/2)^2/2mr^2$.

But because this energy rapidly decreases as r grows, it contributes little to the integral along the section from r_1 to r_2.

We now determine the time τ_e required to eject an electron from a slow trap. The probabilities for the electron to be on the resonance levels for which the time of transfer to other levels is short compared to $\tau_{\hat{n}}$ obey an equilibrium distribution

$$W_{\hat{n}}=g^{-1}\exp[-(\varepsilon_{\hat{n}}-\varepsilon_t)/T],$$

where we have assumed that these levels lie noticeably significant than the ground level ($\varepsilon_{\hat{n}}-\varepsilon_t \ll T$), whose degree of degeneracy is g, that is, $W_{0,1,0}=g^{-1}$. This implies that the time τ_e' to eject an electron through such "equilibrium" resonance states is given by the following formula:

$$\tau_e'^{-1}=g^{-1}\sum_{[n]}\tau_{[n]}^{-1}\exp[-(\varepsilon_{[n]}-\varepsilon_t)/T].$$

Here $[n]$ stands for the complete set of quantum numbers for an "equilibrium" level. For "nonequilibrium" levels of a trap (if it has such levels) and for states in the continuous spectrum, the time that it takes an electron to leave for lower equilibrium levels exceeds the time that it takes to leave the well. We may assume, therefore, that if an electron finds itself in such states, it leaves the well. Distinguishing between equilibrium and nonequilibrium states and using the principle of detailed balance for the latter make it possible to write τ_e in the form

$$\tau_e^{-1}=g^{-1}\sum_{[n]}\tau_{[n]}^{-1}\exp[-(\varepsilon_{[n]}-\varepsilon_t)/T]$$
$$+g^{-1}\sum_{\{n\}}\tau_{l\{n\}}^{-1}\exp[-(\varepsilon_{\{n\}}-\varepsilon_t)/T]. \quad (21)$$

where $\{n\}$ stands for the set of quantum numbers characterizing a nonequilibrium level, and $\tau_{l\{n\}}$ is the time that it takes an electron to move from level $\{n\}$ to lower equilibrium levels. The main process determining $\tau_{l\{n\}}$, the descent of carriers along the energy axis in the course of phonon emission, is the same as in the capture of carriers by attractive centers.[15] However, the characteristic time τ_l of this process may differ drastically both from the similar quantity for attractive centers and from trap to trap and depends on the positioned of the upper levels [Eqs. (16) and (17)] in relation to the top of the barrier and on the shape of the barrier near the top. For traps whose nonequilibrium levels are distributed fairly densely up to the upper equilibrium level, τ_l must be of the order of the phonon emission time, that is, roughly 10^{-12}–10^{-13} s at ~ 300 K. But if the distance from the lower nonequlibrium level to the upper equilibrium one exceeds the optical-phonon energy, τ_l is determined by a multiphonon process and may be several orders of magnitude longer.

The capture time τ_c of a trap for the Fermi filling of levels is linked with τ_e via the relation

$$\tau_c/\tau_e=g^{-1}\exp[(\varepsilon_t-\varepsilon_F)/T].$$

This yields

$$\tau_c^{-1}=\sum_{[n]}\tau_{[n]}^{-1}\exp[-(\varepsilon_{[n]}-\varepsilon_F)/T]$$
$$+\sum_{\{n\}}\tau_{l\{n\}}^{-1}\exp[-(\varepsilon_{\{n\}}-\varepsilon_F)/T]. \quad (22)$$

In deriving (19)–(22) we ignored indirect tunneling of electrons to equilibrium levels. This process is less probable than direct tunneling since the respective expressions for τ_e^{-1} and τ_c^{-1} contain a factor much smaller than $T_{n,L}^{-1}$. The times $T_{n,L}$ are the shortest in this case. At $\varepsilon_1=0.025$ eV we have

$$T_{n,L}=2\cdot 10^{-14}[(n+L+1)^3/Z_{\varepsilon'}^2]\text{ s}.$$

Allowing for $U_2(\mathbf{r})$ [Eqs. (13) and (13a)] lifts the degeneracy of the levels in M and eliminates the quantum numbers n and L. True, for $r<y$ this correction is moderate and has little effect on the values of $\varepsilon_{\hat{n}}$. For this reason it remains expedient to classify the levels using the numbers n, L, and M. The main effect of $U_2(\mathbf{r})$ consists in a considerable increase in both τ_c and τ_e, since even a small increase in the barrier height drives the tunneling probability down. For instance, even for $\theta \ll 1$, when

$$U_2(r)\approx\varepsilon_1 K\theta^2[1-y^2/(r+y)^2],$$

the barrier's transparency, already small in the x direction, rapidly decreases as θ grows. Hence, on the right-hand side of Eq. (19) there appears an additional large factor raised to the $(M+1)$ st power. True, further conclusions are independent of this growth of the pre-exponential factor in τ_c and τ_e, and the factor is ignored below. Finally we note that if in Si[100] the free electrons and the resonance level belong to different valleys, we must include in $\tau_{[n]}^{-1}$ the probability of an intervalley transition.

4. THE STATISTICS OF OBSERVED SLOW TRAPS

To determine the optimal conditions of observation of RTS in submicron MOS structures and the expected RTS parameters, we being by estimating the entropy of a large-scale repulsive fluctuation needed for observing a single slow trap in the structure of such dimensions. Slow traps constitute a vivid example of hybrid localization of carriers in a random Coulomb field.[6] To estimate the probability of their existence, we employ the two-parameter density of states introduced in Ref. 6:

$$N(\varepsilon_t,\varepsilon')\approx\rho_{\text{quantum}}(\varepsilon')P(\varepsilon_t-\varepsilon'),$$

where $P(\varepsilon_t-\varepsilon')$ is the probability density of a classical shift of quantum levels by $\varepsilon_t-\varepsilon'$. In our case $P(\varepsilon_t-\varepsilon')$ can be estimated as follows:[5,13]

$$P(\varepsilon_t-\varepsilon')\sim\exp(-S_{opt})/4\varepsilon_1[\pi\Sigma\ln(d/a_{\varepsilon'})]^{1/2},$$

where S_{opt} is the entropy of the fluctuation that guarantees the required shift. This implies that for given $\overline{Q}$ and T the desired average number $\bar{p}$ of the slow traps in a MOS structure of area A is given, in order of magnitude, by the following formula:

$$\bar{p}\sim A\Delta\varepsilon_t\Delta\varepsilon' N(\varepsilon_t,\varepsilon')$$
$$\sim 50T^2A_{\text{quantum}}(\varepsilon')\exp(-S_{opt})/4\varepsilon_1[\pi\Sigma\ln(d/a_{\varepsilon'})]^{1/2}, \quad (23)$$

where $\Delta\varepsilon_t$ and $\Delta\varepsilon'$ are the ranges of admissible variations in ε_t and ε', with $\Delta\varepsilon_t\Delta\varepsilon'\sim 50T^2$, since RTS observations τ_c and τ_e do not differ very greatly, that is, $\varepsilon_F-\varepsilon_t\leqslant(3\text{–}4)T$, and vary from 10^{-3} s to 10^2 s (i.e., the range of variations of the activation energies in these times is less than $T\ln 10^5$; see Ref. 10). Formula (23) provides a crude estimate for $\bar{p}$, but, because of the logarithmic dependence of S_{opt} on the quantities in (23), is quite sufficient for determining the value of S_{opt} needed for these traps to appear with a required rate. For the same reason the suggested estimate of S_{opt} is also acceptable (at $\varepsilon_1 = 0.025$ eV, $\tilde{\Sigma} \sim 1$, and $A \sim 10^{-8}$–10^{-9} cm^2).

Further theoretical analysis and the data of Ref. 10 show that at roughly 300 K the important values of ε' are approximately 0.5 eV and for such ε' and $\tilde{\Sigma} \sim 1$ we have $\rho_{\text{quantum}}(\varepsilon') \sim 10^9$–$10^{10}$ cm^{-2} eV^{-1} (Sec. 2), so we find that at ~ 300 K

$$\rho_{\text{quantum}}(\varepsilon')(T/\varepsilon_1)^2 \sim 10^9-10^{10}\ \text{cm}^{-2}\,\text{eV}^{-1}.$$

True, this value does not vary very strongly with decreasing T since the important values of ε' decrease (see below) and the corresponding value of $\rho_{\text{quantum}}(\varepsilon')$ grows. For instance, the data of Ref. 9 imply that at ~ 30 K we have $\varepsilon' \sim \varepsilon_1$, and since $\rho_{\text{quantum}}(\varepsilon') \sim 10^{12}$–$10^{13}$ cm^{-2} eV^{-1} for such ε' and $\tilde{\Sigma} \sim 1$, we have

$$\rho_{\text{quantum}}(\varepsilon')(T/\varepsilon_1)^2 \sim 10^{10}-10^{11}\ \text{cm}^{-2}\,\text{eV}^{-1}.$$

Together with (23) this leads to the estimate $S_{opt} \sim \ln(10/\bar{p})$. Apparently, for radiating traps $\bar{p} \sim 0.1$, that is, $S_{opt} \sim 5$. The given fairly small value of $\bar{p}$ corresponds to the face[10] that for practically every sample there exists a temperature range (quite narrow, $\Delta T/T \sim 0.1$) in which the RTS is observed. The observation of traps at smaller values of $\bar{p}$ (larger values of S_{opt}) is unlikely, and the study of traps at $\bar{p} \sim 1$ is hindered by the pile-up of RTS generated by different traps whose charge exchange times differ little.

We begin the search for the most probable traps having very low capture cross sections σ by considering the case of maximum temperatures. The extremely strong barriers needed here exist with a noticeable probability only for minimum screening, and carriers are captured either owing to thermal activation or to tunneling to the upper-most, nonequilibrium, levels, that is, ε_T, the optimum capture energy, is close to the top of the barrier. Hence in (14) we put $b_1(\tilde{B}) \approx 1/2$. This yields

$$\tilde{\varepsilon}_T/4\tilde{B} = \ln(\eta\tilde{d}\tilde{B}/Z_{\varepsilon'}) - 1/2, \tag{24}$$

and using (22), we can link the barrier height to τ_c:

$$\tilde{\varepsilon}_T - \tilde{\varepsilon}_F = (T/\varepsilon_1)\ln(\tau_c/\tau_l), \quad \text{i.e.,} \quad \tau_c = \tau_l \exp[(\varepsilon_T - \varepsilon_F)/T]. \tag{25}$$

Here we have allowed for the fact that the principal term in (22) is the sum over $\{n\}$, which can be estimated as $\tau_l^{-1}\exp[-(\varepsilon_T - \varepsilon_F)/T]$, where τ_l is most likely to lie in the range 10^{-10} to 10^{-12} s (Sec. 3).

Augmenting the system of equations (18) (with $K = \tilde{B}$), (24), and (25) with the expression linking ε' and $Z_{\varepsilon'}$, which we write in the form

$$\tilde{\varepsilon}' = cZ_{\varepsilon'}^2, \tag{26}$$

where c is a factor of order unity,[6] we can use the new system to find $\tilde{B}$ and then, via (15), the entropy of the optimal fluctuation corresponding to the given $\tilde{Q}$ (or ε_F), T, Σ, and τ_c or, as is done below, to find the relation between ε_T and T for a given value of S_{opt}.

Introducing the notation $t = (\tilde{\Sigma}S_{opt})^{1/2}/\tilde{B}$ into Eqs. (18) and (24)–(26), we arrive at an equation for t for traps with ε_t,

$$t^2 = \ln\{\eta\tilde{d}[S_{opt}\tilde{\Sigma}c/t[t(T/\varepsilon_1)\ln(\tau_c/\tau_1) + (2+4\ln 2)(S_{opt}\tilde{\Sigma})^{1/2}]]^{-1/2}\} \tag{27}$$

and an equation linking T with ε_F,

$$(T/\varepsilon_1)\ln(\tau_c/\tau_l) + \tilde{\varepsilon}_F = 2(S_{opt}\tilde{\Sigma})^{1/2}(2t - 1/t). \tag{28}$$

The dependence of t on the structure and trap parameters in the right-hand side of (27) is extremely weak. Apparently, $-1.3 \leqslant t \leqslant 1.5$ holds over the entire range of real values of the parameters [e.g., for $\tilde{d} = 15$, $\eta = 2$, $\tilde{\Sigma} = 1$, $S_{opt} = 5$, $c = 1$, $T = \varepsilon_1 = 0.025$ eV, and $\ln(\tau_c/\tau_l) = 25$ value $t \approx 1.45$. This allows us to replace $2t - t^{-1}$ by 2 in (28) with rather good accuracy. If, in addition, we express ε_F in terms of $\bar{Q}_d$ via (3), we find that Eq. (28) yields

$$T\ln[\tau_c\tau_l^{-1}\bar{Q}_d^{opt}(\pi a_1^2)^{3/4}eE_s2^{-3}T^{-1}\tilde{\Sigma}^{-3/4}] \approx 2(S_{opt})^{1/2}\Delta, \tag{29}$$

which shows that for large values of T the free carrier density optimal for studying slow traps, $\bar{Q}_d^{opt}$, varies according to the following law: $\bar{Q}_d^{opt} \propto \exp(2S_{opt}^{1/2}\Delta/T)$. Finding σ from the formula $\tau_c^{-1} = \sigma v_T\bar{Q}_d eE_s/T$, where v_T is the thermal velocity of the carriers,[10] we see from Eq. (29) that $\sigma = \sigma_0\exp(-\Delta E_B/T)$ holds for optimal traps, with $\Delta E_B \approx 2S_{opt}^{1/2}\Delta$ and $\sigma_0 = (\pi a_1^2)^{3/2}/8\tilde{\Sigma}^{3/4}\tau v_T$. From (29) it follows that, for instance, at $T \sim 320$ K and $\bar{Q}_d^{opt} \sim 10^{10}$ cm^{-2} observation of RTS with $\tau_c \sim 1$ in the majority of samples is possible when $\Delta \approx 0.08$ eV (it is assumed that $S_{opt} \approx 5$, $\tau_l \approx 10^{-10}$ s and $E_s = 10^4$ V/cm), that is, $\Sigma \approx 5 \times 10^{12}$ cm,$^{-2}$ and then the following values are optimal: $\varepsilon' \approx 0.8$ eV, $\varepsilon_F \approx -0.24$ eV, $\Delta E_B \approx 0.35$ eV, and $\sigma_0 \approx 10^{-17}$ cm^2. For comparison we note that the σ vs T dependence obtained for a large number of traps studied at temperatures ranging from 250 to 350 K and times ranging from 10^{-3} to 10^2 s has revealed that σ_0 and ΔE_B lie in the 10^{-14}–10^{-19} cm^2 and 0.2–0.65 eV ranges, respectively.[10] For the results listed in Ref. 9 an estimate yields $\Delta \approx 0.04$ eV. This somewhat overestimates Δ (and, hence, Σ and ε', too), since in deriving Eq. (29) S was determined via (15), which overestimates the entropy of an optimal fluctuation (see Sec. 2) and no allowance was made for an increase in probability of slow-trap formation along the periphery of the inversion channel. In narrow-channel MOSFETs used in RTS studies, peripheral effects are essential.

Equation (29) implies that improving the quality of MOS structures (by lowering Σ) leads to a smaller value of $\bar{Q}_d^{opt}$ at a given temperature. Here the RTS amplitude may prove to lie below the signal detection threshold, that is, $\bar{Q}_d^{opt} < \bar{Q}_d^{thr}$, where $\bar{Q}_d^{thr}$ is the threshold number density of free electrons. The threshold and Σ determine the maximum temperature (for a given τ_c) of stable RTS observation in these structures [Eq. (29) with $\bar{Q}_d^{opt}$ replaced by $\bar{Q}_d^{thr}$]. In a large sample, of course, there may be specimens with very strong barriers that enable observing RTSs at higher temperatures.

At low temperatures Eq. (29) becomes invalid because of the increasing optimal density of carriers, $\bar{Q}_{opt}$, and the strengthening of electron screening, which was not allowed for in Eq. (29). The expression for $\tilde{Q}^{opt} \equiv \pi\bar{Q}^{opt}a_1^2$ can be derived from Eqs. (28) and (2)[2)] (since here $\bar{Q} \approx \bar{Q}_l \gg \bar{Q}_d$):

$$T\ln(\tau_c/\tau_l) \approx 2\Delta[S_{opt}^{1/2} + \ln(2^{3/4}\tilde{\Sigma}^{5/8}/\pi^{1/4}\tilde{Q}^{opt})]. \tag{30}$$

Electron screening lowers the height of a barrier with radius $l_0 \gg \eta d$ by $4\pi e^2 \overline{Q} d/\varepsilon_i$. If instead of $\widetilde{Q}^{\rm opt}$ we substitute into (30) the value at which this lowering is Δ, we find that the temperature below which electron screening noticeably lowers the barrier is

$$T_1 = 2\Delta[S_{opt}^{1/2} + \ln(2^{11/4}\widetilde{\Sigma}^{1/8}\widetilde{d}\varkappa/\pi^{1/4}\varepsilon_i)]/\ln(\tau_c/\tau_l). \quad (31)$$

At $S_{\rm opt} = 5$, $\tilde{d} = 15$, and $\tau_c/\tau_l = 10^{10}$ we have $T_1 \approx 0.3\Delta$, and with $\widetilde{\Sigma} = 2$ ($\Sigma \approx 5\times10^{12}$ cm^{-2}), when $\Delta \approx 0.08$ eV, we have $T_1 = 0.024$ eV. For $T \leqslant T_1$ the temperature dependence of $\overline{Q}_d^{\rm opt}$ becomes much weaker than Eq. (29) would imply.

As electron screening grows, the optimal fluctuation becomes smaller. For $l_0 < \eta d$ we must use Eq. (11) to determine ε_T instead of Eq. (14). Here, too, at maximum temperatures (see below) ε_T is close to the top of the barrier, that is, $\tilde{b}(l_0) \approx 1/2$ [see Eq. (11)]. Hence, for $l_0 < \eta d$, in the system of equations used for $l_0 \gg d$ Eqs. (25) and (26) do not change, $K = \widetilde{Q}\tilde{l}_0$ in Eq. (18), and Eq. (15) must be replaced with (10) and Eq. (24) with

$$\tilde{\varepsilon}_T/4\widetilde{Q}\tilde{l}_0 = \ln(\widetilde{Q}\tilde{l}_0^2/Z_{\varepsilon'}) - 5/2. \quad (32)$$

If we allow for (2), this system reduces to a cumbersome transcendental equation for $\widetilde{Q}^{\rm opt}(T)$. Here we discuss only some of its implications.

In deriving Eq. (11) and subsequent formulas a large parameter was employed, $\tilde{l}_0 \gg 1$. Numerically this condition proves to be more stringent: $\tilde{l}_0 > \tilde{l}_{\rm cr} \equiv S_{\rm opt}^{1/2}\pi^{1/4}2^{-3/4}\widetilde{\Sigma}^{-1/8} \exp u$, where u is the solution to the equation

$$u - \ln[u + S_{opt}^{1/2}(1/2 + \ln 2)]/2 = \ln[2^{7/4}e^{5/2}/\widetilde{\Sigma}^{1/8}\pi^{1/4}S_{opt}c^{1/2}].$$

The quantity u depends weakly on all the parameters. At $\widetilde{\Sigma} = 1$, $c = 1$, and $S_{\rm opt} = 5$ we have $u \approx 2.65$, and $\tilde{l}_{\rm cr} \approx 25$. The conditions $\tilde{l}_{\rm cr} < \tilde{l}_0 < \eta\tilde{d}$ can be met only in MOS structures with $d = 1000$–1200 Å in a moderate temperature interval

$$u < \ln(\tau_c/\tau_l)T/2\Delta < u + S_{opt}^{1/2}\ln(\eta\tilde{d}/\tilde{l}_{cr}),$$

in which

$$\widetilde{Q}^{opt} = 2^{4/4}\widetilde{\Sigma}^{5/8}\pi^{-1/4}\exp\{-[u + (\ln(\tau_c/\tau_l)T/2\Delta - u)/S_{opt}^{1/2}]\}, \quad (33)$$

and

$$\overline{Q}_d \propto \exp[-2u(1 - S_{opt}^{-1/2})\Delta/T].$$

When $d \sim 400$–600 Å, the temperature dependence of $\widetilde{Q}^{\rm opt}(T)$ for $l_0 < \eta d$ can be found only numerically. The temperature of the transition of optimal fluctuations with $l_0 < \eta d$ for such values of d can be estimated as $T_2 \approx 2u\Delta/\ln(\tau_c/\tau_l)$. At $\Delta = 0.08$ eV and $\tau_c/\tau_l = 10^{10}$ we obtain $T_2 \approx 0.022$ eV.

Qualitatively, at lower temperatures the average statistical properties of a trap generating an RTS are as follows. If we wish to observe an RTS with given τ_c and τ_e as T decreases, we must increase $\overline{Q}$, so that the enhanced screening lowers the barrier and, hence, does not allow a sharp increase in τ_c and τ_e. In the process one should also observe a tendency for the transition of ε_T to ever lower lying resonance levels, that is, as T is lowered, an ever growing number of traps should be observed for which the factor $\exp(\Delta E_B/T)$ in the expression for σ is fairly moderate but the factor σ_0 is small. True, traps that capture carriers through low-lying resonance states should also often be encountered for $T \gg T_2$. We have assumed that ε_T coincides with the top of the barrier, which followed from the requirement that $\Lambda(\varepsilon) + \varepsilon/T$ grow with ε if one ignores the energy dependence of the pre-exponential factor in the expression for τ_c. The requirement, as can be verified, has the form $\varepsilon_1/T < (\frac{8}{3})^{1/2}\pi Z_{\varepsilon'}/K^{3/2}$ and is met in the situations just discussed. Actually, for lower quasidiscrete levels, $T_{n,L}$ is of order 10^{-14} s or less, while the real values of τ_l are on the order of 10^{-10} s or even more (see Sec. 3). This leads to a decrease in ε_T even for high temperatures. Finally, for the lowest temperatures and the respective large values of $\overline{Q}^{\rm opt}$, when the Fermi level lies in the allowed band and direct tunneling of electrons to level ε_t becomes possible, traps should be observed that have no thermal-activation dependence in σ ($\Delta E_B = 0$).

Thus, the foregoing implies that (1) among silicon MOSFETs of submicron dimensions when the number density of the charged centers built into the insulator is of order 10^{12} cm,$^{-2}$ many may contain a single fluctuation surface trap with times τ_c and τ_e in the range 10^{-3} to 10^2; (2) the temperature dependence of σ for such charges is of the thermal-activation type observed in RTS experiments, and the optimal values of σ_0 and ΔE_B we have found lie within the range of values measured for different traps (the reason for such a broad spread in values of σ_0 and ΔE_B is also discussed); and (3) the tendency, predicted by the fluctuation theory, for the conduction of the inversion channel to change, whereby an RTS is observed, and the properties of the observed traps to change under temperature variations resemble in many respects those observed in experiments.

Now, knowing the structure and properties of fluctuation surface traps, we can determine the main parameters characterizing an RTS generated by a single trap of this kind. We will also determine what information can be extracted from studying such traps and compare their parameters with the parameters of RTS observed in mesoscopic silicon MOSFETs.

5. THE RTS AMPLITUDE

Since a slow trap remains in filled and empty states for a very long time, it is possible to isolate in the noise of a varying nature of considerably higher frequencies those fairly small jumps in the resistance of the inversion channel, $\Delta R/R$ (with R the resistance of channel and ΔR its variation brought on by the capture of an electron), that are generated by the change in the charge state of the trap. The mechanism of these jumps differs for low and high electron number densities.

We start with the case where $\overline{Q}_d < \overline{Q}_l$. The conductivity of the channel is of the activation type [see Eq. (3)] and external charges are screened by electrons localized at fast surface traps and by the gate electrode. When an electron is captured by a slow trap, its potential raises the fast-trap levels by $u(l)$, and in the course of time intervals long compared to the charge-exchange time of the traps an additional electron density is generated equal, on the average, to

$$\delta\bar{Q}_t(l)=\int_0^\infty d\varepsilon'\rho_b(\varepsilon',l)\{[1+\exp[(U(l)+u(l)-\varepsilon'-\varepsilon_F)/T]]^{-1}$$
$$-[1+\exp[(U(l)-\varepsilon'-\varepsilon_F)/T]]^{-1}\}, \qquad (34)$$

where $U(l)$ is the height of the slow-trap barrier, and $\rho_b(\varepsilon',l)$ the density of states with a binding energy ε' at point l. Here

$$\rho_b(\varepsilon'\,l)=(\bar{Q}_l/2\Delta)\exp[-(\varepsilon_F+\varepsilon')/2\Delta],$$

at $l\sim l_0$ and $\varepsilon'\sim-\varepsilon_F$ and

$$\rho_b(\varepsilon',l)\approx\rho_{\text{quantum}}(\varepsilon')$$

for $l\ll l_0$. From (1) with $u(l)<T<2\Delta$ we have

$$\delta\bar{Q}_t(l)=-u(l)\rho_b[U(l)-\varepsilon_F,l].$$

We see that although the shift $u(l)$ is at its greatest when $l\ll l_0$, with $u(l)\approx e^2/\varkappa l$, the quantity $\delta\bar{Q}_l(l)$ is small because $\rho_b[U(l)-\varepsilon_F,l]$ is exponentially small when $U(l)>2\Delta$. Hence, for $l\gg l_0$ the main screening charge occupies the region with $U(l)\leqslant 2\Delta$ and $\delta\bar{Q}_l(1)\approx-u(l)\bar{Q}_l/2\Delta$. If the length of linear electron screening, $\varkappa\Delta/\pi e^2\bar{Q}_l$, and l_0 are small compared to d, an electron from a slow trap screens fast traps, where the number of electrons decreases by 1. In times τ_c and τ_e the distribution of the electrons over the fast traps reaches equilibrium and outside the barrier, when a slow trap captures an electron, the distance from the Fermi level to the boundary of the allowed band changes, on the average over the sample, by $\delta\varepsilon_F=-2\Delta/A\bar{Q}_l$ [see Eq. (2)], while the relative changes in $\bar{Q}_d$ [see Eq. (3)] and $\Delta R/R$ are equal:

$$\Delta R/R=-\Delta\bar{Q}_d/\bar{Q}_d=2\Delta/T\bar{Q}_lA. \qquad (35)$$

The variation in mobility caused by charge exchange in the trap was ignored here because its contribution $\sim(\Sigma A)^{-1}$ is negligible.

From (35) it follows that $\Delta R/R\propto\bar{Q}^{-1}$. Using Eqs. (2) and (3) and expressing $\bar{Q}_l$ in terms of $\bar{Q}_d$, we find that $\Delta R/R\propto\bar{Q}_d^{-T/2\Delta}$.

For $\bar{Q}\ll\varkappa\Delta/\pi e^2d$ the relative RTS amplitude decreases with $\bar{Q}$, contrary to the case discussed above. The reason is weak electron screening, with the result that a slow trap capturing an electron leads only to a small probability $p_b\ll 1$ of an electron leaving fast traps. We can find this probability by employing the above expression $\delta\bar{Q}_l(l)\approx-u(l)\bar{Q}_l/2\Delta$ and the fact that under weak electron screening (for $p_b\ll 1$) $u(l)\approx e^2/\varkappa l$ for $l<\eta d$ and $u(l)\approx e^2\varepsilon_sd^2/\varepsilon_i^2l^3$ for $l\gg 2\varkappa d/\varepsilon_i$. The result is

$$p_b=2\pi\int_0^\infty dll|\delta Q(l)|=\frac{\pi\bar{Q}_l}{\Delta}\int_0^\infty dllu(l)=\frac{\pi e^2\bar{Q}_ld}{\Delta}\times\begin{cases}2/\varepsilon_i-l_0/\varkappa d & \text{if} \quad l_0\ll d,\\ 2\varepsilon_sd/\varepsilon_il_0^2 & \text{if} \quad l_0\gg\varepsilon_sd/\varepsilon_i.\end{cases}$$

The average deviation of the Fermi level outside the barrier, $\delta\varepsilon_F$, decreases in proportion to p_b. As a result, formula (35) assumes the form

$$\frac{\Delta R}{R}=\frac{4\pi e^2d}{\varepsilon_iTA}\begin{cases}1-\varepsilon_il_0/\varkappa d & \text{if} \quad l_0\ll d,\\ \varepsilon_sd/\varepsilon_il_0 & \text{if} \quad l_0\gg\varepsilon_sd/\varepsilon_i.\end{cases} \qquad (36)$$

Hence, here $\Delta R/R$ depends on the radius of the barrier in the following way: it decreases as l_0 grows, and since l_0 grows as $\bar{Q}$ decreases, $\Delta R/R$ decreases as well. In addition, the derivation of (36) implies that for traps with $l_0<d$ the relative RTS amplitude reaches its maximum at $\bar{Q}_l\sim\varepsilon_i\Delta/2\pi e^2d$ and assumes the following value: $(\Delta R/R)_{\max}\approx 4\pi e^2d/\varepsilon_iTA$. At $d=4\times10^{-6}$ cm, $T=1/40$ eV, and $A=4\times10^{-9}$ cm^2 we have $(\Delta R/R)_{\max}\approx 2\%$. Formulas (35) and (36) give the RTS amplitude averaged over different samples. The amplitude differs from sample to sample because it is determined not by $\delta\varepsilon_F$ averaged over the entire sample outside the barrier but by a similar quantity averaged exclusively over the free-electron flow paths that pass only above the potential wells of the random potential pattern. Hence, in samples in which the flow paths pass next to a barrier, $\Delta R/R$ is greater than the values that follow from (35) and (36). On the other hand, in samples where the paths are far from a barrier, $\Delta R/R$ is smaller than these values. This agrees with the experimental data listed in Refs. 9, 10, and 16.

Now we turn to the case $\bar{Q}_d>\bar{Q}_l$. For real Si:SiO$_2$ structures such a situation is possible if $\bar{Q}>\Sigma$, that is, $\bar{Q}\geqslant(2-4)\times10^{12}$ cm.$^{-2}$ Then the radius of a barrier in a trap is fairly moderate [since $l_0\sim\bar{Q}^{-1/2}$; see condition (5)], and the trap is screened primarily by free charge carriers. The optimal temperature for observing traps with charge-exchange times of the order of $10^{-2}-1$ s for such $\bar{Q}$ is moderate (about 50 K or even lower; see Sec. 4 and Ref. 9), and the free electrons form a $2D$ Fermi liquid whose metallic conduction determines their excitation within a narrow energy band ($\sim T$) near ε_F. When a slow trap captures an electron, the number of free electrons in the sample decreases by 1. The number of conduction electrons does not change in the process since the $2D$ density of states is a constant. Resistance increases because the Fermi wave vector k_F diminishes and because the scattering on the trap barrier intensifies, with the barrier becoming higher owing to electron capture. The jump in resistance, ΔR_c, caused by the decrease in the number of free electrons by 1 can be interpreted as a variation in R brought on by the decrease in the gate potential V_g by the potential of the capacitor formed by the sample with charge e, that is $4\pi ed/\varepsilon_iA$:

$$\Delta R_c/R=(\partial\ln R/\partial V_g)4\pi ed/\varepsilon_iA. \qquad (37)$$

Here we have also allowed for variations in the contribution from different scatterers and the effect of weak carrier localization, which may manifest itself in such samples.

Among free-electron scatterers the slow trap is the most effective. Its scattering cross section is classical and roughly equal to $2l_0$. Since $\tilde{Q}_d\tilde{l}_0^2=Z$, where eZ is the charge of the centers incorporated in the trap, Z increases by 1 when an electron is captured. The respective increase in the scattering cross section can be estimated as $\sim 2\delta\tilde{l}_0=(\tilde{Q}_d\tilde{l}_0)^{-1}$, and the resulting increase in the effective inverse time of scattering at $\sim\hbar k_F(mA\tilde{Q}_d\tilde{l}_0)^{-1}$. This yields

$$\Delta R_l/R=\gamma_t(\hbar\mu/eA)(2/g_v\tilde{Q}_d)^{1/2}\tilde{l}_0^{-1}, \qquad (38)$$

where μ is the electron mobility, g_v the degree of valley degeneracy of band states, and γ_t a factor of order unity.

The sum of (37) and (38) gives the relative RTS amplitude. Apparently, in the experiments conducted so far the first term, (37), dominates. But the final conclusion depends on the values of μ and γ_t.

At liquid helium temperatures the RTS amplitude may be determined by universal conductance fluctuations (UCF). Earlier[17] it was assumed, on the basis of the data listed in Refs. 18 and 19, that the RTS are caused by UCF at large displacements ($k_F\delta l \gg 1$) of a quasipoint scatterer. In the given model the sources of UCF are variations in ε_F (see Ref. 20) and small deformations of the scatterer, whose size is much smaller than the mean free path of the electrons, l_{sc}. The variations in ε_F caused by the decrease in the number of free electrons by 1 are weak and introduce into the UCF amplitude an additional small factor of the order of $4\pi ed/\varepsilon_i A V_{gc}$, where V_{gc} is the halfwidth of the conductance correlation function caused by variations in V_g. For $V_{gc} \approx 0.2B$, $d \approx 6\times 10^{-6}$ cm (see Ref. 17), and $A = 10^{-9}$ cm^{-2} this factor is roughly 0.015. The contribution introduced by the change in the barrier is greater here. On the basis of ideas about the interference between classical Feynman paths and owing to the fact that in a 2*D* system every path passes through a finite fraction of the total number of scatterers,[19] we can assume that changes in the barrier generate UCF whose amplitude contains an additional factor of order $k_F\delta l_0(l_0/l_{sc})$, which reflects the presence of a correlated small shift ($k_F\delta l_0 \ll 1$) of a large number ($\sim l_0/l_{sc}$) of scatterers. Since we have $l_0\delta l_0 \approx (2\pi\bar{Q}_d)^{-1}$, this factor is of order $(k_F l_{sc})^{-1} \ll 1$, that is, the amplitude of the UCF generated by changes in the charge of a fluctuation trap is parametrically small and decreases as a function of $\bar{Q}_d$.

The UCF were employed to explain the RTS amplitude in view of the variation of the latter from trap to trap.[17] For fluctuation traps there may also be other reasons for such variations. The regular variations in the RTS amplitude [see Eqs. (35)–(38)] can be separated in experiments from UCF by isolating the random oscillations in amplitude caused by variations in V_g. The period of these oscillations ($\sim V_{gc}$) for $\bar{Q}_d > \bar{Q}_l$ must be short compared to the characteristic scale of the regular variations in the amplitude. In identifying the mechanism that determines the RTS amplitude it is also possible to use variants of th 1ethod of determining the position of a slow trap by the length of the sample.[21]

6. FIELD AND TEMPERATURE CURVES FOR CAPTURE AND EJECTION TIMES τ_c and τ_e

The source and sink potentials of a MOSFET, if one ignores their difference, specify the universal position of the Fermi level in the inversion channel. When V_g varies by ΔV_g in the direction of inversion, the Fermi level in a metal drops by $e\Delta V_g$ in relation to the position of the Fermi level in the channel, which shifts by a quantity $\Delta\varepsilon_F$ determined by the following formula:

$$\Delta V_g = (4\pi ed/\varepsilon_i)(\Delta\bar{Q} + N_a\Delta w) = (4\pi ed/\varepsilon_i)(\partial\bar{Q}/\partial\varepsilon_F + e^{-2}C_s^{-1})\Delta\varepsilon_F,$$

where Δw is the variation in the thickness of SCR, and $C_s = \varepsilon_s/4\pi w$ the specific capacitance of SCR. This leads to the following expressions for the variations in ε_F and $\bar{Q}$:

$$\partial\varepsilon_F/\partial V_g = (\varepsilon_i/4\pi ed)(\partial\bar{Q}/\partial\varepsilon_F + e^{-2}C_s^{-1})^{-1}, \quad (39)$$

$$\partial\bar{Q}/\partial V_g = (\varepsilon_i/4\pi ed)(\partial\bar{Q}/\partial\varepsilon_F + e^{-2}C_s^{-1})^{-1}\partial\bar{Q}/\partial\varepsilon_F. \quad (40)$$

The increase in ε with V_g caused τ_c to lower [see Eq. (22)]. The effect of changes in $\bar{Q}$ on τ_c and τ_e is more complex. Let us study it for the situation in which $\bar{Q}_l > \bar{Q}_d$. Two cases are of interest, $l_0 < d$ and $l_0 \gg d$.

For $l_0 < d$, as $\bar{Q}$ increases by $\Delta Q\tilde{Q}$, the excess charge density inside the barrier becomes equal to $\xi(l) - \bar{Q} - \Delta\bar{Q}$, as a result of which l_0 decreases by $\Delta l_0 = l_0\Delta\bar{Q}/2\bar{Q}$ [see condition (5)]. This and the variation in the average field at the semiconductor surface equal to $2\pi e\Delta\bar{Q}/\varepsilon_s$ cause the energy to acquire an additional term $\Delta U(\mathbf{r})$. When $r \ll l_0$, the additional energy term is given by the following formula:

$$\Delta U(\mathbf{r}) = -(\pi e^2\Delta\bar{Q}/\varkappa)[l_0 - 2(1+\varkappa/\varepsilon_s)r\cos\theta].$$

This term shifts the energies $\varepsilon_{\hat{n}}$ by

$$\Delta\varepsilon_{\hat{n}} = -(\pi e^2\Delta\bar{Q}/\varkappa)[l_0 - 2(1+\varkappa/\varepsilon_s)r_{\hat{n}}], \quad (41)$$

where $r_{\hat{n}} = \int d\mathbf{r}\Psi_{\hat{n}}^2(\mathbf{r})r\cos\theta$, with $\Psi_{\hat{n}}(\mathbf{r})$ the wave function of state $\hat{n}$. The respective tunneling exponents change by $\Delta\Lambda_{\hat{n}}$:

$$\Delta\Lambda_{\hat{n}} = 2\pi e^2\Delta\bar{Q}(\varkappa^{-1} + \varepsilon_s^{-1})(2m/\hbar^2)^{1/2} \times \int_{x_1}^{x_2} dx\,(x - r_{\hat{n}})[\hat{U}(x) - \varepsilon_{\hat{n}}]^{-1/2}. \quad (42)$$

Employing Eqs. (19)–(22) and (39)–(42), we get

$$\frac{T}{e}\frac{\partial\ln(\tau_c/\tau_e)}{\partial V_g} = \frac{\partial(\varepsilon_t - \varepsilon_F)}{e\partial V_g} = -\frac{\varepsilon_i}{\varkappa}\frac{[l_0 - 2(1+\varkappa/\varepsilon_s)r_{0,1,0} + 2a_t\Sigma^{1/2}/\bar{Q}]}{4d}, \quad (43)$$

$$(T/e)\partial\ln\tau_e/\partial V_g \approx (\varepsilon_i/\varkappa + \varepsilon_i/\varepsilon_s)[(\gamma y - r_{0,1,0})/2d]. \quad (44)$$

In deriving these two formulas we have allowed for (2) and assumed that traps with $l_0 < d$ exist only if $e^2C_s(\partial\bar{Q}/\partial\varepsilon_F) \gg 1$ holds, that is, if $\bar{Q}$ is not very low. Also, when we derived (44), the integral in (42) was considered for the optimal-capture level, $\varepsilon_{\hat{n}} \approx \varepsilon_T$. The factor $x - r_{\hat{n}}$ in the integrand was estimated at $\gamma y - r_{\hat{n}}$, where y stands for the coordinate of the top of the barrier, and γ is a factor equal to unity for small Λ and slowly increasing with Λ. The remaining integral was estimated at $\hbar/(2m)^{1/2}T$, since $\varepsilon = \varepsilon_T$ corresponds to the minimum of $\Lambda(\varepsilon) + \varepsilon/T$ (see Sec. 4). From Eqs. (43) and (44) it follows that τ_e increases with V_g but τ_c decreases. For large $\bar{Q}$ the rates at which τ_c and τ_e change are of the same order of magnitude, while for small $\bar{Q}$ the rate of variation of τ_c is much higher than that of τ_e since in the latter case we have $l_0 \gg y$. If $\partial\varepsilon_F/\partial V_g$ is determined together with $\partial\ln\varepsilon_F/\partial V_g$ and $\partial\ln\tau_e/\partial V_g$ (see, e.g., Refs. 1 and 2), Eq. (43) can be used to find l_0.

For traps with $l_0 \gg d$, which exist when $\bar{Q}$ is small, the additional energy term $\Delta U(\mathbf{r})$ with $r \ll l_0$ is independent of $\mathbf{r}$ to zeroth order in the parameter d/l_0:

$$\Delta U(\mathbf{r}) = -4\pi e^2 d\Delta\bar{Q}/\varepsilon_i.$$

In this approximation τ_e is independent of V_g since the shift of all the levels is the same. But the $\tau_c(V_g)$ dependence for $l_0 \gg d$ is stronger:

$$(T/e)\partial\ln\tau_c/\partial V_g \approx -(1 + 4\pi e^2 d\bar{Q}/\varepsilon_i\Delta)/(\varepsilon_s d/\varepsilon_i w + 4\pi e^2 d\bar{Q}/\varepsilon_i\Delta). \quad (45)$$

The right-hand side for $4\pi e^2 d\bar{Q}/\varepsilon_i\Delta \gg 1$ is equal to unity, then it grows like $\bar{Q}^{-1}$ as $\bar{Q}$ decreases, and for

$4\pi e^2 w\overline{Q}/\varepsilon_s\Delta \ll 1$ reaches its maximum value of $\varepsilon_i w/\varepsilon_s d$. The right-hand sides of (43) and (44) are always small compared to unity.

When $\overline{Q}_d > \overline{Q}_l$, Eqs. (43) and (44) may change quantitatively because at such values of $\overline{Q}_d$ the height of the barrier is not very great ($\leqslant 2\Delta$). Hence, when $\overline{Q}$ increases by $\Delta\overline{Q}$, the drop in the excess charge density inside the barrier is smaller than $\Delta\overline{Q}$ and increases from the edge of the barrier to the center, that is, $r_{0,1,0}$ in (43) acquires an additional factor smaller than unity, and l_0 acquires a still smaller factor. For large values of $\overline{Q}_d$ and at the correspondingly low temperatures of RTS observation, the l_0 to $r_{0,1,0}$ ratio is fairly moderate and $\partial\varepsilon_F/\partial V_g$ is small. This leads to strong cancellation of the terms on the right-hand side of Eq. (43), while Eq. (44) undergoes change caused by the fact that for $2D$ electrons tunneling to the trap in the l plane Eq. (42) is invalid.

The temperature dependence of τ_c and τ_e [see Eqs. (19)–(22)] can primarily be described by straight lines on $\ln\tau_{c,e}$ vs T^{-1} diagrams. These may have breaks corresponding to jumps in ε_T from one level to another. [This leads to jumps in the factor γ in Eq. (44) and breaks in $\ln\tau_{c,e}(V_g)$ curves]. At breaks the slope of the straight lines must be smaller on the lower temperature side.

Using the temperature curves for τ_c and τ_e to determine the energies ε_T, ε_t, and ε_F, we must allow for the temperature dependence of these energies. One reason for this is the large entropy of electronic ionization in Si at room temperature.[22,23] For the band gap of Si we have $\partial\varepsilon_g/\partial T = -\Delta S_{CV}$, with $\Delta S_{CV} \approx 3$ (see Ref. 22), and for the binding energy of the Coulomb center, ε', we have $\partial\varepsilon'/\partial T = -\Delta S_I$, where $\Delta S_I \approx 0$ for a hydrogen like center and increases with the contribution from the central cell, that is, with ε', up to ΔS_{CV} (see Ref. 23). For slow surface traps studied at ~ 300 K the typical values of ε' lie in the 0.5 to 0.7 eV range and, therefore, ΔS_I for such traps can be expected to be high (for similar traps in the bulk of Si, $\Delta S_I \sim 2$), although the effect of SiO_2 and of the vibrations of separate charges in the distributed nucleus in unclear. The random structure of the nuclei may explain why ΔS_I differs from trap to trap. The increase in ε_T with T is caused not only by the diminishing binding energy on the nucleus but also by the increase in the height of the barrier caused by the weakening in electron screening. The latter occurs for the following reasons. The value of ε_F is determined by the position of the Fermi level in the source (sink), which is an n^+-region doped so heavily that it becomes highly degenerate. In it, therefore, the distance from the Fermi level to the bottom of the conduction band is temperature-independent. The position of the latter is determined by the electron affinity energy χ. If we write $\partial\chi/\partial T = \Delta S_\chi$ (the results of van Vechten and Thurmond[23] suggest that $\Delta S_\chi \approx 0.8\ \Delta S_{CV}$), we find that at a constant gate-source voltage as T increases by δT the Fermi level in the inversion channel drops in relation to the Fermi level in the metal by $\Delta S_\chi \delta T$, which is equivalent to V_g changing by $\delta V_g = -\Delta S_\chi \delta T/e$. This yields

$$\partial(\varepsilon_t-\varepsilon_F)/\partial T = -\Delta S_\chi \partial(\varepsilon_t-\varepsilon_F)/\partial e V_g + \Delta S_I. \qquad (46)$$

From Eqs. (46) and (43) [or (45)] we find the temperature variations of the trap level. They, obviously, depend on the electron number density (owing to V_g) and increase as functions of l_0. For instance, for $l_0 \gg d$ and small $\overline{Q}$ [see Eqs. (45) and (46)], $\partial(\varepsilon_t - \varepsilon_F)/\partial T$ may reach $\Delta S_I + \Delta S_\chi \varepsilon_i w/\varepsilon_s d$.

7. DISCUSSION OF THE RESULTS OF RTS STUDIES

Up till now, in accordance with McWhorter's hypothesis,[24] surface traps were assumed to reside in the insulator. This was considered the reason for $1/f$ noise. Long charge-exchange times for the traps and their considerable spread, needed for flicker noise generation,[25] were explained quite logically by the rapid lowering in the probability of tunneling below a high barrier for traps located farther from the semiconductor interface. In the Si:SiO_2 system the barrier for electrons is roughly 3.2 eV high and that for holes ~ 4.7 eV. Estimates and extensive studies of Me–Si_3N_4–SiO_2–Si and Me–SiO_2–Si structures with tunnel-thin layers of SiO_2 (see, e.g., Ref. 1) have shown that the length over which the transmission coefficient decreases by a factor of e is very small even for electrons: $\lambda_t \leqslant 1$ Å. Hence, to describe flicker noise in Si:SiO_2 structures, traps were employed that were at a distance h of roughly 10–20 Å from the interface. Even for such barriers the charge-exchange times, proportional to $\exp(h/\lambda_t)$, are very long.

RTS studies have revealed that the main reason for large charge-exchange times in slow surface traps is the low probability of thermal activation of carriers to a higher energy rather than the temperature-dependent low tunneling probability. [For instance, according to the data of Ref. 10, for traps studied at ~ 300 K it was found that σ is on the order of 10^{-22}–10^{-26} cm^2 and varies like $\sigma = \sigma_0 \exp(-\Delta E_B/T)$, with the constant σ_0 being either large or moderate ($\sim 10^{-14}$–10^{-19} cm^2).] In view of this, starting with Ref. 9, the researchers involved in RTS studies explain the long charge-exchange times in slow surface traps by multiphonon capture of carriers to localized states in the insulator. The capture multiphonon cross sections are low and may be exponentially temperature-dependent.[26,27] The conclusions made in Refs. 26 and 27 refer to the bulk situation, however. If in the process of capture a carrier goes from the semiconductor to the insulator, σ includes not only the low probability of transferring a large amount of energy from the carrier to the lattice but also the probability of spatial tunneling of the carrier in the insulator. As noted earlier, for the Si:SiO_2 system this latter probability is very low if the trap is located farther than 10–20 Å from the interface, and hence even for such values of h there is no rational explanation why the measured values of σ_0 are so high, the more so that the radius of the wave function of the state lying more than 3 eV below the bottom of the conduction band in SiO_2 is of the order of 1 Å.

Yet, proceeding from the fact that the traps reside in the insulator and therefore, when V_g varies their levels shift in relation to the bottom of the conduction band, Ralls *et al.*[9] established h using the formula

$$h = -d\,\partial\varepsilon_t/\partial e V_g. \qquad (47)$$

(Subsequently various refinements were introduced into this formula; see Ref. 10). Most often the obtained values of h lie in the 10–20 Å range, that is, are unacceptably high. In addition, since the range of charge-exchange times of these traps is not too wide, one would expect such a correlation between the thermal activation and tunneling exponentials that the product of the two would change little. Yet, Ralls *et al.*,[9] who studied the traps over a broad temperature range, observed essentially the opposite correlation: traps with a larger value of $\exp(\Delta E_B/T)$ had a greater value of h. And, finally, Kirton and Uren[10] obtained for a number of traps unrealistically large values of h, up to roughly 200 Å.

But if formula (47) or its analogs are not employed and the traps are assumed to be very close to the interface, it is difficult to explain the observed strong dependence (especially for weak inversion) on V_g of a number of parameters of such highly localized states.

The theory of Coulomb fluctuation traps developed in the present paper allows for a unified approach in describing the entire body of data of RTS studies. The following can be added to the conclusions on which this statement rests (see the end of Sec. 4). The theory illuminates the amplitude of the observed RTS, the amplitude variation from trap to trap, and the dependence of the amplitude on the sample's conductivity. It explains the activation dependence of τ_c and τ_e, the observed activation energies, and the characteristic tendency of the activation exponents to be often very large at high T and much smaller at low T (because when carriers are captured through lower resonance states the tunneling exponent is large). For instance, in Ref. 9 for a trap studied in the 101 to 111 K range the interval of variation of the activation exponent was found to be 3×10^9–3×10^{10} for τ_c and 3×10^{12}–3×10^{14} for τ_e, and for a trap studied in the 26 to 34 K range these intervals were 10^4–10^5 and 5×10^5–3×10^7, respectively. Note that a similar tendency is characteristic of multiphoton capture, but here the dependence of $\ln\sigma$ on T^{-1} should differ considerably from linear.[27] For fluctuation traps, on the other hand, this dependence may be close to linear over the entire range of variation of τ_c and τ_e, provided that the difference between the energies of resonance states is great.

The fluctuation theory explains the dependence of τ_c and τ_e on V_g. According to it, over a broad range of surface band bending, τ_c decreases as V_g grows, while τ_e increases. Another conjecture that agrees with the experimental data is that the dependence of τ_c on V_g must be strong for weak inversion and much weaker for strong inversion, when l_0 is smaller. The dependence of τ_e on V_g is weak for all levels of inversion. Comparing formulas (43) and (47), we find that $l_0\sim4\varkappa h/\varepsilon_i$, that is, for $h\sim10$–20 Å we have $l_0\sim100$–200 Å. But for $h\sim200$ Å, then we have $l_0>d$ [see Eqs. (45) and (47)].

The fluctuation theory also makes it clear why the trap level is found to have a drastic temperature shift at roughly 300 K and why $\partial\varepsilon_t/\partial T$ depends strongly on V_g for weak inversion (for traps with $l_0\gg d$). At the same time, there are differences here between theory and the experimental data of Ref. 10. The reason may be that Ref. 10 reports the results of processing the experimental data in a way that does not quite agree with the fluctuation theory of activation conductivity given in Ref. 5. Most likely, however, the fluctuation theory of slow traps needs to be further developed. A detailed verification of agreement of theory and experiment should provide an impetus, especially since the present paper studies only the principal features of the properties of such traps; hence a number of effects have remained outside its scope.

8. CONCLUSION

Agreement between the conclusions of the fluctuation theory and the data gathered in RTS studies is too multifaceted to be accidental, the more so since the studies have provided exceptionally complete information about the kinetic and thermodynamic characteristics of a slow trap obtained in direct measurements in a single experiment and their nontrivial dependence on external conditions, and also a large volume of information on the properties of various traps in a broad range of external conditions and the characteristic change in these properties. All this strongly suggests that slow surface traps are generated by fluctuations of the charge built into the insulator and constitute planar nanostructures built from Coulomb centers according to certain rules. Hence there are three important conclusions:

(a) Studying the structure of the nucleus of a slow surface trap with an RTS as the detecting device may provide unique information about the heterojunction in Si:SiO_2, the centers generating the built-in charge, their potential, and the size of the charge of the nucleus. In addition to the methods already used, one of special interest is the optical spectroscopy of a trap (which is possible for moderate or even weak illumination), since at frequencies of transitions between the trap levels one can expect a strong resonant drop in τ_e. Another variant of the spectroscopy of the resonance levels of a trap consists in determining the optimal-capture energy from a diagram of $\ln\tau_e$ vs T^{-1} and recording its jumps from one resonance level to another caused by variations in T, V_g, magnetic field strength, etc. Also, the temperature curves of the level energies may provide information about the phonon modes near the Si:SiO_2 interface.

Each slow surface trap is a unique instrument for studying such surface electronic properties as electron scattering and band structure. Here we note only the spectroscopic possibilities arising from the fact that τ_c and τ_e are inversely proportional to the density of electronic states at the optimal-capture level. For instance, in the case of metallic conduction and low temperatures, when the optimal-capture level coincides with the trap level, moving the latter in relation to the Fermi level (by varying V_g or the magnetic field strength) and determining their mutual position from the τ_c to τ_e ratio allows a quick estimate of the density of states by observing the variations of τ_c and τ_e proper. This method should prove especially effective when there is a gap in the density of states near the Fermi level.

(b) The present work indirectly corroborates the hypothesis put forward in Ref. 28 that random Coulomb barriers surrounding electron traps are the reason common to all semiconductors for the change in the spectrum of generation-recombination noise to $1/f$ noise. Actually we have discussed here one of the mechanisms suggested in Ref. 28 to explain the formation of slow traps. But to confirm the fact that it is these traps that cause the observed surface flicker noise, we have chosen another avenue, which allows for higher reliability, since the signals of random charge-exchange from a single trap are much more informative than $1/f$ noise, which is a linear combination of signals from many traps. Apparently, in other cases, too, identification of the source of this noise will be more reliable in RTS studies, the more so that RTS similar in many respects to the case just discussed were later observed in heterojunctions,[29] MOS-tunnel diodes,[30] and tunnel (metal–insulator–metal) diodes,[31] and much earlier these signals were often observed in studies of various structures with *p–n* junctions and were named burst, or popcorn noise.

(c) The statistics of formation of fluctuation traps suggests that even with the best MOS structures used in RTS studies the total density of built-in charged centers, $\Sigma = \Sigma_+ + \Sigma_-$, is much higher than $|\Sigma_+ - \Sigma_-|$ (earlier this was implicitly suggested by the results of Refs. 4, 5, and 33 and other papers), which determines the density of the built-in charge. The quantity Σ is an essential criterion for determining the quality of MOS structures used in scientific studies and devices. It determines the characteristics of MOSFETs through the surface carrier mobility, the density of surface states, the hysteresis of the state of the surface, and the noise power. For instance, the power of $1/f$ noise is a very important criterion for improving the quality of MOSFETs used as transducers of weak analog signals, say, in matrix photodetectors. Often it is this noise that determines the threshold of signal detection, which is understandable since transistors operate in this case at low channel conductivity and low temperatures, when $1/f$ noise, according to the theory developed above, is exceptionally strong.

The author is grateful to N. M. Storonskiĭ and R. A. Suris for helpful discussions.

[1)] On the Si[100] surface and for $\bar{Q}_d \ll \bar{Q}_l$ the direction of optimal electron tunneling may deviate from the x axis since the electron effective mass along this axis is much greater than along l.

[2)] Formula (2) was derived in the quasiclassical approximation and is valid for contaminated surfaces with $\bar{\Sigma} \ll 1$. At $\bar{\Sigma} \sim 1$ the pre-exponential factor in (2) is, apparently, somewhat overestimated. Hence the weakening of potential barriers by electron screening can only increase further.

[1]S. Sze, *Physics of Semiconductor Devices*, Wiley, New York (1981), Chaps. 7 and 8.

[2]E. H. Nicollian and J. R. Brews, *MOS (Metal-Oxide Semiconductor) Physics and Technology*, Wiley, New York, (1982), Chaps. 7 and 8.

[3]T. Ando, A. Fowler, and F. Stern, *Electronic Properties of Two-dimensional Systems*, American Physical Society, New York (1982), Chap. 3.

[4]B. Henderson, R. L. Vranch, and R. L. Vranch, Semicond. Sci. Technol. **4**, 1045 (1989).

[5]V. A. Gergel' and R. A. Suris, Zh. Eksp. Teor. Fiz. **84**, 719 (1983) [Sov. Phys. Uspekhi **57**, 415 (1983)].

[6]N. M. Storonskiĭ and B. I. Fuks, Zh. Eksp. Teor. Fiz. **96**, 1880 (1989) [Sov. Phys. Uspekhi **69**, 1061 (1989)].

[7]V. L. Bonch-Bruevich, *Collection of Articles on Solid-State Physics*, Vol. 2, Nauka, Moscow (1959) [in Russian], p. 182.

[8]V. N. Abakumov, V. Karpus, and I. N. Yassievich, Fiz. Tech. Poluprovodn. **22**, 262 (1988) [Sov. Phys. Semicond. **22**, 159 (1988)].

[9]K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, Phys. Rev,. Lett. **52**, 228 (1984).

[10]M. J. Kirton and M. J. Uren, Adv. Phys. **38**, 367 (1989).

[11]B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors*, Springer, Berlin (1984), Chap. 12.

[12]S. T. Pantelides, Rev. Mod. Phys. **50**, 797 (1978).

[13]N. M. Storonskiĭ and B. I. Fuks, Zh. Eksp. Teor. Fiz. **100**, 1873 (1991) [Sov. Phys. Uspekhi **73**, 1036 (1991)].

[14]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, Pergamon, Oxford (1977).

[15]V. N. Abakumov, V. I. Perel' and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **12**, 3 (1978) [Sov. Phys. Semicond. **12**, 1 (1978)].

[16]M. J. Kirton, M. J. Uren, S. Collins, M. Schultz, A. Karmann, and K. Scheffer, Semicond. Sci. Technol. **4**, 1116 (1989).

[17]W. J. Skocpol, P. M. Mankievich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, Phys. Rev. Lett. **56**, 2865 (1986).

[18]B. L. Al'tshuler and B. Z. Spivak, Pis'ma Zh. Eksp. Teor. Fiz. **42**, 363 (1985) [JETP Lett. **42**, 447 (1985)].

[19]S. Feng, P. A. Lee, and A. D. Stone, Phys. Rev. Lett. **56**, 1960 (1986).

[20]P. A. Lee and A. D. Stone, Phys. Rev. Lett. **55**, 1622 (1985).

[21]P. Restle, Appl. Phys. Lett. **53**, 1862 (1988).

[22]V. Heine and J. A. van Vechten, Phys. Rev. B **13**, 1622 (1976).

[23]J. A. van Vechten and C. D. Thurmond, Phys. Rev. B **14**, 3539 (1976).

[24]A. L. McWhorter, in *Semiconductor Surface Physics*, edited by R. H. Kingston, Univ. of Pennsylvania Press, Philadelphia (1975), p. 207.

[25]Sh. M. Kogan, Usp. Fiz. Nauk **145**, 285 (1985) [Sov. Phys. Uspekhi **28**, 170 (1985)]; M. B. Weissman, Rev. Mod. Phys. **60**, 537 (1988).

[26]C. H. Henry and D. V. Lang, Phys. Rev. B **15**, 989 (1977).

[27]V. N. Abakumov, I. A. Merkulov, V. I. Perel', and I. N. Yassievich, Zh. Eksp. Teor. Phys. **89**, 1472 (1985) [Sov. Phys. JETP **62**, 853 (1985)].

[28]B. I. Fuks, Fiz. Tekh. Poluprovodn. **24**, 2036 (1990) [Sov. Phys. Semicond. **24**, 1265 (1990)].

[29]T. Judd, N. R. Couch, P. H. Beton, M. J. Kelly, T. M. Kerr, and M. Pepper, Appl. Phys. Lett. **49**, 1652 (1986).

[30]K. R. Farmer, C. T. Rogers, and R. A. Buhrman, Phys. Rev. Lett. **58**, 2255 (1987).

[31]C. T. Rogers and R. A. Buhrman, Phys. Rev. Lett. **55**, 859 (1985).

[32]M. Buckingham, *Noise in Electronic Devices and Systems*, Ellis Horwood, Chichester (1983).

[33]V. A. Gergel', Fiz. Tekh. Poluprovodn. **17**, 637 (1983) [Sov. Phys. Semicond. **17**, 398 (1983)].

Translated by Eugene Yankovsky