Density of localized states in the surface impurity band of a metal-insulator-semiconductor structure

M. S. Bello, E. I. Levin, B. I. Shklovskiĭ, and A. L. Éfros

A. F. Ioffe Physicotechnical Institute, USSR Academy of Sciences

(Submitted 20 October 1980)
Zh. Eksp. Teor. Fiz. **80**, 1596–1612 (April 1981)

We investigate theoretically the effective density of localized states in a metal-insulator-semiconductor structure, as determined by capacitance and electric-conductivity measurments. We consider a two-dimensional impurity band made up of donors and acceptors located on the surface. Account is taken also of the scatter of energies of non-Coulomb origin. It is shown that the Coulomb interaction can cause the effective state density to differ substantially from the single-electron state density. In the limit of almost complete occupation of the impurity band by electrons, when the density of the charged donors is low, these donors form a spatially ordered structure (a Wigner crystal), thereby strongly affecting the properties of the system. In particular, the Fermi level in the Wigner-crystal level is lowered with increasing electron density; this corresponds to a negative effective density of states. The percolation threshold and the activation energy of the band conductivity in the Wigner-crystal region are analytically calculated.

PACS numbers: 73.40.Qv, 71.20. + c, 71.55. – i

1. INTRODUCTION

Metal–insulator–semiconductor (MIS) structures have become the subject of intensive study in the last decade. By varying the potential V_g applied between the metal and the semiconductor (Fig. 1a) it is possible to vary the electron density on the surface of the semiconductor in such a structure. At low temperature, in a wide range of V_g, the Fermi level E_F is in the region of localized states, as is attested by the activation character of the temperature dependence of the inversion-layer surface conductivity. The electrons are localized either on traps imbedded in the insulator, or on sodium ions that can be specially introduced into the structure[1]; these ions are distributed over the semiconductor surface and form a two-dimensional impurity band (Fig. 1b). By varying V_g one can move the impurity levels relative to the Fermi level, thereby depleting or filling the impurity band. This offers ample opportunities to study the structure of the impurity band, and primarily the state density (SD) in it.

Two methods are used in the main to determine the SD in experiment. According to the first, the SD on the Fermi level is given by the relation

$$G_c(\mu)=dn/d\mu, \tag{1}$$

where n is the surface density of the electrons and μ is the Fermi energy reckoned from some energy that characterizes the band structure on the semiconductor surface. In the present paper we reckon μ from the energy of an isolated impurity, located at a level E_0 below the edge of the unperturbed conduction band. The $n(\mu)$ dependence is determined by analysing the volt-farad characteristics of the MIS structure.[2] The SD determined in this manner will be called capacitive.

The other method is used as a rule in the region where the states are localized.[3] With increasing electron density in the impurity band, a change takes place in the activation energy ε_a of the surface conductivity due to the spilling of the electrons into the conduction band (Fig. 1b). The state density G_{act} is defined as

$$G_{act}=-dn/d\varepsilon_a. \tag{2}$$

With change of the potential V_g, changes occur in the activation energy and in the surface electron density n: $\Delta n=\Delta V_g C_d e^{-1}$, where C_d is the capacitance of the insulator and e is the absolute value of the electron charge. Substituting the changes of both quantities in (2), we obtain G_{act}.

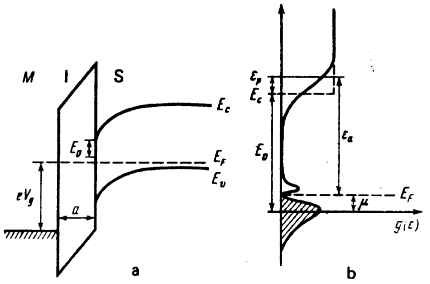
The energy ε_a is equal to the energy distance from the Fermi level to the mobility threshold ε_p of the conduction band (Fig. 1b). Measuring ε_p from the unperturbed bottom of the conduction band, we obtain

$$\varepsilon_a=E_0+\varepsilon_p-\mu. \tag{3}$$

Thus, the SD G_c and G_{act} represent different quantities. Moreover, to determine G_{act} we must know not only the function $\mu(n)$ but also $\varepsilon_p(n)$. If we disregard the electron-arrangement correlations connected with their Coulomb interaction and describe this interaction only with the aid of an average self-consistent potential, then G_c coincides with the single-electron SD $g(\varepsilon)$ on the Fermi level (details follow). However, as indicated earlier,[4,5] the single-electron SD $g(\varepsilon)$ vanishes on the Fermi level because of the correlations. It is clear therefore that the SD G_c and G_{act} defined above do not coincide with $g(\mu)$ and do not have as simple an interpretation as in the absence of interaction between the electrons. The purpose of the present paper is to calculate the SD G_c and G_{act} with this interaction taken into account.

We consider below an inversion-layer model that describes best the situation that arises in experiments such as those of Fowler and Harstein,[1] when the localized states of the electrons are produced on sodium ions that serve as donors in the two-dimensional impurity band produced in an n-inversion layer on a p-Si substrate. We propose a two-dimensional Poisson distribution of the donors and acceptors, with an arbitrary degree of compensation $K=N_A/N_D$ (N_A and N_D are two-dimensional densities). Certain donors are occupied by electrons having a density n, and are neutral; the ion-

0038-5646/81/040822-08$02.40



ized donors are positively charged and all the acceptors are negative. The distances between donors are assumed large compared with the localization radius, so that the overlap of the electron wave functions on the neighboring donors can be neglected and the system can be assumed to be strictly classical. We are actually dealing here with a two-dimensional classical impurity band. The distinguishing feature of the problem is, however, that the considered plane is generally speaking not neutral but has an average surface charge density $\sigma=e(N_D-N_A-n)$. In addition, since a metallic electrode is located on the left, at a distance a from the plane in question (see Fig. 1), the interaction of the charges must be investigated with account taken of the image forces.

Our model constitutes therefore a parallel-plate capacitor with a charge density σ. The capacitor is made up of a plane with randomly distributed donors and acceptors, and of a metallic plate. An emf applied to this capacitor maintains an electrochemical-potential difference E_F between its plates and regulates its charge (Fig. 2). (It is assumed in Fig. 2 that the level E_0 of the isolated donor coincides in the absence of an external voltage with the Fermi level in the metal. It can be shown that this assumption is not essential.)

To generalize the model we introduce a random scatter of the donor levels Φ_i of non-Coulomb origin. The energies Φ_i are assumed to be uniformly distributed in the interval from $-A$ to A. The total energy of the system takes thus the form

$$H=\sum_i^D n_i\Phi_i+{}^1/_2\sum_{ij}^D(1-n_i)(1-n_j)e_{ij}-\sum_i^D\sum_k^A(1-n_i)e_{ik}+{}^1/_2\sum_{k,l}^A e_{kl}. \tag{4}$$

Here e_{ij} is the energy of interaction between two charges with allowance for the image forces:

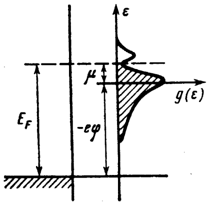


FIG. 2. Energy diagram and state density in the considered model.

$$e_{ij}=\frac{e^2}{\varkappa}\left[\frac{1}{r_{ij}}-\frac{1}{(r_{ij}^2+4a^2)^{1/2}}\right],\quad e_{ii}=-\frac{e^2}{2\varkappa a}, \tag{5}$$

where r_{ij} is the distance between the charges and the plane in question, and $\varkappa$ is the lattice permittivity, which can be assumed for real structures to be equal, with good accuracy, to the arithmetic mean of the permittivities of the insulator and the semiconductor.[3] The donor occupation numbers n_i are equal to unity if the donor is filled, and to zero if its ionized. The letters D and A over the summation sign indicate that the summation is over all the donors or all the acceptors.

To find the electron density n corresponding to a given electrochemical potential E_F, we must find a set $\{n_i\}$ that minimizes the quantity

$$\tilde{H}=H-E_F\sum_i^D n_i. \tag{6}$$

The density is

$$n=\sum_i^D\frac{n_i}{S},$$

where S is the area of the system.

The electrochemical potential E_F is reckoned from the energy of the electron on the donor in the absence of other charges. If the capacitor is not charged, this potential coincides with the Fermi energy μ of interest to us. In a charged capacitor, however, the value of E_F, just as the electron energy on an isolated donor, is raised or lowered (on top of the change due to the filling of new levels) by an amount $-e\varphi$ where $\varphi=4\pi\sigma a\varkappa^{-1}$ is the potential of the parallel-plate capacitor. Therefore the electrochemical potential E_F is connected with the Fermi energy μ by the relation (see Fig. 2)

$$E_F=\mu-4\pi\sigma a e\varkappa^{-1}. \tag{7}$$

The function $E_F(n)$ should be obtained by minimizing expression (6). It is more convenient instead to minimize the energy H_0, which differs from (4) by the amount of energy of a parallel-plate capacitor with charge σ:

$$H_0=H-2\pi a\sigma^2 S\varkappa^{-1}. \tag{8}$$

The minimization, given the electron density n, yields a function $H_0(n)$ such that

$$\mu=\frac{1}{S}\frac{dH_0}{dn}. \tag{9}$$

Minimization of H_0 makes it also possible to obtain a set of single-electron energies ε_i that constitute the potential energies of the electrons on the donors in the ground state (reckoned from the average potential energy $4\pi\sigma a\varkappa^{-1}$):

$$\varepsilon_i=\delta H_0/\delta n_i=\Phi_i-\sum_{j\neq i}^D e_{ij}(1-n_j)+\sum_k^A e_{ik}-e_{ii}/2+4\pi\sigma e\varkappa^{-1}. \tag{10}$$

[We have left out a term $(2n_i-1)2\pi ae^2/\varkappa S$, assuming that $SN_D^{1/2}\gg a$ since the entire system is macroscopic.]

The donor i is empty if $\varepsilon_i>\mu$ and occupied by an electron if $\varepsilon_i<\mu$. According to the earlier studies[4,5] the single-electron state density $g(\varepsilon)$ has a Coulomb gap. In the case of pure Coulomb interaction, $g(\varepsilon)$ vanishes in the vicinity of the Fermi level like

$$g(\varepsilon)=2|\varepsilon-\mu|\varkappa^2/\pi e^4. \tag{11}$$

In the case considered here the metal plate causes screening and transforms the Coulomb interaction into dipole interaction at $r_{ij} \gg a$. Therefore $g(\mu)$ differs from zero and is of the order of $\varkappa/ae^2$. We shall show in this paper that the SD G_c defined by (1) has nothing in common with $g(\mu)$.

The Coulomb gap stems in essence from the fact that the discrete character of the electron charge and the localization of the electron states bring about a correlation in the electron arrangement due to the electron interaction. If it is assumed that all the charges in the system are smeared, then the interaction energy in (4) is equal to the energy of a parallel-plate capacitor with charge σ. The energy H_0 [see (8)] therefore includes only energy of non-Coulomb origin, i.e.,

$$H_0=\sum_i^D \Phi_i n_i = S\int_{-\infty}^{\mu} g_0(\Phi)\Phi\, d\Phi,$$

where $g_0(\Phi)$ is the given SD, and the chemical potential μ is defined by the condition

$$\int_{-\infty}^{\mu} g_0(\Phi)\, d\Phi = n.$$

It follows then from (1) that

$$G_c=g_0(\mu). \tag{12}$$

In this simplest case G_c coincides with the single-electron state density $g_0(\Phi)$ on the Fermi level. However, when account is taken of the correlations in the arrangement of the electrons connected with the electron-electron interaction, this no longer is so. In the general case one can only prove the inequality

$$G_c^{-1}=d\mu/dn>-4\pi e^2 a\varkappa^{-1}. \tag{13}$$

It follows from the fact that the energy H defined in (4) has at a given value of E_F a minimum at a certain density n. At this density we have

$$d^2H/dn^2=dE_F/dn>0.$$

Equation (7) leads then to (13). We shall show below that the observed SD can indeed be negative.

Analytic calculation of ε_p and μ is possible only in certain limiting cases, which will be indicated below. To construct the complete picture, we have undertaken to simulate the system indicated above with a computer, modifying somewhat the program used in the preceding paper[7] to calculate a three-dimensional impurity band. In the next section we describe the program and then report and explain the results for different degrees of compensation.

2. DESCRIPTION OF SIMULATION PROGRAM

Since the computer used by us was incapable of simulating a sufficiently large system, an important factor in the simulation was the choice of the boundary conditions. We have assumed the investigated system to be square in form, and that the remainder of the plane containing this square has a uniform charge of density σ. This enables us to avoid the large electric-field distortions that inevitably arise on the edges of the system if a capacitor with finite dimensions is considered. To calculate the total energy H, it is convenient to represent the obtained system as a uniformly charged plane with charge σ, on which is superimposed as a whole a neutral square consisting of a uniform background with charge density $-\sigma$ on which donors and acceptors are randomly distributed. Then

$$H=H_{sq}+2\pi a\sigma^2 S\varkappa^{-1}. \tag{14}$$

The first term is the inner electrostatic energy of a neutral square, the second is the self-energy of the "large" capacitor (of area S) containing the square. The energy of the charges of the neutral square in the uniform field of the large capacitor is zero. On the other hand it is easily seen that by analogy with the statements made in the Introduction [see (8)] it is necessary to minimize the energy $H_0=H-2\pi a^2 S\varkappa^{-1}$. According to (14), $H_0=H_{sq}$, so that the problem reduces to minimization of the electrostatic energy of a neutral square at a given electron density.

Let

$$\varphi_0(\mathbf{r})=\frac{e}{\varkappa S_0}\int_{sq} d^2\mathbf{r}'\left[\frac{1}{|\mathbf{r}-\mathbf{r}'|}-\frac{1}{((\mathbf{r}-\mathbf{r}')^2+4a^2)^{1/2}}\right] \tag{15}$$

be the potential produced by a square having a uniform charge equal to e. The integration is over the square, its area is S_0. Then the potential produced by the uniform charge $-\sigma$ inside the square (by the background) is $-S_0\sigma\varphi_0(\mathbf{r})/e$. The energy $H_0=H_{sq}$ can be represented in the form

$$H_0={}^1/_2\sum_{i,j}^D(1-n_i)(1-n_j)e_{ij}-\sum_i^D(1-n_i)\sum_k^A e_{ik}$$
$$+\sigma\sum_i^D(1-n_i)\varphi_0(\mathbf{r}_i)+\sigma\sum_k^A\varphi_0(\mathbf{r}_k)+{}^1/_2\sigma^2\bar{\varphi}S_0^2+\sum_i^D n_i\Phi_i, \tag{16}$$

$$\bar{\varphi}=\frac{1}{S_0}\int_{sq} d^2\mathbf{r}\,\varphi_0(\mathbf{r}). \tag{17}$$

The first two terms in (16) describe the interaction of the charged donors and acceptors, the third and fourth describe the interaction of the charged donors and acceptors with the background, and the fifth, the self energy of the background. We have left out of (16) terms that describe the interaction of the charged acceptors, inasmuch as these terms do not contain the occupation numbers and are therefore inessential for the minimization. Algebraic transformations reduce (16) to the form

$$H_0={}^1/_2\sum_{i\neq j}^D q_{ij}n_i n_j+\sum_i^D n_i E_i, \tag{18}$$

where

$$E_i=\sum_j^A q_{ij}-\sum_j^D q_{ij}+{}^1/_2 q_{ii}+\Phi_i, \tag{19}$$

$$q_{ij}=e_{ij}-e\varphi_0(\mathbf{r}_i)-e\varphi_0(\mathbf{r}_j)+e\bar{\varphi}. \tag{20}$$

On going from (16) to (18) we have left out a number of terms that do not depend on n_i. The electron energy on the donor i is of the form

$$\varepsilon_i=\delta H_0/\delta n_i=E_i+\sum_{j\neq i}^D q_{ij}n_j. \tag{21}$$

The algorithm of the minimization program contains much that has been borrowed from an earlier paper,[7] and will not be described here in detail. In the square considered, the random-number generator produced the coordinates of N donors and K_N acceptors. The donor and acceptor charges were assigned values +1 and −1 respectively. Each donor could also be assigned a random value Φ_i uniformly distributed in the interval from $-A$ to A. The integral (15) can be calculated analytically, and (17) numerically. With the aid of (21) we calculated the energies ε_i at all donors at $n_i \equiv 0$. These energies included the minimal energy $\varepsilon_{\min}$. A group of p electrons was then added to the system. The first electron occupied a donor with minimum energy $\varepsilon_{\min}$. All the donor energies were then recalculated with account taken of the potential of this electron and of the change of the background charge, and a new energy $\varepsilon_{\min}$ was found. The second electron occupied a donor with this energy, etc. The change of the total energy H_0 following the landing of each electron was calculated. The system of p electrons obtained through this procedure, generally speaking, is not in the ground state, and a subprogram was used to search for the pseudo-ground state and to obtain an energywise more favorable rearrangement of the electrons. This subprogram was described in detail in Ref. 8. All the decreases produced in the total energy H_0 by these permutations were taken into account. The result were sets of energies ε_i and of occupation numbers n_i corresponding to the pseudo-ground state.[1)] They could be used to determine the SD $g(\varepsilon)$ and the Fermi level μ. The Fermi level was obtained by two different methods. The first corresponded directly to the definition (9). It consisted of dividing by p the change of the total energy H_0 following the landing of p electrons, followed by minimization. In the second method the Fermi level was defined as the arithmetic mean of the maximum energy $\varepsilon_{\max}$ of the occupied donor and the minimum energy $\varepsilon_{\min}$ of the empty donor, i.e., as the energy borderline between the empty and occupied states. The next p electrons were then added and the process repeated. As a rule, $p = N/100$ and one realization of the coordinates of the donors and acceptors yielded a relatively smooth $\mu(n)$ curve. This function was next averaged over several realizations.

3. SIMULATION RESULTS AND THEIR DISCUSSION

The most important results are shown in Fig. 3 in the form of a plot of the Fermi level μ vs the electron density n. We have assumed here that the level spread is only of Coulomb origin and we have put $\Phi_i = 0$. The Fermi level was obtained with the aid of (9). The presented data make it possible to determine G_c with the aid of relation (1).

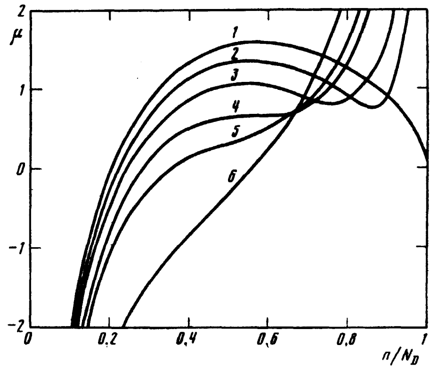


FIG. 3. Fermi level μ (in units of $e^2N_D^{1/2}/\varkappa$) vs. n/N_D at $K=0$ (1), 0.1 (2), 0.2 (3), 0.35 (4), 0.5 (5) and 1 (6). Aggregates of 400 donors were used. Each curve is the result of averaging over 20 realizations.

We begin the discussion with compensation $K \geqslant 0.35$. As seen from Fig. 3, the corresponding curves have two steep sections at large and small electron densities. This can be easily explained. The electrons strive to occupy those donor states whose energies are lowered the most by the potential of the other charged donors and acceptors. If the electron density is low and almost all the donors are empty, these states are produced on the donor pairs, each comprising two donors anomalously close to each other. The electron occupies only one of the donors of each pair, while the second charged donor lowers the energy of this electron. In the three-dimensional case, at a higher degree of compensation, an important role is played also by the large-scale potential (see Ref. 6), but it is easy to verify that in the two-dimensional case it is insignificant, since the mean squared potential of a planar system of randomly arranged Coulomb center has at large distances only a logarithmic divergence, which is eliminated by electron screening or by screening by image charges induced by the metallic surface (see Ref. 9). It is easy to calculate the number of donor pairs that lower the electron energy below the Fermi energy, and thus express μ in terms of n. The probability that another donor is present in a circle of radius $r_\mu = e^2/\varkappa|\mu|$ around a given donor is equal to $\pi N_D r_\mu^2$, where N_D is the surface donor density. The surface pair density is obtained from this by multiplying by N_D and dividing by 2. We obtain $\pi N_D^2 r_\mu^2/2 = n$ or

$$\mu = -\left(\frac{\pi}{2}\right)^{1/2} \frac{e^2 N_D^{1/2}}{\varkappa} \left(\frac{N_D}{n}\right)^{1/2}. \tag{22}$$

This calculation is valid if $n \ll N_D$.

It is easy also to calculate the $\mu(n)$ dependence as $n \to N_D$. In this case almost all the donors are filled, and the rare empty donors are anomalously close to the acceptors. The density of acceptors with a donor at a distance shorter than r_μ is $\pi N_D N_A r_\mu^2$. Therefore

$$\mu = \sqrt{\pi} \frac{e^2 N_D^{1/2}}{\varkappa} \left(\frac{N_A}{N_D - n}\right)^{1/2}. \tag{23}$$

This relation is valid if $N_D - n \ll N_A$. The simulation results agree well with (22) and (23). The reasoning that leads to the asymptotic expressions (22) and (23) offer in fact an exhaustive explanation of the behavior of the $\mu(n)$ curves at $K \geqslant 0.35$.

The most important feature of the curves of Fig. 3 is that at $K \leqslant 0.35$ the $\mu(n)$ dependence has a maximum. According to (1) this means that the SD G_c becomes infinite, and then becomes negative. Let us examine this phenomenon at $K = 0$. As seen from Fig. 3, at small n the $\mu(n)$ plot is adequately described by Eq. (23), but then the plot reaches a maximum and subsequently $\mu \to 0$ from positive values as $N_D - n \to 0$. We shall show below that this phenomenon is due to Wigner crystallization of the holes.[2)]

As $n \to N_D$, the spatial positions of the empty donors (holes) are determined not by the acceptors, of which there are none in the considered case $K=0$, but only by the mutual repulsion of the holes. If the donor arrangement were ordered, then the holes should form in the ground state a Wigner crystal. We are considering randomly disposed donors, but as $n \to N_D$ the hole density tends to zero, becoming much smaller than the donor density. In this situation the disorder of the donors is not very important, since the discreteness of the possible hole positions is much less than the average distance between holes. It is clear therefore that in the ground state the hole arrangement should have a short-range order of the same type as in an ordinary classical Wigner crystal. The only problem is that the deviations from an ideal periodic structure, which are due to fluctuations of the donor positions, can accumulate over distances that are large compared with the period of the Wigner crystal, and can lead to phase randomization. This question is discussed in Appendix 1. The arguments advanced there indicate that an arbitrarily weak disorder of this kind does indeed lead to phase randomization if the dimensionality of space is less than four. In the two-dimensional case of interest to us the phase randomization takes place over a length (see 1.1) $R \sim l(N_D l^2) \gg l$, where $l = (N_D - n)^{-1/2}$ is the period of the Wigner crystal.

Phase randomization over a large distance has very little effect on the binding energy of the Wigner crystal and can be calculated in the same manner as in the ideal case. Assume that the distance a to the metal surface is much larger than the crystal period. We have then the Wigner problem of electron crystallization on a homogeneous positive compensating background. The binding energy is of the form

$$H_0 = -\frac{\alpha}{2}\frac{e^2}{\varkappa}(N_D - n)^{3/2} S. \tag{24}$$

It is known that the minimum energy is possessed by a triangular lattice (see Ref. 13) for which the coefficient $\alpha = 3.92$. For a square lattice $\alpha = 3.90$ and for a hexagonal, $\alpha = 3.78$. Using (9), we obtain

$$\mu = \frac{3\alpha}{4}\frac{e^2}{\varkappa}(N_D - n)^{1/2}. \tag{25}$$

As $n \to N_D$ the chemical potential tends thus to zero from the positive direction, and this explains the result shown in Fig. 3.

At $l \gg a$ the result should be a "dipole" Wigner crystal with a binding energy smaller by a factor $(a/l)^2$ than obtained from (24). We need retain in (4) only the terms that describe the interaction of the charges with their own images. We then obtain from (4) and (8)

$$\mu = 4\pi e^2 \varkappa^{-1} a (N_D - n) + e^2/4\varkappa a. \tag{26}$$

The computer experiment yields values of $\mu(n)$ that agree with (25) and (26) in the corresponding regions. We were unable, however, to obtain an exact numerical value of α, since the computer facilities were inadequate to satisfy simultaneously the conditions $(S_0)^{1/2} \gg a \gg l \gg N_D^{-1/2}$ necessary for the existence of a "good" crystal. At $N=1000$, $a=10$, $l \sim 10$, and $N_D = 1$ we obtained $\alpha = 3.86$.

In the considered impurity-band model the electron is regarded as localized on a donor and its zero-point oscillations have low amplitude. It is this which allows the existence of the dipole crystal, which is impossible in the case of free electrons in the limit when their density is low. For the same reason the melting temperature of such a crystal should be higher than that of an ordinary two-dimensional Wigner crystal. We note that we have used above the definition (9) for μ. The definition of μ as the energy corresponding to half the distance between the empty minimum-energy level and the filled maximum-energy level yields results that agree with the first definition everywhere except in the region of the Wigner crystal. In the latter region these definitions lead to substantially different results. The reason is that in the Wigner-crystal region the filled and empty states are separated by a rigid gap.

We consider now the case of nonzero but small degrees of compensation, $K = N_A/N_D < 0.35$. In the region $N_D - n \ll N_A$, all the holes are near acceptors and formula (23) is valid. In the region $N_A \ll N_D - n \ll N_D$ the holes for which there were not enough acceptors form a Wigner crystal and formula (25) is valid. Finally, at n comparable with N_D the Wigner crystal vanishes (just as at $K=0$). The result is a minimum on the $\mu(n)$ curve at a density n close to $N_D - N_A$ (see curves 2 and 3 in Fig. 3).

We have discussed so far results of computer experiments with a zero non-Coulomb energy scatter ($\Phi_i \equiv 0$). It is clear that at a large scatter $A \gg e^2\varkappa^{-1}N_D^{1/2}$ one can neglect in (4) all terms but the first. Then, just as in the derivation of (12), $G_c = g_0(\mu)$. Thus, with increasing A the section with negative G_c should vanish, just as when the degree of compensation is increased. Figure 4 shows the results for $K=0$ and $\Phi_i \neq 0$, which illustrate this situation.

To conclude this section we present a plot of the single-electron DS $g(\varepsilon)$ obtained with the aid of our computer experiments (Fig. 5), and emphasize once more that

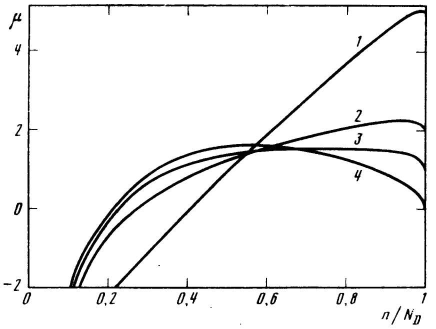


FIG. 4. Dependence of μ (in units of $e^2N_D^{1/2}/\varkappa$) on n/N_D at $K=0$ and at values of A respectively 5, 2, 1, and 0 (curves 1, 2, 3, and 4).

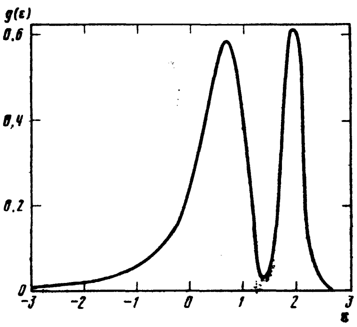


FIG. 5. Single-electron state density g (in units of $N_D^{1/2}\varkappa e^{-2}$) vs. the energy ε (in units of $e^2 N_D^{1/2}/\varkappa$) at $n = 0.7\,N_D$, $c = 10N_D^{-1/2}$, and $K = 0$.

the DC G_c has nothing in common with $g(\varepsilon)$. An exception is the case of a large non-Coulomb scatter of the levels, $A \gg e^2\varkappa N_D^{1/2}$. In this case G_c coincides with $g_0(\mu)$, and $g(\varepsilon)$ differs from g_0 only in a small vicinity of the Fermi level[8] while allowance for the Coulomb interaction at $g_0 < \varkappa a^{-1} e^{-2}$ leads to no change whatever in $g(\varepsilon)$.

4. CONCLUSION

The curves in Fig. 3 yield the dependence of the capacitive DS $G_c = (d\mu/dn)^{-1}$ on the electron density in the inversion layer. A detailed experimental study of these dependences can yield the impurity-band compensation (which serves here as a parameter) or an estimate of the width A of the donor-energy scatter of non-Coulomb origin; this scatter can be due to the different distances between the donors and the semiconductor surface. The most interesting property of these curves is the negative DS in the region of the Wigner crystal.

The state density G_{act}, strictly speaking, cannot be obtained with the aid of the performed calculations. According to (2) and (3), besides the relation $\mu(n)$ determined above, we must find also the mobility threshold $\varepsilon_p(n)$. It is easy to verify that the potential relief produced in the conduction band is classical at low density of the impurity states. Indeed, the probability of tunneling through a typical relief hump of height $\gamma = e^2 N_D \varkappa^{-1}$ and of width $N_D^{-1/2}$ is equal to $W = e^{-Q}$, where $Q = 2\hbar^{-1} N_D^{-1/2}(2\gamma m)^{1/2} \approx (N_D a_B^2)^{-1/4}$ and a_B is the radius of the impurity state. At sufficiently low N_D, the probability W is much less than unity. Therefore the potential relief of the bottom of the conduction band can be regarded as classical. The mobility threshold then coincides with the classical percolation level (see Chap. 5 of Ref. 6). To find it we must calculate the potential energy of the electron produced at the bottom of the conduction band by the charged impurities, and solve the continual percolation problem (see Ref. 6). We present here only the calculations of the percolation level in the Wigner-crystal region, which can be performed analytically.

It must be borne in mind that when the charged donors are ordered, narrow allowed bands are produced below the percolation level. These bands, generally speaking, lead to a spreading of the wave packet, and hence contribute to the conductivity. If, however, the Wigner-crystal period is large enough, the mobility in these bands is low and it can be assumed that the conduction is effected only by the electrons of energy higher than the percolation level.

TABLE I. Values of the coefficients α, η, γ, and β for three lattices.

Lattice	α	η	γ	β
Triangular	3.92 *	1.43	1.31	1,63
Quadratic	3,90 *	1.62	1.14	1,78
Hexagonal	3,78	2,03	0,68	2,16

*Data of Ref. 13.

In the calculation of the percolation level, the potential can be regarded as static, inasmuch as over times typical of band conductivity there are not transitions of electrons between donors. The calculation of the percolation level is described in Appendix 2. We have obtained

$$\varepsilon_p = \gamma \frac{e^2}{\varkappa}(N_D - n)^{1/2}, \tag{27}$$

where the constant γ is expressed in terms of the constant α of (24). Accordingly

$$\varepsilon_a = E_0 + \varepsilon_p - \mu = E_0 - \beta \frac{e^2}{\varkappa}(N_D - n)^{1/2}, \tag{28}$$

where $\beta = 3\alpha/4 - \gamma$. According to (2)

$$G_{act}^{-1} = -\frac{\beta}{2}\frac{e^2}{\varkappa}(N_D - n)^{-1/2}. \tag{29}$$

The calculation results for the percolation level are listed in Table I. It is seen from it that $\beta > 0$, so that the SD determined by measuring the change of the activation energy of the band conductivity with change of electron density is also negative in the Wigner-crystal region. This means that at low degree of compensation the activation energy first decreases with electron density, and then increases if the hole density is low. It seems to us that the impurity band produced on the silicon surface by sodium ions (see Ref. 1) is a suitable object for the observation of this effect. Unfortunately, ε_a was measured in Ref. 1 at only one value of V_g corresponding to the maximum current.

An interesting question from the theoretical viewpoint is the low-temperature electric conductivity that is not due to spilling into the conduction band. At intermediate degrees of compensation it should have a hopping character without any special features; in the Wigner-crystal region, however, the hole motion should more readily be correlated. We shall not deal here with this question.

The authors take the opportunity to thank M. Feĭgel'man and D. E. Khmel'nitskiĭ for a discussion of the stability of a Wigner crystal.

APPENDIX 1: INSTABILITY OF A LOCALIZED-HOLE WIGNER CRYSTAL TO WEAK DISORDER OF THE LOCALIZATION CENTERS

We consider the case when the total donor density N_D is very high compared with the density $p \equiv N_D - n$ of the ionized donors (holes), but such that the average distance $r_D = N_D^{-1/2}$ between donors is much larger than the radius of the localization of the electron on the donor. The problem remains classical in the sense indicated in the Introduction. The potential produced by the polarization of the neutral donors will be neglected, since the total energy includes only the interaction between holes that can be located on any of the donors, as well as the interaction of the holes with the homogeneous negative background. It is required to find the hole configuration corresponding to the ground state of the system. If the holes could occupy arbitrary points in space, then the ground state would correspond to the ordered structure called a Wigner or Coulomb crystal. The hole can be located only on donors, but the donor density is very high, so that the disorder due to the random distribution of the donor is very small. Undoubtedly, therefore, at distances large compared with the average distance between holes, an ordered structure should exist. The question is whether this structure has an infinite correlation radius, or does the disorder lead at large distances to phase randomization, so that strictly speaking there is no crystal.

We consider now an ideal lattice with a number of lattice site per unit volume corresponding to the given hole density p, and assume that it penetrates through a space (of arbitrary dimensionality d) in which donors are randomly distributed. We examine the donor configuration that results from placing holes on donors that are closest to the sites of the ideal lattice. This configuration corresponds to an infinite correlation radius. Owing to the disorder, the peaks of the correlation function have finite width, but this width is small compared with the distances between the peaks and, most importantly, does not increase with the number of the peak. We now must determine whether this structure corresponds to the ground state.

For this purpose, we reason as follows. Assume that the considered ideal lattice is not infinite, but forms a hypercube with linear dimension $R \gg l \equiv p^{1/2}$. We calculate the electrostatic energy of the interaction of the holes located near a lattice site. We then displace our lattice relative to the donor distribution specified in the space, move the holes to the donors closest to the lattice sites, and calculate the energy anew. The ideal crystal corresponds to the minimum energy. Therefore a displacement of the holes by a distance on the order of r_D from the sites of the ideal lattice leads to an energy increase of second order in the parameter r_D/l. The characteristic energy increase per cell should be of the order of

$$\frac{e^2}{l}\left(\frac{r_D}{l}\right)^2,$$

and the average increase of the hypercube energy of the order of

$$\frac{e^2}{l}\left(\frac{r_D}{l}\right)^2\left(\frac{R}{l}\right)^d.$$

Since the donor configuration in the vicinity of each lattice site can be regarded as statistically independent, the dispersion of the hypercube energy is of the order

$$\frac{e^2}{l}\left(\frac{r_D}{l}\right)^2\left(\frac{R}{l}\right)^{d/2}.$$

This is the amount by which the hypercube energy can change (on the average) following the indicated shift. It follows therefore that if a hole crystal is placed in a region of space with linear dimension R not in the manner imposed by the neighboring regions, but by subjecting it to a translation of order l such that its position is most favorable at the given donor configuration, then the resultant energy gain is of the order of

$$\frac{e^2}{l}\left(\frac{r_D}{l}\right)^2\left(\frac{R}{l}\right)^{d/2}.$$

This, however, corresponds to a complete phase randomization that takes place over a length R.

If this randomization is effected jumpwise in the hyperplane that separates regions of size R, then this obviously leads to a strain-energy loss on the order of

$$\left(\frac{e^2}{l}\right)\left(\frac{R}{l}\right)^{d-1}.$$

The loss can be much smaller if the phase randomization proceeds via a smooth deformation of the Wigner crystal. In this case the deformation should not change the volume, for otherwise the negative background will not compensate for the charge of the crystal and the electrostatic energy will increase greatly. To ensure phase randomization over the distance R, the shear components of the strain tensor must be of the order l/R. The strain energy per crystal cell is of the order of

$$\frac{e^2}{l}\left(\frac{l}{R}\right)^2,$$

and the total energy loss is of the order of

$$\frac{e^2}{l}\left(\frac{l}{R}\right)^2\left(\frac{R}{l}\right)^d.$$

Comparing the energy gain and loss, we find that at $d < 4$ there should be no ideal Wigner crystal, and the correlation radius R_c (obtained by equating the gain and the loss) is of the form

$$R_c \sim l\left(\frac{l}{r_D}\right)^{4/(4-d)} \quad (d<4). \tag{1.1}$$

The correlation function should vanish at $R \gg R_c$ if $d < 4$.

APPENDIX 2: PERCOLATION LEVEL IN CONDUCTION BAND FOR ORDERED CHARGED DONORS

We consider a case when the holes form a Wigner crystal (assumed to be ideal), and calculate the percolation level in the conduction band in the potential produced by the charged lattice sites. We propose below to calculate the percolation levels for a quadratic, triangular, and hexagonal lattice.

1. *Quadratic lattice.* The equipotential lines for an electron in the conduction band are shown in Fig. 6a. The points 1 are lattice sites (holes) and correspond to a minimum of the electron potential energy in the conduction band, while points 2 are the maxima of the potential energy. The only nonclosed equipotential

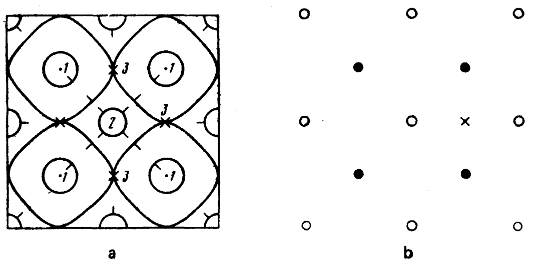


FIG. 6. Illustrating the calculation of the percolation level in a quadratic lattice.

line corresponds to the percolation level. To determine this level it suffices therefore to find the energy at the point 3. According to (24) the potential energy at the point 1 is $\varphi_1(p)=\alpha e^2\varkappa^{-1}p^{1/2}$, where p is the hole density. We represent the energies at the points 2 and 3 in the form

$$\varphi_2(p)=\eta e^2\varkappa^{-1}p^{1/2},\quad \varphi_3(p)=\gamma e^2\varkappa^{-1}p^{1/2}. \tag{2.1}$$

The problem is to express η and γ in terms of the coefficient α of Eq. (24).

We regard the quadratic lattice with density $2p$ (Fig. 6b) as a superposition of two quadratic lattices with densities p (their sites are shown by dark and light circles respectively). Consider a black site on Fig. 6b. It is a point of type 1 with respect to the black sublattice, of type 2 with respect to the white, and of type 1 with respect to the combined one. The energy at this site is equal, on the one hand, to $\varphi_1(2p)$ and on the other to the sum of the potential energies produced by the individual sublattice, i.e., $\varphi_1(p)+\varphi_2(p)$. Using (2.1), we get

$$\alpha e^2\varkappa^{-1}(2p)^{1/2}=\alpha e^2\varkappa^{-1}p^{1/2}+\eta e^2\varkappa^{-1}p^{1/2}, \tag{2.2}$$

from which it follows that

$$\eta=(\sqrt{2}-1)\alpha. \tag{2.3}$$

We consider now the energy marked by the cross in Fig. 6b. For both sublattices it is a point of type 2, but for their sum it is of type 2. Thus,

$$\varphi_2(2p)=2\varphi_3(p),\quad \gamma=\frac{\eta}{\sqrt{2}}=\frac{\sqrt{2}-1}{\sqrt{2}}\alpha. \tag{2.4}$$

2. *Triangular and hexagonal lattices*. Figure 7a shows the unit cell of a triangular lattice, and the characteristic points are designated in the same manner as for the quadratic lattice. Figure 7b shows the subdivision of a triangular lattice with density $3p$ into two sublattices. The energy per site is

$$\begin{gathered}\varphi_1(3p)=\varphi_1(p)+2\varphi_2(p),\\ \eta=(\sqrt{3}-1)\alpha/2.\end{gathered} \tag{2.5}$$

We can write analogously with the aid of Fig. 7c

$$\begin{gathered}\varphi_1(4p)=\varphi_1(p)+3\varphi_3(p),\\ \gamma=\alpha/3.\end{gathered} \tag{2.6}$$

The result for a hexagonal lattice can be obtained in similar fashion. Moreover, since the hexagonal lattice can be represented as a superposition of two triangular ones, the coefficient α for it (α^6) can be expressed in terms of the coefficient α for the triangular lattice

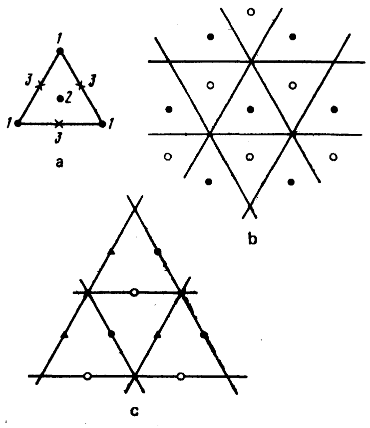


FIG. 7. Illustrating the calculation of the percolation level in a triangular lattice.

(α^3). We present here only the final results:

$$\begin{aligned}\alpha_6&=(\sqrt{3}+1)\alpha_3/2\sqrt{2},\\ \eta_6&=(\sqrt{3}-1)^2\alpha_6,\\ \gamma_6&=\eta_6/3.\end{aligned} \tag{2.7}$$

Using the results of Ref. 13, in which the coefficients α were calculated for a triangular and quadric lattice, we can calculate all the coefficients of interest to us. They are listed in Table I.

[1)] As shown in Ref. 8, at the values of N employed here the properties of the pseudo-ground states hardly differ from those of the ground state.

[2)] The possibility of Wigner crystallization of localized carriers in the three-dimensional case was first mentioned by Batyev.[10] Wigner crystallization of localized carriers in an MIS structure and the ensuing consequences were reported by Chenskiĭ and Tkach[11] and by us.[12]

---

[1] A. Harstein and A. B. Fowler, Surface Science 73, 19 (1978).

[2] R. A. Suris, Materials of the 7th Winter School of Physicotech. Inst. on Semiconductor Physics, Leningrad, 1975, pp. 245–285. M. Pepper, Contemp. Phys. 18, 423 (1977).

[3] C. J. Adkins, J. Phys. C: Solid St. Phys. 11, 851 (1978).

[4] A. L. Éfros and B. I. Shklovskiĭ, *ibid.* 8, L49 (1975).

[5] A. L. Efros, *ibid.* 9, 2021 (1976).

[6] B. I. Shklovskiĭ and A. L. Éfros, Élektronnye svoistva legirovannykh poluprovodnikov (Electronic Properties of Doped Semiconductors), Nauka, 1979.

[7] A. L. Éfros, Nguyen Van Lien, and B. I. Shklovskiĭ, J. Phys. C: Solid St. Phys. 12, 1869 (1979).

[8] S. D. Baranovskii, A. L. Éfros, B. L. Gelmont, and B. I. Shklovskiĭ, J. Phys. C: Solid St. Phys. 12, 1023 (1979).

[9] V. A. Gergel' and R. A. Suris, Zh. Eksp. Teor. Fiz. 75, 192 (1978) [Sov. Phys. JETP 48, 95 (1978)].

[10] E. G. Batyev, Pis'ma Zh. Eksp. Teor. Fiz. 29, 560 (1979) [JETP Lett. 29, 510 (1979)].

[11] E. V. Chenskiĭ, and Yu. A. Tkach, Tenth Conf. on Semiconductor Theory, Novosibirsk, 1980, Abstracts, Part 2, p. 181.

[12] M. S. Bello, E. I. Levin, B. I. Shklovskiĭ, and A. L. Éfros, *ibid.* part 1, p. 59.

[13] G. Messner, H. Namaizawa, and M. Voss, Phys. Rev. B13, 1370 (1976).

Note added in proof (3 March 1981). Chenskiĭ and Tkach have recently published a paper (Sov. Phys. JETP 52, 915 (1980) containing qualitatively similar results.

Translated by J. G. Adashko