PHYSICAL REVIEW B VOLUME 26, NUMBER 12 15 DECEMBER 1982

Influence of charged impurities on Si inversion-layer electrons

B. Vinter*
Physik Department, Technische Universität München, D-8046 Garching, Federal Republic of Germany
(Received 8 February 1982)

A theoretical investigation of various effects of charged impurities in the oxide on electrons in an inversion layer in the SiO_2-Si(100) metal-oxide-semiconductor structure has been carried out in the quantum limit. For very low electron density we determine ground- and excited-bound-state energies and wave functions exactly and show their dependence on surface field and impurity distance from the interface. In the high-density limit we treat screening and many-body effects by applying the local-density-functional method. The self-consistent potential gives very good agreement with experimentally determined scattering rates and has one fourfold-occupied bound state. The degeneracy of the bound state is not lifted, even if spin- or valley-density-functional methods are applied. For interfaces under stress, where another subband can be occupied, we study the effect of an impurity on the two-component electron gas. While direct comparison with experiments is not possible, the results seem to indicate that in order to understand mobility measurements under stress, scattering mechanisms additional to those present without stress have to be taken into account. The calculations show that there is only a bound state associated with subband 0′ when no states associated with subband 0 are occupied.

I. INTRODUCTION

In a metal-oxide-semiconductor (MOS) structure one forms a channel of carriers in the semiconductor very close to the oxide-semiconductor interface when a voltage is applied between metal and semiconductor. An important technological and fundamental problem is to understand the mobility of the carriers in the channel. In n-channel inversion layers on Si surface roughness, acoustic phonon and impurity scattering are thought to determine the low-field mobility. At low temperatures, where the phonon scattering can be neglected, impurity scattering is dominant for low inversion-layer densities, whereas surface-roughness scattering increases with increasing density, so that the mobility shows the characteristic mobility peak[1] for N_{inv} somewhere between 10^{12} and 10^{13} cm^{-2}.

A major step forward in the experimental study of the mobility has been the technique of drifting Na^+ ions through the oxide to the oxide-semiconductor interface in a controlled manner.[2,3] By variation of oxide charge, substrate bias, and temperature it is then possible to separate the contributions to the scattering rates from the three mechanisms.[4] Of course, this allows a much better quantitative test of theoretical calculations of scattering rates. Apart from giving rise to impurity scattering the Na^+ ions show other interesting properties: Near the interface they can form bound states occupied by inversion-layer electrons; for sufficiently high impurity concentrations an impurity band can be formed and its conductivity by hopping can be studied.[5] In the last few years investigations on the impurity bound states by infrared absorption have also appeared.[6–8]

In this paper we study theoretically, in detail, the influence of such charged oxide impurities on electrons in the inversion layer at low temperature. Much theoretical work on this problem has been done by several authors in various approximations and models.[9–19] The general goal of our work has been to resort to numerical methods to do as realistic a calculation as seems possible with all nonessential approximations eliminated. This has been achieved for two general situations: First, the number of electrons in the inversion layer is much smaller than the number of impurities. In this case the impurities cannot be screened by the electrons, and wave functions and binding energies can be determined by solving the Schrödinger equation for one electron with a given potential. We show in Sec. II how we have done this, and in Sec. III results are presented on wave functions and energies of the ground state and several excited states as a function of impurity distance from the interface and surface field. Our results on the ground state do not differ much from those of simpler theories,[12] but most of the excited states have not been described before, and we follow the crossover behavior from perturbed hydrogeniclike states to perturbed subband states as the depletion-layer po-

tential increases relative to the impurity potential. Second, we consider the situation in which the number of inversion-layer electrons is much larger than the number of impurities. In this case we must take into account the interactions between the electrons. It is well known[20–22] that many-body effects play an important role even without impurities, so that these should be incorporated in a theory of the screening of the impurity potential. Earlier theories have invariably used linear screening[9,15–19] and, in most cases, a Born approximation for the scattering rate. Here we use the local-density-functional method to calculate self-consistently the screened potential. Furthermore, we make a complete phase-shift analysis of the scattering of electrons on the screened potential. The method is described in Sec. IV, and in Sec. V we show that the mobility calculated in this way agrees considerably better with the experimental measurement of Hartstein *et al.*[4] than earlier theories do.

Another result of the method is one bound state lying less than 1 meV below the lowest subband bottom. This state is fourfold degenerate like all the subband states, and to investigate this somewhat surprising result we have carried out extended calculations which, in principle, allow lifting of the degeneracy. These valley-density-functional methods are described in Sec. VI, but the results only substantiate the fourfold degeneracy. Similar methods can be used to investigate the influence of an impurity on inversion-layer electrons which, by application of uniaxial strain, occupy two subbands originating from different valleys of bulk Si. In Sec. VII we present results for this case; no directly comparable experiment exists, but it seems that additional scattering mechanisms come into play when two subbands are close in energy. In the concluding section we also include remarks on some further calculations meant to illuminate the fourfold occupancy of the bound state.

Before we go on to the main part of the paper let us mention the approximations we have made. If at all possible, elimination of these would require much more difficult calculations and better knowledge of the detailed atomic structure of the interface. We use in all cases the effective mass approximation and represent the interface by an infinitely steep barrier; all electrons other than the inversion-layer electrons are accounted for by the macroscopic dielectric constants of the media; the local-density-functional formalisms are used uncritically by taking the eigenvalues of the effective Schrödinger equation to represent one-particle excitation energies of the system; the distance between impurities is assumed to be so large that multiple scattering can be neglected, and the impurity is taken to be a point charge in the oxide. These are orthodox approximations in most inversion-layer theories, and we see no reason why they should be more radical for the problem at hand.

II. SOLUTION WITHOUT SCREENING

We assume a positive ion of charge Ze to be situated at the distance z_0 from the interface $z=0$. In addition, the depletion charge provides a depletion field which can be taken to be constant near the interface. Furthermore, an electron in the semiconductor sees its own image. The system has cylindrical symmetry around the interface normal through the impurity which we choose to be the z axis. Our one-electron Hamiltonian is then the following:

$$H=-\frac{\hbar^2}{2m_t}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]-\frac{\hbar^2}{2m_l}\frac{\partial^2}{\partial z^2}+V_{\text{depl}}(z)+V_{\text{im}}(z)+U_{\text{imp}}(R,z) \tag{1}$$

where $m_t=0.1905m_e$ and $m_l=0.916m_e$ are the transverse and longitudinal masses of Si;

$$V_{\text{depl}}=e^2N_{\text{depl}}z/\kappa_s\epsilon_0\ ,$$

where N_{depl} is the depletion charge and $\kappa_s=11.7$, the relative permittivity of Si, is the depletion potential;

$$V_{\text{im}}=(\kappa_s-\kappa_i)/16\pi\kappa_s\epsilon_0(\kappa_s+\kappa_i)z$$

is the image potential with the relative permittivity of the insulator κ_i taken to be 3.9; and

$$U_{\text{imp}}(R,z)=-Ze^2/4\pi\bar{\kappa}\epsilon_0[R^2+(z+z_0)^2]^{1/2}$$

is the impurity potential, where $\vec{\mathrm{R}}=(x,y)$ and $\bar{\kappa}=(\kappa_s+\kappa_i)/2$. In addition, the interface to the insulator is assumed to provide an infinite potential barrier, so that the wave functions vanish at $z=0$.

It is convenient to introduce scaled atomic units in which $\hbar=2m_t=e^2/8\pi\epsilon_0\bar{\kappa}=1$; the unit of length is the effective Bohr radius $a_0^*=a_0\bar{\kappa}(m_e/m_t)$ $=21.7$ Å; and the unit of energy is the effective Rydberg $\mathcal{R}^*=\mathcal{R}(m_t/m_e)/\bar{\kappa}^2=42.6$ meV. In cylindrical coordinates (R,θ,z) we then have $H=H_0+H_1$, where

$$H_0(z)=-\frac{m_t}{m_l}\frac{\partial^2}{\partial z^2}+V_{\text{depl}}(z)+V_{\text{im}}(z)\ ,$$
$$H_1=-\frac{\partial^2}{\partial R^2}-\frac{1}{R}\frac{\partial}{\partial R}-\frac{1}{R^2}\frac{\partial^2}{\partial\theta^2}+U(R,z)\ , \tag{2}$$

with

$$U(R,z)=-2Z/(R^2+(z+z_0)^2)^{1/2}\ .$$

In order to solve the Schrödinger equation $H\psi=E\psi$ we first take advantage of the cylindrical symmetry to classify the solutions according to the "magnetic" quantum number

$$m=0,\pm1,\pm2,\ldots:\ \psi(R,\theta,z)$$
$$=e^{im\theta}\varphi_m(R,z)\ .$$

Furthermore, solving the one-dimensional Schrödinger equation corresponding to H_0 we get the familiar complete set of subband wave functions $H_0\zeta_n(z)=E_n\zeta_n(z)$. We use this set to expand the z dependence of the function $\varphi_m(R,z)$, i.e.,

$$\psi_m(R,\theta,z)=e^{im\theta}\sum_{n=0}^{\infty}f_n^{(m)}(R)\zeta_n(z)\ . \tag{3}$$

Inserting this expansion in the Schrödinger equation and taking the inner product with $\zeta_N(z)$ we obtain an infinite set of coupled ordinary differential equations for the radial parts:

$$\left[\frac{d^2}{dR^2}+\frac{1}{R}\frac{d}{dR}+\left[(E-E_N)-\frac{m^2}{R^2}\right]\right]f_N(R)$$
$$=\sum_{n=0}^{\infty}U_{Nn}(R)f_n(R) \tag{4}$$

for $N=0, 1, 2, \ldots,$ and with

$$U_{Nn}(R)=\int\zeta_N^*(z)U(R,z)\zeta_n(z)dz\ . \tag{5}$$

We have dropped the superscript m in $f_n^{(m)}(R)$.

If we consider the undisturbed system, i.e., $U(R,z)=0$, the differential equations are decoupled and are of Bessel type. Thus the complete solution for equation N is

$$f_N(R)=\begin{cases}CJ_m(k_NR)+SN_m(k_NR) & \text{for } E>E_N \quad (6)\\ CI_m(k_NR)+SK_m(k_NR) & \text{for } E\le E_N \quad (7)\end{cases}$$

where $k_N=|E-E_N|^{1/2}$, C and S are arbitrary constants, and J, N, I, and K are the standard Bessel functions of integer order m. A permissible wave function can only be found for $E>E_N$ and $S=0$, since the other Bessel functions have singularities at 0 or ∞, and we recover the inversion layer wave functions written in cylindrical coordinates

$$\psi_{mN}(R,\theta,z)=C_Ne^{im\theta}J_m(k_NR)\zeta_N(z) \tag{8}$$

with the energy $E=E_N+k_N^2$.

Suppose we search for a solution to Eqs. (4) and (5) representing a bound state, i.e., a state with $E<E_0$. The Green's function $G_N(R,R',E)$ corresponding to the Nth equation

$$\left[\frac{d^2}{dR^2}+\frac{1}{R}\frac{d}{dR}+\left[(E-E_N)-\frac{m^2}{R^2}\right]\right]G_N(R,R',E)$$
$$=\frac{\delta(R-R')}{R} \tag{9}$$

is then given by

$$G_N(R,R',E)=\begin{cases}-K_m(k_NR')I_m(k_NR)\ , & R<R'\\ -I_m(k_NR')K_m(k_NR)\ , & R'<R\end{cases} \tag{10}$$

so that we can write the solution as

$$f_N(R)=\int_0^{\infty}R'dR'G_N(R,R',E)\sum_n U_{Nn}(R')f_n(R')$$
$$=-K_m(k_NR)\int_0^{R}R'dR'I_m(k_NR')\sum_n U_{Nn}(R')f_n(R')-I_m(k_NR)\int_R^{\infty}R'dR'K_m(k_NR')$$
$$\times\sum_n U_{Nn}(R')f_n(R')\ . \tag{11}$$

If we express the solution in the form analogous to (7)

$$f_N(R)=C_N(R)I_m(k_NR)+S_N(R)K_n(k_NR)\ , \tag{12}$$

we have

$$C_N(R)=-\int_R^{\infty}R'dR'K_m(k_NR')\sum_n U_{Nn}(R')f_n(R')\ ,$$
$$S_N(R)=-\int_0^{R}R'dR'I_m(k_NR')\sum_n U_{Nn}(R')f_n(R') \tag{13}$$

or

$$\frac{dC_N}{dR}=RK_m(k_NR)\sum_n U_{Nn}(R)f_n(R)\ ,$$
$$\frac{dS_N}{dR}=-RI_m(k_NR)\sum_n U_{Nn}(R)f_n(R)\ . \tag{14}$$

The boundary conditions are $C_N(\infty)=S_N(0)=0$ to obtain a normalizable wave function. Furthermore, the solution must be continuous and have a continuous derivative, which is equivalent to having $C_N(R)$ and $S_N(R)$ continuous everywhere. All these conditions can only be fulfilled for special values of E, so that the energy spectrum below E_0 is discrete.

In practice we have to make the controllable approximation of truncating the system of equations. For a given energy we first choose $C_n(0)=\delta_{nn'}$ and $S_n(0)=0$ and integrate Eq. (14) from $R=0$ to some intermediate point R_0 with the results $C_{nn'}(R_0-)$ and $S_{nn'}(R_0-)$. Similarly, we choose $S_n(\infty)=\delta_{nn'}$ and $C_n(\infty)=0$ and integrate from $R=\infty$ to R_0 with the results $C_{nn'}(R_0+)$ and $S_{nn'}(R_0+)$. This is done for all n'. The condition of continuity at R_0 now requires that we choose $C_n(0)$ and $S_n(\infty)$ so that for all n'

$$\sum_n C_{nn'}(R_0-)C_n(0)=\sum_n C_{nn'}(R_0+)S_n(\infty)\ ,$$
$$\sum_n S_{nn'}(R_0-)C_n(0)=\sum_n S_{nn'}(R_0+)S_n(\infty)\ . \tag{15}$$

This is a set of homogeneous linear equations to determine $C_n(0)$ and $S_n(\infty)$. It only has a nonzero solution when the determinant, which is a function of the energy, vanishes. We therefore find the energy of a bound state by iterating the procedure until a zero of the determinant has been found. Then the boundary values $C_n(0)$ and $S_n(0)$, and thereby the total wave function, are determined from (15) and from normalization.

There are two sources of inaccuracy in this method. First, we must truncate the system of equations to a finite number; this inaccuracy can be checked by trying different numbers of subbands. Second, we do not integrate to infinity but only to a finite $R_{\max}$. This corresponds to setting the potential equal to zero for $R>R_{\max}$ and plays no role for states that are tightly bound inside $R_{\max}$. For the long-range Coulomb potential one would expect infinitely many bound states and the very shallow states cannot be described correctly for a finite $R_{\max}$. Trying to determine these would be an academic exercise which does not interest us here. The states are not lost; they are pushed upwards in energy into the continuum when we cut off the potential at $R_{\max}$.

For energies between E_0 and E_1 the Green's function $G_0(R,R',E)$ defined by (9) is not given by (10) but by

$$G_0(R,R',E)=\begin{cases}\dfrac{\pi}{2}N_m(k_0R')J_m(k_0R), & R<R'\\ \dfrac{\pi}{2}J_m(k_0R')N_m(k_0R), & R'<R\ ,\end{cases} \tag{16}$$

where $k_0=(E-E_0)^{1/2}$, so that for $f_0(R)$ we get

$$f_0(R)=\frac{\pi}{2}N_m(k_0R)\int_0^R R'dR'J_m(k_0R')\sum_n U_{0n}(R)f_n(R)$$
$$+\frac{\pi}{2}J_m(k_0R)\int_R^\infty R'dR'N_m(k_0R')\sum_n U_{0n}(R)f_n(R)+C_\infty J_m(k_0R)\ . \tag{17}$$

By analogy with Eqs. (12)–(14) we write

$$f_0(R)=C_0(R)J_m(k_0R)+S_0(R)N_m(k_0R) \tag{18}$$

and obtain

$$\frac{dC_0}{dR}=-\frac{\pi}{2}RN_m(k_0R)\sum_n U_{0n}(R)f_n(R)\ ,$$
$$\frac{dS_0}{dR}=\frac{\pi}{2}RJ_m(k_0R)\sum_n U_{0n}(R)f_n(R)\ , \tag{19}$$

where the boundary condition is now only $S_0(0)=0$. The method of solution is the same as for bound states but, in addition, we also choose $C_N(\infty)=S_N(\infty)=0$ for all N except $C_0(\infty)=1$, and integrate from infinity to R_0 with the results $c_n(R_0+)$ and $s_n(R_0+)$. The continuity conditions at R_0 now require $C_n(0)$, $S_n(\infty)$, and $C_0(\infty)$ to fulfill for all n':

$$\sum_n C_{nn'}(R_0-)C_n(0)=\sum_n C_{nn'}(R_0+)S_n(\infty)+c_n(R_0+)C_0(\infty)\,,$$
$$\sum_n S_{nn'}(R_0-)C_n(0)=\sum_n S_{nn'}(R_0+)S_n(\infty)+s_n(R_0+)C_0(\infty)\,. \tag{20}$$

In these linear equations there is one more variable than there are equations, so there are solutions for any energy, i.e., the spectrum is continuous. The constants $C_n(0)$, $S_n(\infty)$, and $C_0(\infty)$ are determined by (20) to within a constant factor which has to be obtained from normalization of the wave function. Asymptotically we have the following:

$$f_0(R)=C_0(\infty)J_m(k_0R)+S_0(\infty)N_m(k_0R)\,,$$
$$f_n(R)=0,\quad n>0 \tag{21}$$

i.e., the undisturbed subband wave function with a phase shift

$$\eta=-\arctan S_0(\infty)/C_0(\infty)\,.$$

If we choose the normalization of the extended wave functions to be

$$\langle E'm' \mid Em\rangle=\delta_{mm'}\delta(E-E')\,,$$

we have

$$C_0(\infty)^2+S_0(\infty)^2=(4\pi)^{-1}\,,$$

so that, asymptotically, our total wave function of energy E is

$$\psi_{mE}(R,\theta,z)=(4\pi)^{-1/2}e^{im\theta}\times J_m(k_0R+\eta_m(E))\zeta_0(z)\,, \tag{22}$$

and the scattering solution is completely determined.

The standard derivation[23] yields the induced density of states in the lowest subband to be

$$\delta D(E)=\frac{g_sg_v}{\pi}\sum_{m=-\infty}^{\infty}\frac{d\eta_m}{dE}\,, \tag{23}$$

where $g_s=2$ is the spin degeneracy and g_v is the valley degeneracy ($g_v=2$ in (100) Si).

III. RESULTS WITHOUT SCREENING

In Fig. 1 we show the energy of the lowest-lying bound states for $m=0$ and $m=1$ as a function of the distance z_0 of the impurity from the interface, and with the depletion-layer density as parameter. The binding energy E_0-E increases as the electron gets closer to the impurity, either when the impurity is closer to the interface or when the depletion field presses the electron towards the interface. The energies agree fairly well with those of Martin and Wallis.[12] A direct comparison[24] showed our binding energies to be about 5% larger than theirs. That their error is not larger is, to a great extent, due to their use of a different value of the variational parameter for the z dependence of their wave function. Our binding energies agree very well with those of Lipari.[13] Obviously, his expansion is more efficient when the depletion field is small, whereas ours is more efficient when the impurity potential is small. With sufficiently many terms in the expansions, both methods are exact, of course. We have used five subbands in all the calculations presented here. For $N_{\rm depl}=2\times10^{11}$ cm^{-2} and $z_0=0$, which may be considered a bad case for our

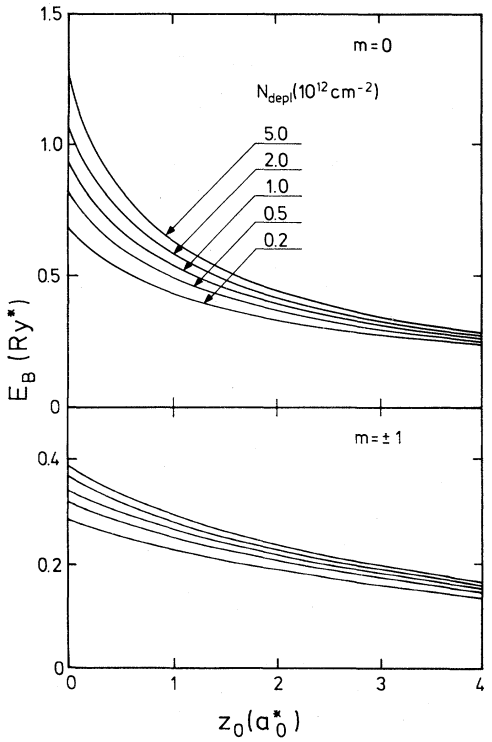

FIG. 1. Binding energy $E_B=E_0-E$ of lowest-lying bound states as a function of distance z_0 of impurity from interface for various values of depletion-layer density $N_{\rm depl}$. Upper frame: states having quantum number $m=0$; lower frame: states having $m=\pm1$. Values of $N_{\rm depl}$ are the same in lower as in upper frame.

method, the binding energy of the lowest state changes by 1%, if we use four subbands instead of five.

The interplay of the impurity and depletion potentials is illustrated in Fig. 2, where we compare the probability densities $|\psi(R,z)|^2$ of the three lowest-lying states having $m=0$ for $N_{\text{depl}}=5\times10^{11}$ cm^{-2} and for $z_0=0$ and $z_0=1$ a.u., respectively. In both cases the ground state looks qualitatively the same; it is essentially half a $2p$-like hydrogen wave function which has been squeezed closer to the surface by the depletion field. The second excited states are, qualitatively, completely different: for $z_0=1$ a.u. the wave function is essentially a product of the lowest subband wave function and a radial function which has two nodes; for the $z_0=0$ a.u. case, however, where the impurity potential is stronger, the wave function looks more like half an excited p state of hydrogen squeezed in the z direction. Finally, the first excited state in between is a mixture for $z_0=0$. For $m\neq0$ states, the wave functions are zero for $R=0$, so the influence of the impurity potential is not so strong and, effectively, the depletion field dominates.

IV. THE DENSITY-FUNCTIONAL METHOD

We now turn to the problem of the effects of an impurity when there are many electrons in the inversion layer. We assume that there are so many electrons that the impurity potential is completely screened out within the distance between impurities and, effectively, we treat just one impurity with infinitely many electrons.

Even without the impurity the electrons contribute substantially to the inversion-layer potential as first calculated by Stern and Howard[9] and Stern,[25] and furthermore, it has been demonstrated in several works[20–22] that effects of exchange and correlation also play an important role in the system. The most versatile simple approximation to treat all these effects is the local-density-functional formalism derived by Kohn and Sham[26] and first extended to the inversion layer problem by Ando.[21] In that method, the direct Hartree potential of the electrons is treated exactly, and all effects of exchange and correlation are approximated by a local potential which mainly depends on the local three-dimensional density. It is then, in principle, simple to write down the equations when the impurity is included; the effective Hamiltonian can be written analogously to Eq. (2); $H=H_0+H_1$, where

$$H_0(z)=-\frac{m_t}{m_l}\frac{\partial^2}{\partial z^2}+V_{\text{depl}}(z)+V_H^0(z)+V_{\text{xc}}^0(z)\,, \tag{24}$$

$$H_1=-\frac{\partial^2}{\partial R^2}-\frac{1}{R}\frac{\partial}{\partial R}-\frac{1}{R^2}\frac{\partial^2}{\partial\theta^2}+U(R,z)+V_H^1(R,z)+V_{\text{xc}}^1(R,z)\,, \tag{25}$$

where the new terms are the Hartree potential $V_H=V_H^0+V_H^1$ to be found from Poisson's equation and the exchange-correlation potential $V_{\text{xc}}=V_{\text{xc}}^0+V_{\text{xc}}^1$ to be derived from homogeneous electron-gas properties. We have divided these potentials, so that V_H^0 and V_{xc}^0 are the potentials if no impuri-

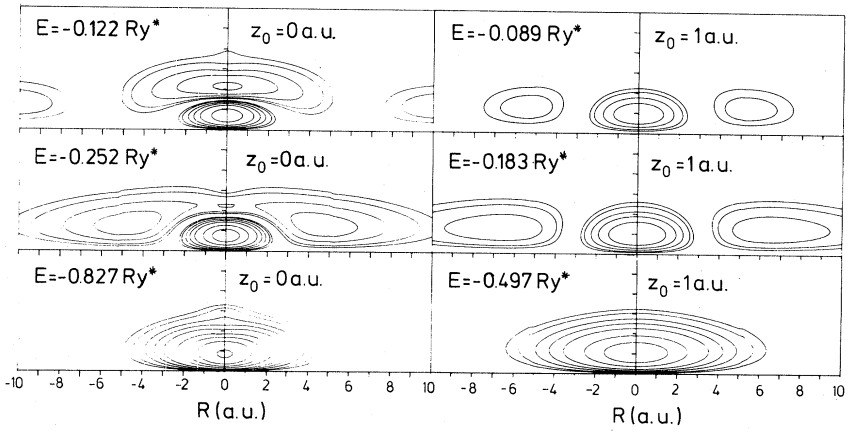

FIG. 2. Probability density plot of three lowest states having $m=0$. Left frames: distance of impurity from interface $z_0=0$. Right frames: $z_0=1$ a.u., $N_{\text{depl}}=5\times10^{11}$ cm^{-2}. The contour plots are logarithmic with three contours per decade.

ty were present, and V_H^1 and V_{xc}^1 are the changes induced by the bare impurity potential $U(R,z)$. The density-functional method requires that the equations be solved self-consistently with the ground-state density determined by

$$\rho(R,z)=\sum_{E<E_F} |\psi_E(R,\theta,z)|^2 , \tag{26}$$

where ψ_E is an eigenfunction of H with energy E, and the sum runs over all states with energy less than the Fermi energy E_F. On the other hand, the density determines the Hartree and exchange-correlation potentials as described later in this section, so the solution has to be found by iterating Eqs. (24)–(26) until the potentials and density do not change.

We see that in order to solve these equations (i) we must be able to solve the effective Schrödinger equation $H\psi_E=E\psi_E$ for all energies below E_F; (ii) when the density is given by (26) we must be able to solve Poisson's equation to determine the Hartree potential; (iii) from the density we must be able to determine the exchange-correlation potential; and (iv) we must have a converging iteration procedure to obtain self-consistency.

The first point is solved by a natural extension of what was done in Sec. II. We use (24) to obtain

$$\rho^1(R,z)=g_v g_s\left[\sum_{BS}|\psi_{Em}(R,\theta,z)|^2+\sum_m\int_0^{E_F}[|\psi_{Em}(R,\theta,z)|^2-|J_m(k_0R)\zeta_0(z)|^2/4\pi]dE\right] \tag{28}$$

where the first sum runs over all bound states, $k_0^2=E-E_0$, and the normalization defined in Eq. (22) has been used. Because of our assumption of having infinitely many electrons per impurity we have the same Fermi energy $E_F-E_0=4\pi N_{inv}/g_v g_s$ as in the undisturbed system.

Second, we have to solve Poisson's equation for the change in the Hartree potential $V_H^1(R,z)$ due to the induced charge $\rho^1(R,z)$:

$$\nabla^2 V_H^1=-8\pi\frac{\bar{\kappa}}{\kappa_s}\rho^1(R,z) , \tag{29}$$

valid in the semiconductor $z>0$. The asymptotic behavior for $R\to\infty$ or $z\to\infty$ is

$$V_H^1(R,z)\sim 2Z/(R^2+z^2)^{1/2} \tag{30}$$

to secure macroscopic charge neutrality. On the insulator side the potential should fulfill Laplace's equation with the same asymptotic behavior. The boundary conditions at the interface require the continuity of V_H^1 and $\kappa\partial V_H^1/\partial z$. The latter boundary condition is taken care of by the method of images: to obtain the potential in the semiconductor we replace the two media by one with permittivity κ_s and add the image charge

$$\rho_{im}^1(R,z)=(\kappa_s-\kappa_i)/(\kappa_s+\kappa_i)\rho^1(R,-z)$$

for $z<0$, so that Eq. (29) becomes

$$\nabla^2 V_H^1=-8\pi\frac{\bar{\kappa}}{\kappa_s}[\rho^{-1}(R,z)+\rho_{im}^1(R,z)] , \tag{31}$$

to be solved in the whole space with the asymptotic behavior given in Eq. (30).

We could now proceed the solve the Poisson equation by some method. However, we have to construct an iteration towards self-consistency and the long range of the Coulomb interaction is inconvenient, since a small deviation at one point affects

a complete set of subband wave functions $H_0\zeta_n(z)=E_n\zeta_n(z)$. In contrast to the calculation in Sec. II we have to do this self-consistently, i.e., we find the Hartree potential by solving a one-dimensional Poisson equation and an exchange-correlation potential from the density given by

$$\rho^0(z)=N_{inv}|\zeta_0(z)|^2$$
$$=g_v g_s(E_F-E_0)|\zeta_0(z)|^2/4\pi , \tag{27}$$

where we have restricted our calculation to inversion-layer densities N_{inv} for which only the lowest subband is occupied. This is the subband problem without impurity treated first by Ando.[21] To solve the Schrödinger equation with the impurity (24), (25) we proceed as described in Sec. II and use the subband wave functions to expand the wave function ψ_E. The perturbing potential is then the screened impurity potential $U(R,z)+V_H^1(R,z)+V_{xc}^1(R,z)$ instead of just the bare potential $U(R,z)$, but in other respects Eqs. (4) and (5) remain structurally the same and can be solved by the methods described previously. Once we have the wave functions we obtain the density

$$\rho(R,z)=\rho^0(R,z)+\rho^1(R,z) ,$$

where

the potential essentially everywhere. We therefore introduce the trick suggested and employed in spherical symmetry by Manninen *et al.*[27] and add $-k^2V_H^1$ on both sides of Eq. (31):

$$(\nabla^2-k^2)V_H^1=-8\pi\frac{\bar{\kappa}}{\kappa_s}[\rho^1+\rho_{\rm im}]-k^2V_H^1 \tag{32}$$

where k^2 is a parameter that can be freely chosen. The Green's function corresponding to the left-hand side is now a screened short-range interaction, so that small deviations in one point only affect the potential in the neighborhood. On the right-hand side we insert the charge of the present iteration and the Hartree potential from the previous iteration. After convergence the potential does not change from one iteration to the next and Eq. (32) is exactly the same as the Poisson equation (31).

In order to solve Eq. (32) we expand in spherical harmonics. Setting $R=r\cos\theta$ and $z=r\sin\theta$ we have

$$\begin{aligned} V_H^1(R,z)&=\sum_{l=0}^{\infty}V_l(r)Y_{l0}(\theta)\ ,\\ \rho^1(R,z)+\rho_{\rm im}(R,z)&=\sum_{l=0}^{\infty}\rho_l(r)Y_{l0}(\theta)\ , \end{aligned} \tag{33}$$

so that

$$\begin{aligned} \rho_l(r)&=2\pi\int_0^{\pi}\sin\theta\,d\theta[\rho^1(r,\theta)+\rho_{\rm im}(r,\theta)]Y_{l0}(\theta)\\ &=\sqrt{2\pi}\left[\int_0^1\rho^1(r,\mu)\bar{P}_l(\mu)d\mu+\frac{\kappa_s-\kappa_i}{\kappa_s+\kappa_i}\int_{-1}^0\rho^1(r,-\mu)\bar{P}_l(\mu)d\mu\right]\\ &=\left[1+(-1)^l\frac{\kappa_s-\kappa_i}{\kappa_s+\kappa_i}\right]\sqrt{2\pi}\int_0^1\rho^1(r,\mu)\bar{P}_l(\mu)d\mu\ , \end{aligned} \tag{34}$$

where $\bar{P}_l(\mu)$ are the normalized Legendre polynomials. Inserted in Eq. (32) the expansions lead to uncoupled differential equations for each l:

$$\frac{d^2V_l}{dr^2}+\frac{2}{r}\frac{dV_l}{dr}-\left[k^2+\frac{l(l+1)}{r^2}\right]V_l=-8\pi\frac{\bar{\kappa}}{\kappa_s}\rho_l(r)-k^2V_l(r)\ . \tag{35}$$

The Green's function corresponding to the left side is

$$G_l(r,r')=\begin{cases} r'^2\dfrac{2k}{\pi}i_l(kr')k_l(kr), & r>r'\\ r'^2\dfrac{2k}{\pi}k_l(kr')i_l(kr), & r<r' \end{cases} \tag{36}$$

where i_l and k_l are the spherical Bessel functions. The solution of Eq. (35) is then

$$\begin{aligned} V_l(r)=&\frac{2k}{\pi}i_l(kr)\int_r^{\infty}r'^2k_l(kr')\left[8\pi\frac{\bar{\kappa}}{\kappa_s}\rho_l(r')+k^2V_l(r')\right]dr'\\ &+\frac{2k}{\pi}k_l(kr)\int_0^r r'^2i_l(kr')\left[8\pi\frac{\bar{\kappa}}{\kappa_s}\rho_l(r')+k^2V_l(r')\right]dr'\ , \end{aligned} \tag{37}$$

or

$$V_l(r)=C_l(r)i_l(kr)+S_l(r)k_l(kr) \tag{38}$$

with

$$\frac{dC_l}{dr}=-\frac{2k}{\pi}\left[8\pi\frac{\bar{\kappa}}{\kappa_s}\rho_l(r)+k^2V_l(r)\right]\ , \tag{39}$$

$$\frac{dS_l}{dr}=\frac{2k}{\pi}\left[8\pi\frac{\bar{\kappa}}{\kappa_s}\rho_l(r)+k^2V_l(r)\right]\ . \tag{40}$$

The boundary conditions are $C_l(\infty)=S_l(0)=0$. In practice we restrict the calculations to $r<r_m$ with some large r_m and assume that the potential outside r_m has the form given by Eq. (31). We therefore integrate (40) from $r=0$ to $r=r_m$ and thereupon integrate (39) backwards using the boundary condition which follows:

$$C_l(r_m)=\frac{[V_0(r_m)\delta_{l0}-S_l(r_m)k_l(kr_m)]}{i_l(kr_m)}\ , \tag{41}$$

with $V_0(r_m)=\sqrt{4\pi}\,2Z/r_m$. In Eqs. (35)–(40) V_l on the right-hand sides is taken to be the potential from the previous iteration.

The third point to be treated is the exchange-correlation potential. According to the general theory of Kohn and Sham[26] the best local potential in the limit of slowly varying density is the exchange correlation contribution to the chemical potential of a homogeneous electron gas having the local density. In our case we have the problem that our electrons are moving in a polarizable background and that we have two media of different polarizability. Thus two electrons deep in the semiconductor interact in a background of dielectric constant κ_s while two electrons very close to the interface effectively have a background dielectric constant $\bar{\kappa}$. This makes a substantial difference, and one must somehow incorporate this dependence on distances from the interface. Here we follow completely the approximation used by Ando[21]; two electrons at $\vec{r}_1=(\vec{R}_1,z_1)$ and $\vec{r}_2=(\vec{R}_2,z_2)$ interact via the Coulomb interaction as

$$V(\vec{r}_1,\vec{r}_2)=\frac{e^2}{4\pi\epsilon_0\kappa_s}\left[\frac{1}{|\vec{r}_1-\vec{r}_2|}+\frac{\kappa_s-\kappa_i}{\kappa_s+\kappa_i}\frac{1}{[(\vec{r}_1-\vec{r}_2)^2+4z_1z_2]^{1/2}}\right] \tag{42}$$

where the image term makes the interaction depend not only on $\vec{r}_1-\vec{r}_2$ but also explicitly on z_1 and z_2. The approximation now consists of taking $z_1\cong z_2=z$ in the $4z_1z_2$ term and then calculating the contribution of exchange and correlation to the chemical potential of a homogeneous electron gas with the interaction described by (42), where z enters as a parameter. We then obtain an exchange-correlation potential which depends on the local density $\rho(R,z)$ and explicity on z.

For the results presented in the next section we have directly taken over the exchange-correlation potential calculated by Ando.[21] It is a special case of the various potentials which will be derived in Sec. VI and we shall postpone details to that section.

The final point is how to obtain convergence in the iteration of the potentials and the density. We have already incorporated this in the discussion of the Hartree-potential calculation. By choosing a large enough free parameter k the changes from iteration to iteration are slow and gradually vanish independent of the starting potential. A value of k between 1 and 1.5 turns out to give the fastest convergence; smaller values of k lead to divergence and greater values slow down the convergence rate.

V. RESULTS WITH SCREENING

We first show some representative results for an impurity of charge $Z=1$ situated on the interface $z_0=0$. For $N_{\text{depl}}=3.6\times10^{11}$ cm^{-2} and an inversion layer density $N_{\text{inv}}=3\times10^{12}$ cm^{-2} we show in Fig. 3 the total charge distribution

$$\rho(R,z)=\rho^0(z)+\rho^1(R,z)$$

and the induced charge density $\rho^1(R,z)$. It can be seen that there is a pileup of charge near the impurity, and also that the charge moves closer to the surface there. This is most evident in the induced charge, where there are regions behind the pileup which are depleted of charge. In the direction parallel to the interface these regions gradually lead to Friedel oscillations[28] of period π/k_F where $k_F=E_F^{1/2}$ is the Fermi wave vector.

In Fig. 4 we show the phase shift $\eta_0(E)$ for $m=0$ wave functions and the total phase shift $\sum_{m=-\infty}^{\infty}\eta_m(E)$. The arrow below 0 (the edge E_0 of the lowest subband) represents the bound state which has $m=0$. According to Eq. (28) the bound state contains four electrons because of the spin and valley degeneracy. The theory as formulated in Sec. IV does not allow a lifting of this degeneracy, so at first one might consider this an artifact of the theory. However, more refined methods which we describe in Secs. VI and VII and which do not have this degeneracy built in also lead to a fourfold degenerate bound state, so one should not discard the result even though it is unexpected. The sum of the phase shifts in the continuum are therefore negative, and we see that the Friedel sum rule (see, e.g., Ref. 9) is very well fulfilled; at the Fermi energy the phase shift is $-3\pi/4$, which means that three electrons have been taken out of the continuum, so the total induced charge is $+1$ electron. This is a necessary condition for self-consistency of our results. The induced density of states is $\propto d\eta/dE$ [Eq. (23)], so, mainly, states near the bottom of the subband have been taken out to form the bound state, which is very shallow.

The scattering of the subband wave function leads to a finite mobility parallel to the interface. The transport scattering time τ_{tr} has been shown by Stern and Howard[9] to be given by (in SI units)

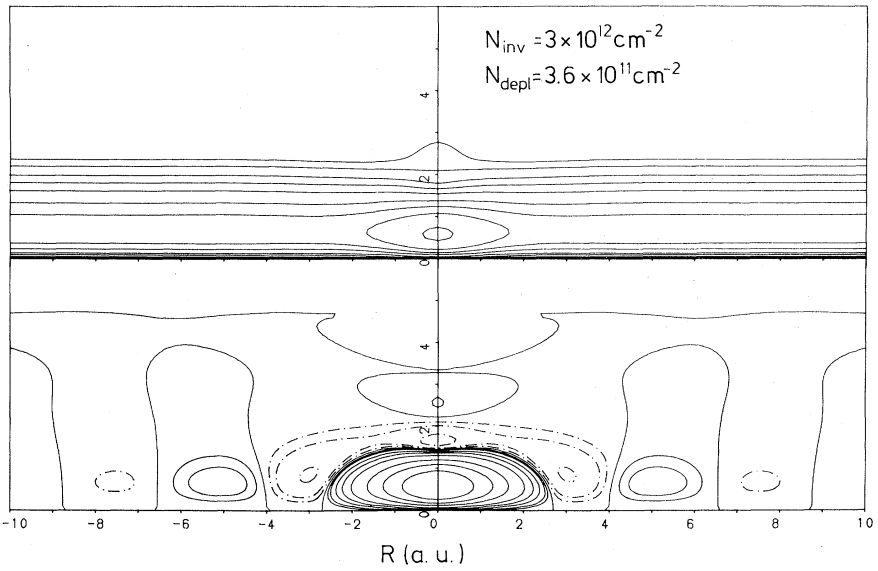

FIG. 3. Electronic density in the inversion layer near an impurity at the origin. $N_{\text{inv}}=3\times10^{12}$ cm^{-2}, $N_{\text{depl}}=3.6\times10^{11}$ cm^{-2}. Upper frame: total density. Lower frame: induced density. The contour plots are logarithmic with three contours per decade, lowest value 10^{-3} electrons/a_0^{*3} in the induced density negative values are denoted by broken curves, and the contour for zero density is also shown.

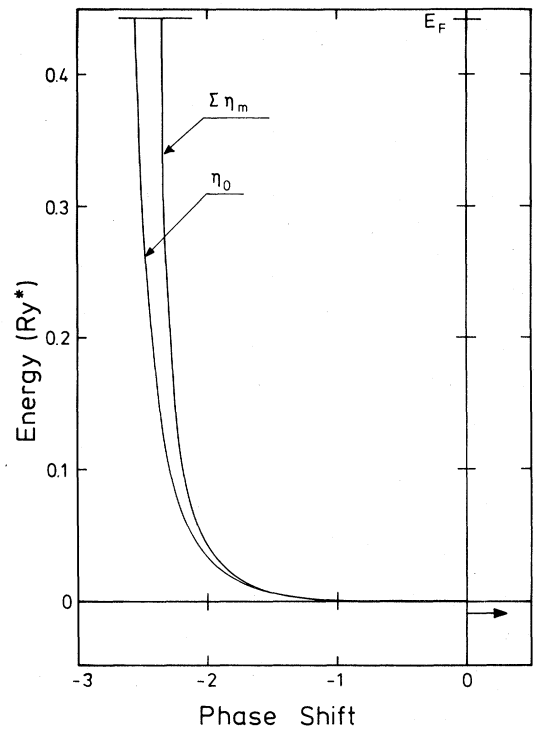

FIG. 4. Phase shift η_0 for $m=0$ states and total phase shift $\sum\eta_m$ vs energy measured from the bottom of the lowest subband. $N_{\text{inv}}=3\times10^{12}$ cm^{-2}, $N_{\text{depl}}=3.6\times10^{11}$ cm^{-2}. E_F is the Fermi energy. The bound state below the subband is denoted by an arrow.

$$\frac{1}{\tau_{\text{tr}}}=N_{\text{ox}}\frac{4\hbar}{m_t}\sum_{m=0}^{\infty}\sin^2[\eta_m(E_F)-\eta_{m+1}(E_F)] \tag{43}$$

where N_{ox} is the areal density of impurities in the oxide.

In Fig. 5 we compare our results, assuming the impurities to lie on the interface, with the most recent and most relevant experimental results by Hartstein *et al.*[4] By drifting known amounts of Na^+ to the interface they were able to separate the

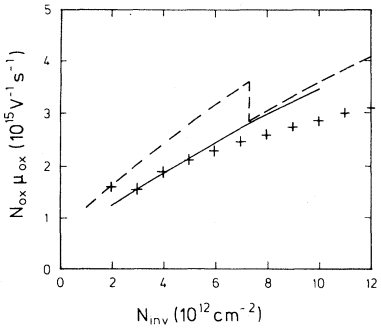

FIG. 5. Mobility due to scattering by oxide charge as a function of inversion-layer density. Full curve: present work. Broken curve: linear screening (Mori and Ando, Ref. 18). Experimental points determined by Hartstein *et al.* (Ref. 4). $N_{\text{depl}}=3.6\times10^{11}$ cm^{-2}.

surface roughness, oxide charge, and phonon scattering contributions, and the figure shows the oxide-charge scattering contribution linear in N_{ox}, i.e., the impurity scattering in the limit of few impurities. Also shown is the most recent linear-screening theory by Mori and Ando.[18] It can be seen that our calculation is a considerable improvement, and that the population of subband 1 invoked by those authors for $N_{inv} \geq 7.3\times10^{12}$ cm^{-2} is not necessary to obtain quantitative agreement with the experiment. When exchange and correlation are taken into account, subband one is not populated below 10×10^{12} cm^{-2} in this case.

The energy of the bound state is shown in Fig. 6. The state is very shallow and has a long range. The binding energy $-E_B$ increases with N_{inv} because the electrons are pressed closer to the impurity at higher densities. For the mobility, however, this effect is more than balanced by the increase in Fermi velocity of the electrons, so that the mobility increases with N_{inv}.

VI. VALLEY-DENSITY FUNCTIONAL METHODS

The fourfold degeneracy of the bound state which comes out of the calculation described in the preceding section leaves one with an uneasy feeling, since one is used to having only one electron bound to an impurity in a semiconductor. For this reason and for other purposes we want to extend the method to allow for a lifting of the degeneracy. The spin-density-functional method[29,30] is an extension of this type and here we can proceed analogously because the valley degeneracy acts very similarly to the spin degeneracy. The reason for this can be understood by considering a homogeneous Si crystal. The conduction electrons have the

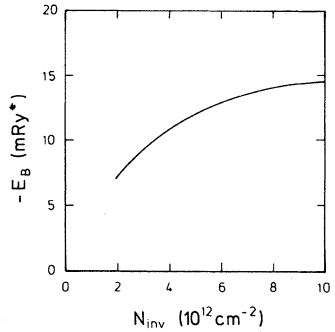

FIG. 6. Energy E_B of bound state relative to bottom of lowest subband as a function of inversion-layer density. $N_{depl}=3.6\times10^{11}$ cm^{-2}.

following wave functions close to the valley minima at $\pm k_0\hat{z}$:

$$\psi_{\vec{k}s\pm}(\vec{r}) \cong e^{i\vec{k}\cdot\vec{r}} e^{\pm ik_0 z} u(\vec{r})\eta_s \ , \tag{44}$$

where $\vec{k}$ is the wave vector measured from the valley minimum, η_s is a spin function, and $u(\vec{r})$ is a cell periodic function. For such functions the Coulomb-interaction matrix element is calculated to be

$$V(q) \cong \begin{cases} \dfrac{e^2}{\kappa_s \epsilon_0 q^2}\delta_{ss'} & \text{for intravalley scattering} \\ \dfrac{e^2}{4\kappa_s\epsilon_0(\pi/a-k_0)^2}\delta_{ss'} & \\ & \text{for intervally scattering} \ . \end{cases} \tag{45}$$

Here $\vec{q}$ is the momentum transferred in the scattering process and π/a is half the Brillouin-zone size in the [100] direction. Because of the large finite value of $\pi/a - k_0$ the intervalley scattering process is much less important and we shall neglect it completely. It is then clear that the valley index $\pm$ is analogous to the spin, which also cannot be changed by the electron-electron interaction.

Because of this analogy it is now straightforward to use the methods applied earlier to the spin-polarized homogeneous electron gas[29,30] in the system of a homogeneous electron gas which has valley degrees of freedom and is allowed to have unequal population of the valleys. For our present purpose we can assume that the spin degrees of freedom are equally populated and only two valleys are occupied. The homogeneous electron gas is then described by the parameter density $n=n^+ + n^-$ and "valley polarization" $\xi=(n^+-n^-)/n$, where $n^\pm$ is the electron density in each valley.

In order to construct an exchange-correlation potential we need to calculate the contributions of exchange and correlation to the chemical potential $\mu_{xc}^{\pm}(n,\xi)$ of the homogeneous valley-polarized electron gas. To obtain manageable expressions for these energies we have to assume that the valleys are isotropic with an effective mass m. The Fermi wave vector in each valley is then

$$k_F^{\pm} = (3\pi^2 n^{\pm})^{1/3} = \left[\frac{3\pi^2}{2}n\right]^{1/3}(1\pm\xi)^{1/3}$$

$$\equiv k_F(1\pm\xi)^{1/3} \ . \tag{46}$$

Furthermore, we shall make use of the plasmon-pole approximation,[31,32] so with careful consideration of powers of 2 we essentially take over the expressions of Gunnarsson and Lundqvist[30] and of Lundqvist[31] for the effective plasmon ω_q and for the exchange and correlation contributions to the chemical potential $\mu_{xc}^{\pm}=\mu_x^{\pm}+\mu_c^{\pm}$ as follows:

$$(\hbar\omega_q)^2=(\hbar\omega_p)^2+\frac{8\epsilon_F}{(1+\xi)^{1/3}+(1-\xi)^{1/3}}\frac{\hbar^2q^2}{2m}+\left[\frac{\hbar^2q^2}{2m}\right]^2, \tag{47}$$

$$\mu_x^{\pm}=-\frac{1}{4\pi^2k_F^{\pm}}\int_0^{\infty}q\,dqV(q)\int_{|k_F^{\pm}-q|}^{k_F^{\pm}+q}p\,dp\,\Theta(k_F^{\pm}-p)\,, \tag{48}$$

$$\mu_c^{\pm}=-\frac{m\omega_p^2}{8\pi^2\hbar k_F^{\pm}}\left[\int_0^{2k_F^{\pm}}\frac{q\,dq}{\omega_q}V(q)\ln\left|\frac{q^2+2k_F^{\pm}q+2m\omega_q/\hbar}{q^2-2k_F^{\pm}-2m\omega_q/\hbar}\right|+\int_{2k_F^{\pm}}^{\infty}\frac{q\,dq}{\omega_q}V(q)\ln\left|\frac{q^2+2k_F^{\pm}q+2m\omega_q/\hbar}{q^2-2k_F^{\pm}q+2m\omega_q/\hbar}\right|\right]$$

where $\omega_p^2=ne^2/\kappa_s\epsilon_0m$, $\epsilon_F=\hbar^2k_F^2/2m$, and $V(q)=e^2/\kappa_s\epsilon_0q^2$.

As mentioned in Sec. IV the presence of the insulator presents a further complication because electrons closer to the interface interact more strongly than electrons deep in the semiconductor. We make the same approximations as Ando,[21] i.e., we replace the actual electron-electron interaction of Eq. (42) by

$$V_{\rm eff}(|\vec{r}_1-\vec{r}_2|,z)=\frac{e^2}{4\pi\kappa_s\epsilon_0}\left[\frac{1}{|\vec{r}_1-\vec{r}_2|}+\frac{\kappa_s-\kappa_i}{\kappa_s+\kappa_i}\frac{1}{[(\vec{r}_1-\vec{r}_2)^2+4z^2]^{1/2}}\right], \tag{49}$$

so that for each z we have a homogeneous electron gas with the interaction given by Eq. (49). We can then still use Eqs. (47) and (48) for $\mu_{xc}^{\pm}$, if we replace $V(q)=e^2/\kappa_s\epsilon_0q^2$ by the Fourier transform of $V_{\rm eff}$:

$$V_{\rm eff}(q,z)=\frac{e^2}{\bar{\kappa}\epsilon_0q^2}\left[\frac{1}{2}\left[1+\frac{\kappa_i}{\kappa_s}\right]+\frac{1}{2}\left[1-\frac{\kappa_i}{\kappa_s}\right]2qzK_1(2qz)\right], \tag{50}$$

where $K_1(x)$ is the modified Bessel function of the second kind.

In Fig. 7 we have plotted some results on the exchange-correlation potentials $\mu_{xc}^{\pm}$. We have introduced scaled atomic units $a_0'=4\pi\bar{\kappa}\epsilon_0\hbar^2/me^2$ and $\mathscr{R}'=e^2/8\pi\bar{\kappa}\epsilon_0a_0'$, and the usual dimensionless parameter $r_s=(3/4\pi n)^{1/3}/a_0'$. Qualitatively, the results resemble those of the spin-polarized electron gas, and for a thorough discussion of the implications of these results on the homogeneous gas we refer to that work.[30] Quantitatively, it is important to notice that our results for $\xi=0$ agree very well with those of Ando,[21] who used a more complicated and supposedly better approximation for the screened interaction. It is therefore safe to assume that the plasmon-pole approximation is quite sufficient.

In the actual calculations we have used the fol-

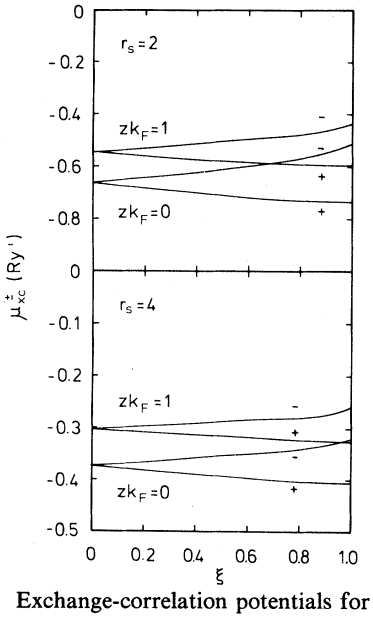

FIG. 7. Exchange-correlation potentials for a valley-polarized electron gas as a function of polarization ξ for two values of density and two values of z. + denotes the more- and − the less-occupied valley.

lowing interpolation formula for $\mu_{xc}^{\pm}$:

$$\mu_{xc}^{\pm}=f(k_F z)\mu_x(r_s)\left[\beta(r_s)\pm\tfrac{1}{2}\delta(r_s)\frac{\xi}{1\pm 0.3\xi}\right], \tag{51}$$

$$\mu_x=-2k_F/\pi ,$$

$$\beta(r_s)=1+0.0862 r_s \ln(1+15.24/r_s) ,$$

$$\delta(r_s)=1-0.012 r_s-0.404 r_s/(1+0.59 r_s) ,$$

$$f(k_F z)=\frac{\bar{\kappa}}{\kappa_s}+\left[1-\frac{\bar{\kappa}}{\kappa_s}\right]\Big/\left[1+\left[\frac{2k_F z}{\Delta(r_s)}\right]^{1.3}\right] ,$$

$$\Delta(r_s)=2.2-0.17 r_s ,$$

where $\mu_{xc}^{\pm}$ is in $\mathscr{R}'$ and $k_F=(9\pi)^{1/3}/2r_s$.

For $z=0$ and $z=\infty$ this expression is accurate to better than 2% for $1<r_s<5$. For intermediate values of z the agreement is not quite so good, but in view of the approximate treatment of the interaction via the image we are content with a z-dependent factor on the expression for $z=0$. It should be mentioned that our results on the inversion layer do not depend much on whether one uses the conductivity mass or the density-of-states mass for m in a_0' and $\mathscr{R}'$.

For the problem of the inhomogeneous system of electrons in the inversion layer and in the potential of an impurity it is now straightforward to apply the general theory of Kohn and Sham[26] as extended to the spin-density-functional formalism.[30] Instead of the Hamiltonian of Eqs. (24) and (25) we get effective Hamiltonians for each valley:

$$H_0^{\pm}=-\frac{m_t}{m_l}\frac{\partial^2}{\partial z^2}+V_{\text{depl}}(z)+V_N^0(z)+V_{xc}^{0\pm}(z) , \tag{52}$$

$$H_1^{\pm}=-\frac{\partial^2}{\partial R^2}-\frac{1}{R}\frac{\partial}{\partial R}-\frac{1}{R^2}\frac{\partial^2}{\partial\theta^2}+U(R,z)+V_H^1(R,z)+V_{xc}^{1\pm}(R,z) , \tag{53}$$

in which

$$V_{xc}^{\pm}(R,z)=\mu_{xc}^{\pm}(n(R,z),\xi(R,z);z)$$

are the local exchange-correlation contributions to the chemical potential, which, as shown above, are functions of the local density, the local valley polarization, and z. Correspondingly, we obtain different solutions to the effective Scrödinger equations $H^{\pm}\psi_E^{\pm}=E^{\pm}\psi_E^{\pm}$ and the total density is given by

$$\rho(R,z)=\sum_{E<E_F}(|\psi_E^{+}|^2+|\psi_E^{-}|^2) , \tag{54}$$

where the sum runs over all states below the Fermi level. In order to solve Eqs. (52)–(54) we can use the same methods as in Sec. IV; essentially the work has only been doubled by treating the two valleys separately. Clearly this formulation has the possibility of different population of the two valleys and of only doubly degenerate bound states.

Before we discuss our results in Sec. VII let us mention two further analogous extensions of the local-density-functional formulation which we have also applied. First, the valley polarization in Eqs. (52)–(54) is not the most general, since each state is still doubly spin degenerate. In principle, one could introduce a formalism in which each spin and valley is treated separately, but what we are really searching for is a solution in which we have just one nondegenerate bound state; to allow such a possibility we can restrict the general formulation to having one valley and spin $(+\uparrow)$ locally occupied differently from the three other $[(+\downarrow),(-\uparrow),(-\downarrow)]$ possibilities. This division leads to equations of the same kind as (52)–(54), only the exchange-correlation potentials are different and must be calculated anew. We shall not present calculations in detail but only show a typical result in Fig. 8. The exchange-correlation potentials are shown for a fixed density $r_s=2$ and for $z=0$ as a function of the "polarization" defined as

$$\xi=3\frac{n_1-n_2}{n} , \tag{55}$$

where n_1 is the density in $(+\uparrow)$ and n_2 is the density in each of the three other combinations, so that $n=n_1+3n_2$. It can be seen by comparison with Fig. 7 that one gains exchange-correlation energy by having all electrons in $(+\uparrow)$, i.e., by mak-

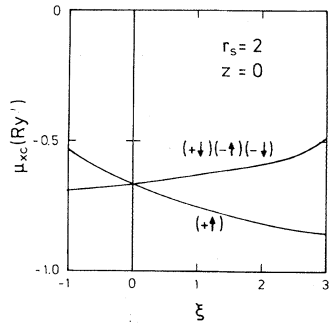

FIG. 8. Example of exchange correlation potentials for a spin- and valley-polarized electron gas as a function of polarization ξ defined in Eq. (55).

ing k_F in $(+\uparrow)$ as large as possible. Of course, this gain occurs at the cost of a higher kinetic energy.

The second application goes beyond what we have described up until now. It has been shown most directly by Stallhofer *et al.*[33] that by applying a suitably directed uniaxial strain on the interface one can lower two of the four valleys which give rise to the primed subbands relative to the two from which the unprimed subbands originate. It is then possible to populate the lowest subband of the primed set. In the natural extension of the local-density-functional method one introduces one exchange-correlation potential V_{xc} for the electrons in the unprimed subbands, and another V'_{xc} for those in the primed subbands. As described in Ref. 34 these potentials then depend on the total density $n+n'$, the polarization $\xi=(n-n')/(n+n')$, and explicitly on z because of the image interaction. We have calculated these potentials analogously to Eqs. (47)–(50) with appropriate changes to take into account the total of four valleys instead of two. For $\xi=1$ we get, with very small deviations, the results of Ando[21] who also calculated the potential V'_{xc} for $\xi=1$.

When an impurity is present at the interface it scatters electrons in both unprimed and primed subbands and the electrons contribute differently to screening. To describe this situation self-consistently we solve the equations analogous to (52)–(53), $H\psi=E\psi$ and $H'\psi'=(E'+S)\psi'$, where

$$H=-\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]-\frac{m_t}{m_l}\frac{\partial^2}{\partial z^2}+V_{\text{depl}}+V_H+V_{xc}\,, \tag{56}$$

$$H'=-\frac{m_t}{m_d}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]-\frac{\partial^2}{\partial z^2}+V_{\text{depl}}+V_H+V'_{xc}\,, \tag{57}$$

S is the strain induced bulk shift of the valleys from which the primed subband originates, and $m_d=(m_t m_l)^{1/2}$ is the density-of-states mass parallel to the interface for the primed subbands. The mass for motion perpendicular to the interface is m_t.[25] In these equations we have neglected the anisotropy parallel to the interface in order to keep the cylindrical symmetry. The equations can be solved by the methods described in Sec. IV.

VII. VALLEY-POLARIZED RESULTS

We have shown earlier[35] that when there is no impurity the Hamiltonian $H_0^{\pm}$ [Eqs. (52)] leads to a lifting of the valley degeneracy at low densities $N_{\text{inv}}\lesssim 3\times 10^{11}$ cm^{-2}. This valley condensation is due to the fact that for low densities it is favorable to increase the Fermi level and thereby the magnitude of the exchange energy by having all electrons in one valley. The gain in energy is larger than the loss in kinetic energy. This transition, which can only occur at low temperatures,[36] has many important implications as discussed in Ref. 35. It has not yet been observed on Si(100) surfaces,[37] but indirectly, various measurements indicate that the valley degeneracy is lifted in the localization regime, e.g., Ref. 38. This is not the problem we want to investigate here, but for our purposes it is important to notice that it is possible within our formalism to lift the valley degeneracy.

With an impurity on the interface we have solved Eqs. (52) and (53) for densities above $N_{\text{inv}}=10^{12}$ cm^{-2}. We have earlier reported[39] that the only self-consistent solution that we have been able to find is the completely unpolarized one in which the bound state contains four electrons. If we run the calculations with the constraint that the bound state be only doubly occupied, we find a polarized solution which, however, has an unoccupied state below E_0, i.e., it is not the self-consistent ground state but represents an excited state with a double "core hole." Similar results are obtained, if we allow only one electron in the bound state and use the exchange-correlation potentials described in connection with Fig. 8, where a polarization of one valley and spin is assumed. We do not expect that a calculation in which all four spin and valley degrees of freedom were allowed to polarize independently would lead to any other result than the completely unpolarized one which we discussed in Sec. V. We therefore conclude that the ground state of our model has a very shallow fourfold degenerate bound state. The very good agreement found with the mobility data (Fig. 5) indicates that this is probably the true ground state, and that the bound state is so shallow that it is difficult to distinguish from a continuum state.

A comparison with the corresponding problem of a charged impurity in a three-dimensional metal is valuable. Self-consistent calculations show that a bound state below the continuum is doubly degenerate, even when spin-density functional methods are used[40]; this is analogous to our situation. In jellium with high density the impurity po-

tential can be so strongly screened that no bound state exists[41]; for strictly two dimensions an attractive potential is expected to have at least one bound state,[42] and because our system is strongly two-dimensional we probably cannot avoid having a bound state.

In Sec. VIII we briefly report some results which further substantiate our claim that the bound state is fourfold degenerate. Experimentally, no bound state has been observed at such high densities; the measurements of McCombe and co-workers[6,7] show only a bound state at low densities, which for higher densities ($\gtrsim 1.5\times10^{12}$ cm^{-2}) merges with the subband transitions. This is at least not in contradiction with our results.

For an inversion layer under uniaxial stress we have earlier reported results[39] on the mobility and bound state when only subband 0′ is occupied. The scattering time τ' was found to be considerably larger than in subband 0 because the electrons are at a greater distance from the impurity. If we take the anisotropy of the mass parallel to the interface into account by defining a longitudinal mobility $\mu_l = e\tau'/m_l$ and a transverse mobility $\mu_t = e\tau'/m_t$, it was found that the longitudinal mobility lies below and the transverse mobility lies above the mobility one has when only subband 0 is filled. The bound state is even more shallow when 0′ is filled than when 0 is filled, because of the different distances from the impurity.

In Fig. 9 we show results for a fixed density $N_{\rm inv}=3\times10^{12}$ cm^{-2} in the intermediate case when the strain is such that both subbands 0 and 0′ are occupied. As a function of the strain-induced energy shift S of the valleys from which the primed subbands originate relative to the valleys of the unprimed subbands, we show in the upper frame the position of the subband edges and the bound state. In the lower frame is shown the mobility of the electrons in subband 0, the longitudinal mobility of the electrons in subband 0′, and the average (longitudinal) mobility. The trends are not very surprising in view of the model; as long as there are electrons in subband 0 they screen the impurity so well that when one starts filling 0′ the electrons in 0′ have a very high mobility of about 50 times that of the electrons in 0. With increasing strain the screening by the electrons in subband 0 becomes less and less efficient, which leads to a reduction of all the mobilities and to an increase in binding energy of the bound state relative to the subband bottom. Finally, for $S > 60$ meV, when the bound state associated with subband 0 is also empty, we have the lower longitudinal mobility

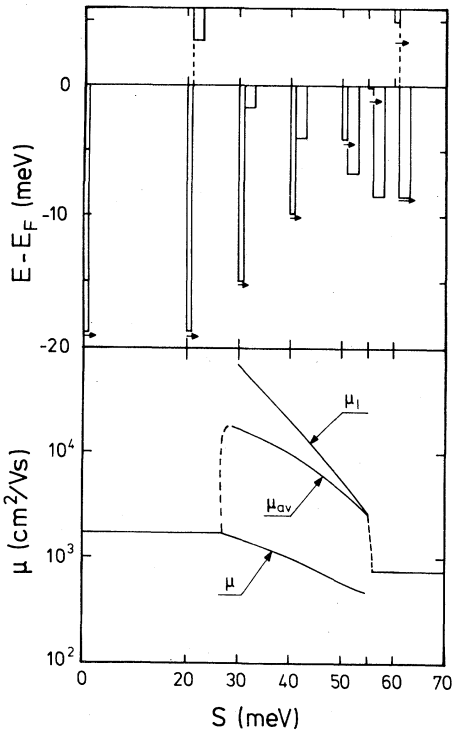

FIG. 9. Results on the system under uniaxial stress. S is the stress parameter defined in Eq. (57). In the upper frame we show the positions of subband edges and bound states relative to the Fermi energy. The thin columns represent subband 0, the thicker columns represent subband 0′. In the lower frame we show the impurity contribution to the mobility of electrons in subband 0 (μ), in subband 0′ (μ_l'), and the average mobility ($\mu_{\rm av}$). The mobility is calculated for $N_{\rm ox}=10^{12}$cm^{-2}. $N_{\rm inv}=3\times10^{12}$ cm^{-2}, $N_{\rm depl}=3.6\times10^{11}$ cm^{-2}.

and the very shallow, fourfold-occupied bound state associated with subband 0′. It should be mentioned that since intervalley scattering is neglected in our approximation, resonances below 0′ in the continuum of subband 0 would show up as bound states.

The trends in the mobility are the same as those found by Takada,[17] but there are important differences between his work and the present work, especially with respect to the bound states; Takada assumes that only one electron can occupy one bound state. He then finds relatively strongly bound states ($E_B \sim 10$ meV), one associated with subband 0 and one associated with subband 0′, and assumes that only the lower of these two states is occupied. Clearly, this involves rather unfounded assumptions about the interaction between electrons in bound states and between the electrons in the subbands and a bound electron. In our calculation the

screened potentials are determined self-consistently, and as long as the bound state associated with subband 0 is occupied, the potential has no bound state associated with subband 0′. Also, the scattering by the neutral impurity treated separately by Takada is consistently included in our calculation.

Experimentally, there have not been investigations on the separate effect of oxide-charge scattering under stress, so it is not possible to compare directly with our calculations. Gesch *et al.*[43] have observed that the mobility as a function of inversion-layer density has two maxima at intermediate strain which indicate population of two different subbands, but at the lower densities, where impurity scattering is thought to dominate, they see an increased resistivity when both subbands are occupied, which is contrary to the results of both Takada[17] and the present calculation. Recent measurements by Kastalsky and Fang[44] show qualitatively the same behavior. The marked decrease in mobility at lower densities, even for strains so low that one would not expect transfer of electrons to subband 0′, led the latter authors to suggest that tail states below $E_{0'}$ can be occupied by a localized electron, which acts as a charged impurity.[45] We should like to point at yet another possible explanation. Suppose the interface contains built-in strains which are randomly distributed in size and direction.[46] Then $E_{0'0}$ is not fixed but varies in the inversion layer, and for a moderate extra strain, domains of the inversion layer get into the mixed state. With increasing external strain the domains increase until some domains contain only electrons in 0′, and at sufficiently high strain all electrons are in 0′. In the intermediate regime the conductivity is now also a percolation process for electrons in subband 0 and 0′ separately, analogous to what has been suggested[35] in simple (100) surfaces at valley-condensation densities.[47] The fact that the mobility reduction is more pronounced at lower densities can be understood as an effect of exchange and correlation which enhances $|E_{0'0}|$, so that each domain contains mainly one type of electron and is highly resistive for the other type. It is difficult to distinguish this model from that of Kastalsky and Fang[44] on the basis of the mobility results alone, but the model presented here explains also the Shubnikov–de Haas results both when the direction of the external strain is such that the 0′ subband is doubly valley degenerate and when it is such that 0′ is fourfold degenerate.[46] After this digression let us note with satisfaction that the mobilities measured when only 0′ is occupied agree qualitatively with our calculation: The longitudinal mobility is lower and the transverse mobility is higher than the mobility measured when only subband 0 is occupied.[43]

VIII. CONCLUSION

We have studied the influence of impurities on electrons in an inversion layer on Si(100) in two limits. First the very low density limit where there are fewer electrons than impurities, so that the system can be treated as a one-electron problem. Our results for the most deeply bound states agree very well with those of more approximate theories; for excited bound states with $m=0$ our results are new and illustrate nicely the interplay of impurity potential and depletion-layer potential. The excitation spectrum of such bound electrons shows an interesting dependence on polarization for which some results have been published,[48] but we postpone details of those calculations and comparison with experiments to a later publication. Second, we have studied the limit of many electrons compared to the number of impurities and used several versions of density-functional methods to investigate scattering and binding of electrons, when screening is of decisive importance. For the impurity-limited mobility of electrons in subband 0 we have obtained considerable improvement over simpler linear theories[16,18] when compared to the most recent measurements.[4] Detailed calculations, which allow valley and/or spin splitting, all lead to the result that there is one bound state below the lowest subband, which is fourfold occupied and so shallow that it is difficult to distinguish from a state in the bottom of the continuum. One might raise two objections to this result. The first might be that even valley-density-functional methods cannot describe these questions of degeneracy correctly. We tend not to believe this objection, mainly because the methods do show a lifting of the degeneracy for the system without impurities,[35] and because of the success the spin-density functional method has had in correctly describing, e.g., magnetic impurities in metals.[49] Also a more strongly bound, singly occupied bound state would probably lead to stronger scattering and ruin the good agreement for the mobility.

The second objection might be that since the bound state is very extended the model of just one impurity cannot be valid and interaction with other impurities might twist the potential in such a way that a polarized solution emerges. We have made a simple attempt to investigate this point; it is not

possible to calculate the self-consistent potential of an arbitrary complex of impurities without going considerably beyond the cylindrical symmetry we have made heavy use of up until now, so instead, we have considered two distributions of impurities with a total charge of four: one positive point charge surrounded by a circular line charge with a total of three elementary charges, and a circular line charge containing four elementary charges. These distributions conserve cylindrical symmetry and no serious change in the programs is necessary. For both distributions the results are as follows: When the radius of the ring is small, we find *one* strongly bound state which is fourfold occupied. With increasing radius this state becomes more and more shallow but remains fourfold degenerate. When the radius becomes of the order of $8-10a_0^*$, the binding energy is about the same as for a single impurity and an additional bound state having $m=\pm 1$ emerges. While not being a proof, these results are a strong indication that in the limit of very few impurities there is more than one electron in the bound state around each impurity in support of the seemingly unusual occupation of the bound state found in the earlier sections.

Also in the high-density limit we have studied the effects of an impurity on the inversion layer, when external stress allows occupation of subband 0′. As one could expect from earlier calculations[17] the results showed a strongly increased mobility for the 0′ electrons when the electrons in subband 0 very efficiently screen the impurity potential. As for the bound state our calculations show always one fourfold-occupied bound state; as long as there is an occupied bound state below subband 0, the potential does not have a bound state below 0′. It would be highly interesting to see drifted Na^+ experiments similar to those of Hartstein *et al.*[4] under stress, so that the impurity scattering could be measured separately. At the moment a direct comparison with experiments is not possible, but our results do not seem to be observed, so our model is probably too simple.

There remains the most difficult problem of describing the intermediate range between very low inversion-layer density and high density $N_{\text{inv}} \gtrsim 2\times 10^{12}$ cm^{-2}. The outstanding challenge here is how to handle the incomplete screening and the details of occupation of more than one bound state at an impurity, when the number of electrons is somewhat larger than the number of impurities. Coming from the low-density side one must describe the system within a Hubbard model,[50] which, however, contains essentially unknown parameters and can at most give a qualitative idea of the behavior of the system. Apart from such general work the only attempt to discuss in a self-consistent way what happens was done by Stern.[15] He used a linearized screening model and assumed the impurities to give rise to a fluctuating potential which causes band tailing on the subband and possibly an impurity band. Furthermore, he assumed at most only one singly occupied bound state per impurity with a chosen binding energy independent of the number of electrons. With this scheme he could only qualitatively model the impurity peak observed below threshold by Fowler *et al.*[5] in dc conductivity. But the model does not address problems like changes in binding energy and occupation of bound states with increasing screening.

Our methods do not seem easily adaptable to this problem. If we try to lower the inversion-layer density below $N_{\text{inv}} \simeq 10^{12}$ cm^{-2} we run into numerical problems and cannot obtain convergence. At present it is not clear whether this is purely numerical and could be solved or whether it is due to a too small area of integration around the impurity. Anyway, a model having only one impurity whose potential cannot be perfectly screened must miss essential features of the physics, so we have not pursued this direction with much effort. The problem of realistically describing the screening and occupation of impurities at low densities thus remains a challenge. Needless to say this is not the only problem in connection with electronic transport in inversion layers at low densities. The most prominent ones are the questions of weak localization and the Anderson transition in two dimensions. The literature on these subjects is expanding so rapidly that we shall only refer to Gold and Götze[51] and references therein.

ACKNOWLEDGMENT

This work was supported by Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 128.

*Present address: THOMSON-CSF, B.P. 10, F-91401 Orsay Cedex, France.
[1]F. F. Fang and A. B. Fowler, Phys. Rev. 169, 619 (1968).
[2]A. Hartstein and A. B. Fowler, J. Phys. C 8, L249 (1975).
[3]A. Hartstein and A. B. Fowler, Phys. Rev. Lett. 34, 1435 (1975).
[4]A. Hartstein, A. B. Fowler, and M. Albert, Surf. Sci. 98, 181 (1980).
[5]A. B. Fowler and A. Hartstein, Philos. Mag. B 42, 949 (1980), and references therein.
[6]B. D. McCombe and D. E. Shafer, in *Proceedings of the 14th International Conference on the Physics of Semiconductors, Edinburgh 1978,* edited by B. L. H. Wilson (Institute of Physics, Bristol 1979), p. 1227.
[7]B. D. McCombe and T. Cole, Surf. Sci. 98, 469 (1980).
[8]H. R. Chang and F. Koch, Surf. Sci. 113, 144 (1982).
[9]F. Stern and W. Howard, Phys. Rev. 163, 816 (1967).
[10]A. V. Chaplik, Zh. Eksp. Teor. Fiz. 60, 1845 (1971) [Sov. Phys.—JETP 33, 997 (1971)].
[11]A. V. Chaplik and M. V. Entin, Zh. Eksp. Teor. Fiz. 61, 2496 (1971), [Sov. Phys.—JETP 34, 1335 (1972)].
[12]B. G. Martin and R. F. Wallis, Phys. Rev. B 18, 5644 (1978).
[13]N. O. Lipari, J. Vac. Sci. Technol. 15, 1412 (1978).
[14]G. M. Kramer and R. F. Wallis, in *Proceedings of the 14th International Conference on the Physics of Semiconductors, Edinburgh 1978.* edited by B. L. H. Wilson (Institute of Physics, Bristol 1979), p. 1243; Surf. Sci. 113, 148 (1982).
[15]F. Stern, Surf. Sci. 58, 162 (1976).
[16]F. Stern, Surf. Sci. 73, 197 (1978).
[17]Y. Takada, J. Phys. Soc. Jpn. 46, 114 (1979).
[18]S. Mori and T. Ando, Phys. Rev. B 19, 6433 (1979).
[19]O. Hipólito and V. B. Campos, Phys. Rev. B 19, 3083 (1979).
[20]B. Vinter, Phys. Rev. B 13, 4447 (1976); 15, 3947 (1977).
[21]T. Ando, Phys. Rev. B 13, 3468 (1976).
[22]F. J. Ohkawa, Surf. Sci. 58, 326 (1976).
[23]W. Jones and N. H. March, *Theoretical Solid State Physics* (Wiley, New York, 1973), p. 996.
[24]B. Vinter, Solid State Commun. 28, 861 (1978).
[25]F. Stern, Phys. Rev. B 5, 4891 (1972).
[26]W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
[27]M. Manninen, R. Nieminen, P. Hautojärvi, and J. Arponen, Phys. Rev. B 12, 4012 (1975).
[28]A. Blandin, E. Daniel, and J. Friedel, Philos. Mag. 4, 180 (1959).
[29]U. von Barth and L. Hedin, J. Phys. C 5, 1629 (1972).
[30]O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B 13, 4274 (1976).
[31]B. I. Lundqvist, Phys. Konden. Mater. 6, 163 (1967); 6, 206 (1967).
[32]A. W. Overhauser, Phys. Rev. B 3, 1888 (1971).
[33]P. Stallhofer, J. P. Kotthaus, and G. Abstreiter, Solid State Commun. 32, 655 (1979).
[34]B. Vinter, Solid State Commun. 32, 651 (1979).
[35]W. L. Bloss, L. J. Sham, and B. Vinter, Phys. Rev. Lett. 43, 1529 (1979).
[36]S. Das Sarma and B. Vinter, Phys. Rev. B 26, 960 (1982).
[37]T. Cole, B. D. McCombe, J. J. Quinn, and R. K. Kalia, Phys. Rev. Lett. 46, 1096 (1981).
[38]A. Gold, S. J. Allen, B. A. Wilson, and D. C. Tsui, Phys. Rev. B 25, 3519 (1982).
[39]B. Vinter, Surf. Sci. 98, 197 (1980).
[40]J. K. Nørskov, Phys. Rev. B 20, 446 (1979).
[41]C. O. Almbladh, U. von Barth, Z. D. Popovic, and M. J. Stott, Phys. Rev. B 14, 2250 (1976).
[42]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, London, 1959), p. 156.
[43]H. Gesch, I. Eisele, and G. Dorda, Surf. Sci. 73, 81 (1978).
[44]A. Kastalsky and F. F. Fang, Surf. Sci. 113, 153 (1982).
[45]This mechanism was first suggested by A. B. Fowler, Phys. Rev. Lett. 34, 15 (1975).
[46]B. Vinter and A. W. Overhauser, Phys. Rev. Lett. 44, 47 (1980).
[47]Fluctuations in density have also been shown by J. R. Brews [J. Appl. Phys. 46, 2181 (1975); 46, 2193 (1975)] to lead to a reduction in mobility.
[48]B. Vinter, Surf. Sci. 113, 140 (1982).
[49]R. Podloucky, R. Zeller, and P. H. Dederichs, Phys. Rev. B 22, 5777 (1980).
[50]J. Hubbard, Proc. R. Soc. London, Ser. A 276, 238 (1963).
[51]A. Gold and W. Götze, J. Phys. C 14, 4049 (1981).