

## Electronic ground state of inversion layers in many-valley semiconductors

M. J. Kelly\* and L. M. Falicov

Department of Physics, † University of California, Berkeley, California 94720

(Received 22 July 1976)

We present a self-consistent many-body calculation of the possible ground states in the inversion layers of many-valley semiconductors (as typified by Si) in the metal-oxide-semiconductor configuration. A complex phase diagram results when the various electronic interactions are regarded as free parameters; for experimentally reasonable values, a charge-density wave is the ground state at the Si(111)-SiO<sub>2</sub> interface with the experimentally observed valley degeneracy and cyclotron mass. At the Si(100)-SiO<sub>2</sub> interface, the paramagnetic state remains stable over the experimentally accessible region of the phase diagram. Applications to other systems are briefly mentioned and experimental tests of the theory are suggested.

### I. INTRODUCTION

When a gate voltage is applied across a metal-oxide-semiconductor device, the complicated re-arrangement of electrons that takes place in the semiconductor is obtained from a self-consistent solution of Poisson's equation and the Schrödinger equation. The induced charge at the metal-oxide interface gives rise to a potential gradient in the oxide and band bending in the semiconductor. The self-consistent solution includes an inversion layer of electrons trapped in the bent bands very near the oxide-semiconductor interface. The potential of a test charge as a function of position is shown in Fig. 1 together with the typical dimensions of the system. From the latter it is clear that quantum effects will be prominent.<sup>1</sup> Using an adaptation of an effective-mass theory of isolated impurities in semiconductors due to Luttinger and Kohn,<sup>2</sup> Stern and Howard<sup>3</sup> treated the self-consistent inversion-layer potential as a similar perturbation, and have described the quantized two-dimensional *n*-type inversion layers of silicon as a function of the surface crystallography.<sup>4</sup> We describe aspects of this theory in more detail below as certain features are incorporated into our own approach.

Until recently, most experiments on *n*-type Si inversion layers were performed at the Si(100)-SiO<sub>2</sub> interface, to exploit a high electron mobility there.<sup>5</sup> At this interface, two of the six anisotropic conduction-band valleys, along the  $\langle 100 \rangle$  axes in the bulk Brillouin zone, present a high effective mass normal to the surface. The theory<sup>3</sup> predicts that these two valleys will be occupied in the inversion layer, and that the other four should remain empty until a large gate voltage is applied. Shubnikov-de Haas studies of this system do indeed reveal an occupied valley degeneracy of two.<sup>6</sup> Occupation of the four valleys with lighter effective mass has not been seen, although excited levels of heavier mass valleys have been identified.<sup>7</sup>

At the Si(111) surface all six  $\langle 100 \rangle$  valleys present the same normal effective mass. Instead of all six being occupied, experiments<sup>8-10</sup> reveal an occupied valley degeneracy of  $2 \pm 0.2$ . In both these latter systems, the cyclotron masses are 20% higher<sup>8</sup> than those of the isolated valleys that the Stern and Howard theory would predict. Less well characterized data are also available for tellurium accumulation layers<sup>11</sup> with higher cyclotron masses and lower occupation degeneracies being obtained.

We present a many-body calculation of the possible ground states that arise when electrons in the different valleys become correlated. The inversion-layer geometry shares a number of features in common with the layered compounds, where charge-density-wave ground states have been identified.<sup>12</sup> A particular type of charge-density wave can explain all the experimental facts observed above for the Si(111)-SiO<sub>2</sub> system. For the Si(100)-SiO<sub>2</sub> system, the ground state is paramagnetic for reasonable values of the electronic interactions, and thus reduces to the Stern and Howard<sup>3</sup> theory in this case.

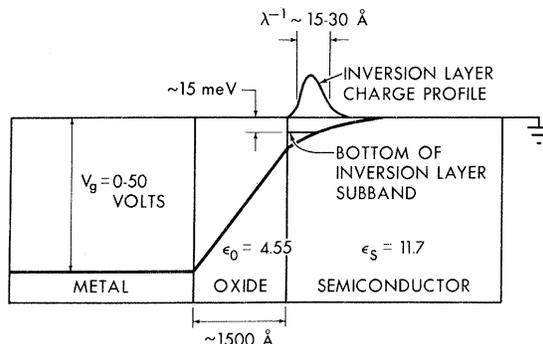


FIG. 1. Metal-oxide-semiconductor system with typical dimensions, including those of the inversion-layer charge-density profile. The electrostatic potential energy of a test charge is also included schematically.

The plan of the paper is as follows: in Sec. II, we describe the variational calculation of the paramagnetic ground state that incorporates elements of the existing theories. We further indicate what small differences occur in the theory, at the present level of sophistication, if a nonparamagnetic ground state exists. We conclude this section with a brief mention of the failure of two possible extensions of the simple theory to account for the lower occupation degeneracy of the valleys. In Sec. III, we describe a self-consistent unrestricted Hartree-Fock calculation of ground state of a Si-SiO<sub>2</sub> inversion layer as exemplified by the Si(111)-SiO<sub>2</sub> system. This calculation leads to a phase diagram as a function of the electronic interactions viewed as free parameters, so in Sec. IV we discuss the regions of the phase diagram that are experimentally accessible, and compare the results with the data. In Sec. V we discuss applications of the theory to the Si(100)-SiO<sub>2</sub> and Si(110)-SiO<sub>2</sub> systems. In Sec. VI we comment on the limitations of the present theory and suggest a range of experiments that will test it, shedding light, in particular, on the intervalley exchange interaction in silicon.

## II. PARAMAGNETIC THEORY OF THE INVERSION LAYER

In the presence of a weakly perturbing potential  $U(\vec{r})$ , such as in the neighborhood of a phosphorous impurity atom in a silicon lattice, Luttinger and Kohn<sup>2</sup> showed that the impurity electron wave function could be written

$$\psi(\vec{r}) = \sum_{\text{valley } v} \alpha_v F_v(\vec{r}) \phi_v(\vec{r}), \quad (1)$$

where  $\phi_v(\vec{r})$  is the full Bloch function at the  $v$ th conduction-band minimum along the  $\Gamma\Delta X$  axis of the bulk Brillouin zone— $v$  sums over the six valleys.  $F_v(\vec{r})$  is a relatively smooth envelope function which satisfied the effective-mass Schrödinger equation

$$[\epsilon_v(-i\vec{\nabla}) + U(\vec{r})]F_v(\vec{r}) = EF_v(\vec{r}), \quad (2)$$

where  $\epsilon_v(\vec{k})$  is the second-order expansion of the conduction band around the  $v$ th minimum. In Eq. (1) the  $\alpha_v$  determine the linear combination of valleys that go into the full wave function. A crucial assumption in obtaining Eqs. (1) and (2) is that  $U(\vec{r})$  is smooth on the scale of a unit cell dimension in silicon.

Stern and Howard<sup>3</sup> discussed the solution of Eq. (2) in the inversion layer geometry where  $U(\vec{r})$  is now the self-consistent potential solution of Poisson's equation on the Si side of the Si-SiO<sub>2</sub> interface. We outline an equivalent variational calculation as the fully self-consistent solutions of Eq.

(2) can only be obtained numerically.<sup>4</sup>

We assume that the  $\alpha_v$  contain the normalization of  $\psi(\vec{r})$ , and that  $F(\vec{r})$  has the form, for the silicon half space  $z > 0$ ,

$$F(\vec{r}, \lambda, \vec{k}_{\parallel}) = (1/\sqrt{\mathcal{Q}})(2\lambda^{3/2})_z e^{-\lambda z} e^{i\vec{k}_{\parallel} \cdot \vec{r}}. \quad (3)$$

Here  $\lambda$  is the variational parameter that gives the over-all envelope width;  $\vec{k}_{\parallel}$  refers to the (small) value of the crystal momentum parallel to the surface for the inversion-layer subbands;  $\mathcal{Q}$  is the cross-section area of the oxide-semiconductor interface. In fact,  $F(\vec{r})$  has a further  $z$ -dependent phase factor which produces a correction to the effective-mass parameters in the equations below [see Eqs. (8)–(13) of Ref. 3 for details].

With a total number density of  $n$  electrons per unit area in the inversion layer (and also at the metal-oxide interface), the electrostatic potential set up by a charge profile described in Eq. (3) is, for  $z > 0$ ,

$$\Phi(z) = \frac{-4\pi n e}{\epsilon_s} 4\lambda^3 e^{-2\lambda z} \left( \frac{z^2}{4\lambda^2} + \frac{z}{2\lambda^3} + \frac{3}{8\lambda^4} \right), \quad (4)$$

where  $\epsilon_s$  is the static dielectric constant for the semiconductor. We ignore the fine structure on the scale of a unit cell introduced by the Bloch function in Eq. (1). We match this potential and its normal derivative across the oxide interface. The oxide layer has a thickness  $\delta$  and static dielectric constant  $\epsilon_o$ . We set the potential at the metal-oxide interface to the gate voltage  $V_g$  as we have assumed that the potential vanishes at  $z \rightarrow +\infty$ . We thus obtain a relation between the gate voltage, the number density and the profile constant  $\lambda$ :

$$n = \frac{V_g}{4\pi e(\delta/\epsilon_o + 3/2\lambda\epsilon_s)}. \quad (5)$$

This result is trivially the capacitance relation obtained if all the inversion-layer electrons were at a distance  $1.5\lambda^{-1}$  from the oxide interface.

We proceed to calculate the Hartree total energy of the inversion-layer electrons. The one-electron term, relative to the conduction-band minimum at  $z \rightarrow \infty$ , is

$$E_1 = \sum_{\substack{\text{occupied} \\ \vec{k}_{\parallel} \text{ states}}} \langle \psi(\vec{r}) | \epsilon(-i\vec{\nabla}) + |e| \Phi(z) | \psi(\vec{r}) \rangle. \quad (6)$$

Provided  $F(\vec{r})$  is smooth on the unit-cell scale, we can exploit the Bloch function orthogonality at the different valleys and obtain

$$E_1 = \sum_{\substack{\text{valleys } v \\ \vec{k}_{\parallel} \text{ occupied}}} |\alpha_v|^2 \langle F_v(\vec{r}) | \epsilon_v(i\vec{\nabla}) + |e| \Phi(z) | F_v(\vec{r}) \rangle. \quad (7)$$

If  $\theta_{ij}^v$  is the  $ij$ th element of the inverse effective-mass tensor at the  $v$ th valley, then

$$E_1 = \frac{\hbar^2}{2} \sum_{k_{\parallel} \text{ occupied}} |\alpha_v|^2 \times (\theta_{xx}^v k_x^2 + \theta_{yy}^v k_y^2 + 2\theta_{xy}^v k_x k_y + \theta_{zz}^v \lambda^2) + \frac{15\pi n^2 e^2}{8\epsilon_s \lambda} \mathcal{G}. \quad (8)$$

In the last term we have summed over the  $n$  occupied states per unit area. We note here that the longitudinal-transverse mass mixing introduced by a  $z$ -dependent phase factor in Eq. (2) requires us to replace the bulk  $\theta_{xx}$  with  $(\theta_{xx} - \theta_{xz}^2/\theta_{zz})$ ,  $\theta_{yy}$  with  $(\theta_{yy} - \theta_{yz}^2/\theta_{zz})$  and  $\theta_{xy}$  with  $(\theta_{xy} - \theta_{xz}\theta_{yz}/\theta_{zz})$ . (Again, see Ref. 3.) We can now write

$$\frac{E_1}{\mathcal{G}} = \frac{\pi \hbar^2 n^2}{2\nu m_d} + \frac{\hbar^2 n \lambda^2}{2m_3} + \frac{15\pi n^2 e^2}{8\epsilon_s \lambda}, \quad (9)$$

where we have assumed  $\theta_{zz}^v = \theta_{zz} = 1/m_3$ , valid when all occupied valleys are equivalent, and where  $m_d = (\theta_{xx}\theta_{yy} - \theta_{xy}^2)^{-1/2}$  is the density-of-states mass of each of the  $\nu$  occupied valleys. We obtain the electron-electron contribution to the total energy by an indirect means. The Hartree total energy can be written as the sum of kinetic energies, the external potential and one-half of the total electron-electron interaction. It is also written as the sum over the kinetic energies and the self-consistent potential contribution minus half the electron-electron terms. Since the external potential is merely the gate voltage  $V_g$  we arrive at the total Hartree energy per unit area as

$$\frac{E_T}{\mathcal{G}} = \frac{\pi \hbar^2 n^2}{2\nu m_d} + \frac{\hbar^2 n \lambda^2}{2m_3} - \frac{1}{2} n e V_g + \frac{15\pi n^2 e^2}{16\epsilon_s \lambda}. \quad (10)$$

We minimize  $E_T$  with respect to  $\lambda$ , subject to the constraint in Eq. (5). For typical metal-oxide-semiconductor devices<sup>8,9</sup>  $\delta \sim 1400$  Å, and with  $\epsilon_s = 11.7$  and  $\epsilon_0 = 4.55$ ,  $\lambda^{-1}$  is of the order of 15–30 Å and the smoothness assumption for  $F(\vec{r})$  is barely satisfied. For such values of  $\lambda$ , Eq. (5) provides an almost linear variation of  $n$  with  $V_g$ , and in the linear limit  $\lambda$  varies as  $(n)^{1/3}$  for the minimized value of  $E_T$ . The dependence of  $\lambda$  and  $n$  on  $m_d$  and  $\nu$  is very weak with the consequence that, in the cases of more complex ground states where  $\nu$  and  $m_d$  both change from their paramagnetic values, the effect on  $\lambda$  and the corresponding electron-electron terms calculated in Sec. IV is negligible. In Sec. III, it is precisely  $\nu$  and  $m_d$  that are determined. By contrast, the dependence of  $\lambda$  and the one-electron energy levels derived from Eq. (10) on the value of  $m_3$  is very sensitive. For this reason, valleys with different  $m_3$  are split well apart and the  $\nu$  and  $m_d$  need be obtained only for those valleys with the largest  $m_3$ .

We calculated the intervalley Coulomb and ex-

change integrals for the interaction of two electrons with wave functions described by Eq. (1). We then performed a variational calculation on the  $\alpha_v$  coefficients. The intervalley exchange terms contribute to energy-level splittings if no symmetries are broken, and these exchange terms are an order of magnitude too small to alter the valley occupancy, even though the degeneracy is lifted. A further generalization of the Kohn-Luttinger<sup>2</sup> theory along the lines proposed by Gubanov and Davydov<sup>13</sup> was also examined. At the (111) surface the six valleys project onto different parts of the surface Brillouin zone [see Fig. 2(a)], so equivalent valleys with different occupations can only be obtained from symmetry-breaking distortions of the type introduced by density waves.

### III. MANY-BODY THEORY OF THE INVERSION-LAYER GROUND STATE

If electron correlations can alter the ground state in the layered compounds, we might expect similar effects in our almost two-dimensional system. We present here a two-dimensional many-body calculation of the possible ground states as a function of the following four electronic interaction energies: (i)  $V_1$ , the intravalley direct interaction which is dominated by the Coulomb re-

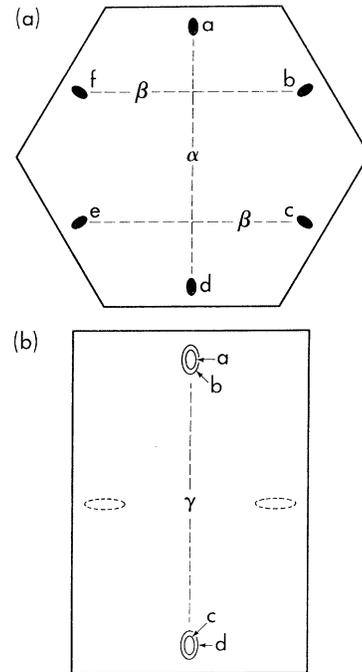


FIG. 2. Surface Brillouin zones for Si at (a) the (111) surface showing the six equivalent valleys, and (b) the (110) surface where the four valleys on the vertical axis are equivalent. The density-wave valley coupling for  $DW\alpha$ ,  $DW\beta$ , and  $DW\gamma$  are also marked.



The self-consistent solutions of Eq. (15) are much too many for an exhaustive study. Even the much smaller secular equation considered by Penn<sup>16</sup> had a wealth of structure in the resulting phase diagrams. We have confined ourselves to the examination of a number of less exotic self-consistent solutions of Eq. (15) which minimize the total energy [Eq. (13)]. These solutions are characterized by nonvanishing values of the correlations and occupations defined in Eq. (12). To begin with, we have not examined any possible mixed spin-charge-density wave solutions which have nonvanishing values of  $\rho_x$  or  $\tau_{xy}$  in the secular equation. Partially ferromagnetic solutions are generally unstable in simpler two-dimensional problems, and we have assumed this to be the case here.

We now summarize the salient features of some solutions:

(i) At small values of  $U$  and  $V$ , the paramagnetic solution is stable as it minimizes the kinetic-energy contribution to the total energy. The six pockets are equally occupied with spin-up and -down electrons, and all correlations vanish. The total energy per unit area takes the form

$$E_T/\mathcal{Q} = n^2 [(\pi \hbar^2 / 12 m_d) + (\frac{5}{12} U_1 + \frac{1}{24} V_1 - \frac{1}{24} U_{2\alpha} - \frac{1}{6} U_{2\beta}) \mathcal{Q}]. \quad (16)$$

(ii) At very large values of  $U$  and  $V$ , the stable solution is the ferromagnetic occupation of one valley which, although it maximizes the kinetic-energy contribution to the total energy, allows the electron-electron contribution to vanish. The total energy per unit area is now

$$E_T/\mathcal{Q} = \pi \hbar^2 n^2 / m_d. \quad (17)$$

In both these solutions the effective mass  $m^*$  is the density of states mass of an isolated pocket, i.e.,  $m^* = m_d = 0.358 m_e$ .

(iii) At intermediate values of  $U_1$ , ordinary charge-density waves (CDW) are stable for  $U_2 < 0$ , while spin-density waves (SDW) are stable for  $U_2 > 0$ . There are two types of either density wave, DW $\alpha$  and DW $\beta$ , driven, respectively, by  $U_{2\alpha}$  and  $U_{2\beta}$ . The DW $\alpha$  solution couples two opposite valleys pushing one below the Fermi level; the other four valleys remain uncoupled and unoccupied. The DW $\beta$  state couples four valleys into two pairs, pushing two below the Fermi level while the remaining two valleys are uncoupled and unoccupied [see Fig. 2(a)].

(iv) The DW $\alpha$  states have an occupational multiplicity  $\nu = 1$  and are defined by

$$\Delta_{ad\uparrow} = \pm \Delta_{ad\downarrow} = \frac{1}{4} n \mathcal{Q}; \quad A_\sigma = D_\sigma = \frac{1}{4} n \mathcal{Q}, \quad (18)$$

with all other occupations and correlations van-

ishing. The total energy per unit area for the CDW $\alpha$  solution ( $\Delta_{ad\uparrow} = \Delta_{ad\downarrow}$ ) is

$$E_T/\mathcal{Q} = n^2 [(\pi \hbar^2 / 2 m_d) + (\frac{1}{8} U_1 + \frac{1}{8} V_1 + \frac{1}{8} U_{2\alpha}) \mathcal{Q}], \quad (19)$$

and for the SDW $\alpha$  solution ( $\Delta_{ad\uparrow} = -\Delta_{ad\downarrow}$ ) is

$$E_T/\mathcal{Q} = n^2 [(\pi \hbar^2 / 2 m_d) + (\frac{1}{8} U_1 + \frac{1}{8} V_1 - \frac{1}{8} U_{2\alpha}) \mathcal{Q}], \quad (20)$$

with  $m^* = m_d = 0.358 m_e$ .

(v) The DW $\beta$  states have an occupation multiplicity  $\nu = 2$ , and are defined by

$$\begin{aligned} \Delta_{bf\uparrow} = \Delta_{ce\uparrow} = \pm \Delta_{bf\downarrow} = \pm \Delta_{ce\downarrow} &\cong \frac{1}{8} n \mathcal{Q} \\ - \left( \sum_k F_k [\epsilon_b(k) - \epsilon_f(k)]^2 \right) &[8(U_1 + 2W)^2 n^2 \mathcal{Q}^2]^{-1}, \\ B_\sigma = C_\sigma = E_\sigma = F_\sigma &= \frac{1}{8} n \mathcal{Q}, \end{aligned} \quad (21)$$

where (i)  $W = 0$  and  $\Delta_{bf\uparrow} = -\Delta_{bf\downarrow}$ , etc., for the SDW $\beta$  solution, i.e.,  $U_{2\beta} > 0$ , and (ii)  $W = -U_{2\beta}$  and  $\Delta_{bf\uparrow} = \Delta_{bf\downarrow}$ , etc., for the CDW $\beta$  solution when  $U_{2\beta} < 0$ . Again, all other occupations and correlations vanish. We can write the total energy per unit area for CDW $\beta$  as

$$E_T/\mathcal{Q} = n^2 [(\pi \hbar^2 / 4 m^*) + (\frac{5}{16} U_1 + \frac{1}{16} V_1 - \frac{1}{16} U_{2\alpha}) \mathcal{Q}], \quad (22)$$

and for SDW $\beta$  as

$$E_T/\mathcal{Q} = n^2 [(\pi \hbar^2 / 4 m^*) + (\frac{5}{16} U_1 + \frac{1}{16} V_1 - \frac{1}{16} U_{2\alpha} - \frac{1}{8} U_{2\beta}) \mathcal{Q}], \quad (23)$$

with the important fact that  $m^*$  here is now a function of  $U$  and  $V$  as described in subsection (viii) below.

(vi) In Fig. 3 we present the phase diagram for the Si(111)-SiO<sub>2</sub> system as a function of  $U_1$  and  $U_{2\beta}$  in the case where  $U_1 = V_1$  (Sec. IV), and for three fixed ratios of  $U_{2\alpha}/U_{2\beta}$ . We recall that  $U$  and  $V$  are functions of  $\lambda$  and  $n$ , but if viewed as independent parameters the total energy scales as ( $n^2$ ).

(vii) We have tested a number of other types of solutions with the following conclusions:

(a) The DW solutions above, which can be characterized by a single  $\vec{q}$ -vector correlating electrons, are stable with respect to the onset of a second DW until  $U$  and  $V$  are very large, by which time the ferromagnetic solution is stable.

(b) The triple DW's defined by

$$\Delta_{ac\uparrow} = \Delta_{bd\uparrow} = \Delta_{ce\uparrow} = \Delta_{df\uparrow} = \Delta_{ea\uparrow} = \Delta_{fb\uparrow} \neq 0, \quad \Delta_{xy\uparrow} = \pm \Delta_{xy\downarrow}, \quad (24)$$

are such that SDW cannot be stabilized at all, and the CDW is always unstable with respect to the existing solutions above.

(c) At large values of  $U$  and  $V$ , a complex mixed spin-charge ferromagnetic solution with  $\nu = 1$  can be obtained. The state is defined by

$$\Delta_{xy\uparrow} = n\mathcal{Q} \left\{ \frac{1}{6} - O\left[\frac{\epsilon(k)}{U}\right]^2 \right\}, \quad (25a)$$

where the second term is a correction of order  $[\epsilon(k)/U]^2$

$$\begin{aligned} \Delta_{xy\uparrow} &\equiv \Delta_{xy\uparrow\uparrow} \equiv 0, \\ X_{\uparrow} &= \frac{1}{6}n\mathcal{Q}, \quad X_{\uparrow} = X_{\uparrow\uparrow} \equiv 0, \end{aligned} \quad (25b)$$

but remains unstable with respect to the ferromagnetic occupation of one isolated pocket [solution (ii) above].

(viii) We have calculated cyclotron masses for the different solutions by determining the areas of the Fermi distribution and their derivatives with respect to the Fermi energy. For the paramagnetic and ferromagnetic solutions  $m^* = 0.358$ , the

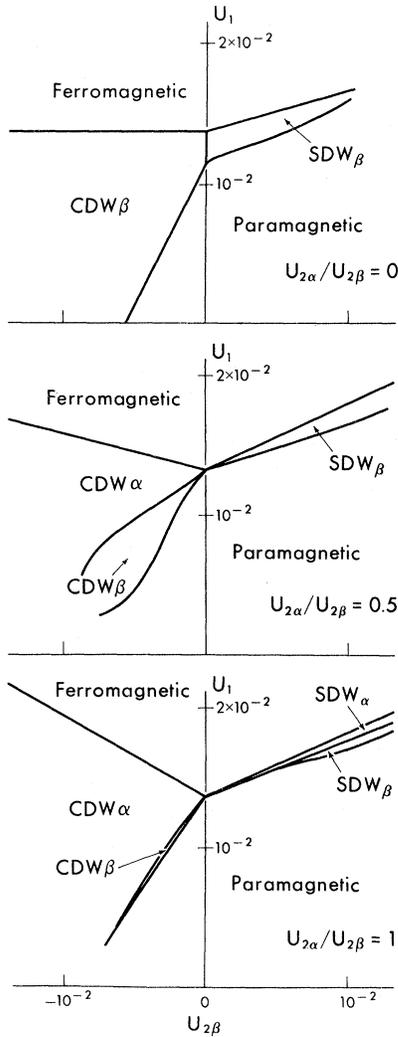


FIG. 3. Phase diagram for the Si(111)-SiO<sub>2</sub> inversion-layer ground state. In all cases  $U_1 = V_1$  and the three diagrams correspond to the ratios  $r = U_{2\alpha}/U_{2\beta}$  of (a)  $r = 0$ , (b)  $r = 0.5$ , (c)  $r = 1$ . Axes in eV for normalizing area  $\mathcal{Q} = 10^{-12} \text{ cm}^{-2}$ .

value for an isolated valley; this is also the value for the DW $\alpha$  solution where the off-diagonal matrix elements in Eq. (15) split the  $\epsilon_a(k)$  and  $\epsilon_d(k)$  bands rigidly apart. For the DW $\beta$  solutions, the effective masses vary with  $(U_1 - 2U_{2\beta})$  for the CDW $\beta$  state and with  $U_1$  for the SDW $\beta$  state. In Fig. 4, we plot the lines of constant effective mass on the assumption that DW $\beta$  is stable over the entire phase diagram. Of course, this figure must be used with the appropriate graph in Fig. 3 in order to obtain the actual range of variation of the effective masses for the CDW $\beta$  and SDW $\beta$  ground state.

#### IV. PHYSICALLY ACCESSIBLE REGION OF THE PHASE DIAGRAM

In Sec. II, we obtained the values of the number density per unit area  $n$ , and the profile constant  $\lambda$ , as functions of the gate voltage  $V_g$ . We now proceed with these values to determine the behavior of the  $U$ 's and  $V$ 's of Sec. III as a function of  $V_g$ . We note that the intravalley direct interaction is dominated by the Coulomb interaction between two electrons in the same valley, with an equivalent result holding for the direct intervalley interaction. We thus calculate

$$W_{xy}(\vec{q}) = \int \int v(\vec{r}_{12}) e^{i\vec{q}\cdot\vec{r}_{12}} |\psi_x(\vec{r}_1)|^2 |\psi_y(\vec{r}_2)|^2 d\vec{r}_1 d\vec{r}_2, \quad (26)$$

where  $x$  and  $y$  now refer to two valleys, the  $\psi$  have the form of a Bloch function multiplied by the

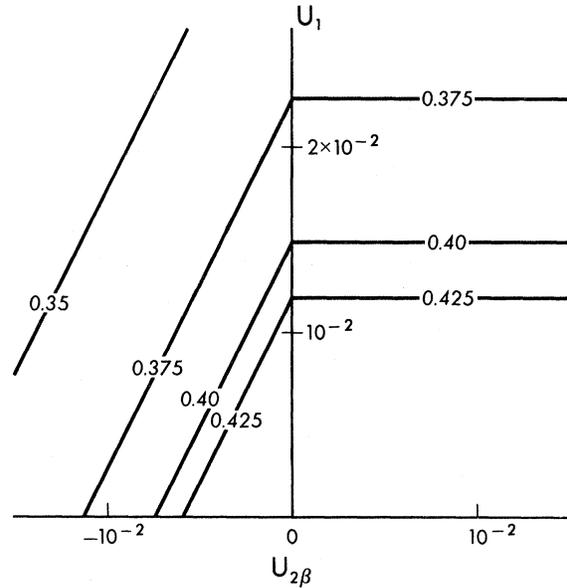


FIG. 4. Lines of constant cyclotron mass for the DW $\beta$  solution. When used with Fig. 3, the ground-state cyclotron mass variation can be obtained.

smooth envelope function  $F$  [Eq. (3)], and with  $\vec{r}_{12} = \vec{r}_1 - \vec{r}_2$ ,  $v(\vec{r}_{12})$  is the appropriately screened Coulomb interaction

$$v(r_{12}) = e^2 e^{-sr_{12}} / r_{12} \epsilon_s, \quad r_{12} = |\vec{r}_{12}|. \quad (27)$$

The parameters of Sec. III are just

$$U_1 = W_{xy}(\vec{q}=0) \quad \text{and} \quad V_1 = W_{xx}(\vec{q}=0). \quad (28)$$

Because of the relative smoothness of  $F(\vec{r})$ , we can perform the  $\vec{r}_1$  and  $\vec{r}_2$  integrations independently over each unit cell using average values for  $v(r_{12})$ , and  $F(\vec{r}_1)$  and  $F(\vec{r}_2)$ . The Bloch functions integrate to the corresponding normalization constant, so that in this scheme we obtain

$$U_1 = V_1 = \frac{e^2 \pi}{\mathcal{Q} \epsilon_s} \left( \frac{128 \lambda^6}{s(4\lambda^2 - s^2)^3} + \frac{\lambda(1.5s^4 - 20s^2\lambda^2 + 120\lambda^4)}{(s^2 - 4\lambda^2)^3} \right) \\ = \frac{e^2 \pi}{\mathcal{Q} \epsilon_s} \frac{\lambda}{s} \frac{16\lambda^2 + 9\lambda s + 1.5s^2}{(2\lambda + s)^3}. \quad (29)$$

In our quasi-two-dimensional, low-electron-density problem, the conventional approaches<sup>17</sup> to screening are of limited validity. In (29),  $s$  is the inverse screening length in the inversion layer but due to the whole MOS system. For  $s \sim 0.06 \text{ \AA}^{-1}$  (a phenomenological choice of the order of the inverse interelectronic spacing),  $U_1$  increases from  $4 \times 10^{-3}$  to  $8 \times 10^{-3}$  eV as the gate voltage increases from 4 to 40 V. In (29) we have chosen  $\mathcal{Q} = 10^{-12} \text{ cm}^2$ .

The intervalley electron-electron exchange contribution to  $U_2$  is positive, but is small for two reasons. First the integral is now of the form

$$W_{xy}(q) = \int \int v(r_{12}) e^{i\vec{q} \cdot \vec{r}_{12}} \psi_x^*(\vec{r}_1) \\ \times \psi_y(\vec{r}_1) \psi_y^*(\vec{r}_2) \psi_x(\vec{r}_2) d\vec{r}_1 d\vec{r}_2,$$

and the trick of integrating the periodic parts of the Bloch functions over unit cells contributes overlap terms as factors. Secondly, we need evaluate this integral for  $|\vec{q}|$  equal to the intervalley  $\vec{k}$  vector, and  $W$  falls off as  $(\lambda/q)^2$  for large  $q$ —here  $\lambda/q \sim \frac{1}{20}$ .

The intervalley exchange mediated by the electron-phonon interaction is known to be large and negative, and may even be greater in absolute value than the direct Coulomb interaction.<sup>14</sup> Of more importance to us is the ratio  $U_{2\alpha}/U_{2\beta}$  which determines the amount of phase space available to the DW $\beta$  states. A relatively crude estimate of  $U_{2\beta}$  will suffice to determine the ground state if  $U_{2\alpha}/U_{2\beta}$  is somewhat less than unity.

From the form of the electron-phonon-electron intervalley exchange (Eqs. II.15 and II.22 of Ref. 14), we see that  $U_2$  varies as  $(\omega_{\Delta k})^{-2}$ , where  $\omega_{\Delta k}$  is the appropriate phonon frequency of wave vector

equal to the intervalley  $k$  vector. By using both the selection rules of Lax and Hopfield<sup>18</sup> and the phonon spectrum of Martin,<sup>19</sup> we do indeed estimate  $U_{2\alpha}/U_{2\beta}$  to be somewhat less than unity. We cannot be more precise as assumptions concerning the equality of deformation potentials<sup>14</sup> and the operation of bulk selection rules near the interface constitute gross approximations.

It is clear that for values of  $U_1$  and  $-U_{2\beta}$  between  $4 \times 10^{-3}$  and  $8 \times 10^{-3}$  eV, and for  $U_{2\alpha}/U_{2\beta} \lesssim 0.5$ , we are in the CDW $\beta$  regime for the experimental range of gate voltages. In this regime the valley degeneracy is two and the effective mass decreases slowly from 0.4 as the gate voltage is increased. All these facts agree with experiment. If the screening parameter  $s$  were to increase significantly with  $n$ , then  $U_1$  could decrease and  $m^*$  increase with gate voltage, contrary to experiment. Further experimental tests of our theory are discussed in Sec. VI.

## V. OTHER SYSTEMS

A four-valley Hamiltonian can be set up for the Si(110)-SiO<sub>2</sub> system—see Fig. 2(b) for the position of the valleys in the Brillouin zone. Here the two-dimensional effective-mass tensors for the four heavy-normal-mass valleys are identical, and the possible density-wave solutions, while they alter the occupied valley degeneracy, do not affect the effective masses. The off-diagonal matrix elements in the secular equation merely split the bands rigidly apart. In Fig. 5 we present the phase diagram of the Si(110)-SiO<sub>2</sub> system with the ratio  $U_{2\alpha}/U_{2\beta} = 0.5$ . The DW $\gamma$  solutions couple the four valleys in all combinations allowed by the nonzero  $q$ -vector labeled  $\gamma$  in Fig. 2(b). The resulting valley occupancy is unity. There is also a regime in which the paramagnetic occupation of two valleys

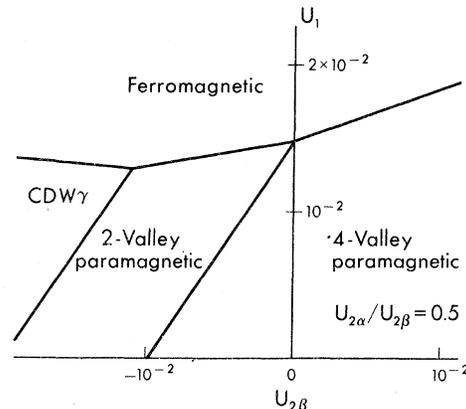


FIG. 5. Phase diagram for the Si(110)-SiO<sub>2</sub> inversion layer ground state, for  $U_1 = V_1$  and the ratio  $U_{2\alpha}/U_{2\beta} = 0.5$ . Again axes are in eV for normalizing area  $\mathcal{Q} = 10^{-12} \text{ cm}^2$ .

(with a  $U_{2\alpha}$  exchange between them) is stable. The effective mass remains at  $0.342m_0$ . The range of  $U$ 's and  $V$ 's is similar to that for the Si(111)-SiO<sub>2</sub> system, so that again, provided that  $|U_{2\beta}| \gtrsim U_1$ , we are in the regime of the two-valley paramagnetic solution.

It is important to realize that this state of asymmetrical occupation of the valleys, which remain paramagnetic in character, is very similar in structure to the ferromagnetic state. The interaction terms are strong enough to unbalance valley occupation (as opposed to spin occupation) and break the symmetry accordingly. Although the resulting state is not magnetic, many of the properties of this state can be considered a ferromagneticlike extension of symmetry breaking in a many-valley system.

For the Si(100)-SiO<sub>2</sub> system, the cubic symmetry of Si, combined with the anisotropy in the effective-mass expansions of the valley, make the paramagnetic occupation of the two high-normal-mass valleys the stable solution up to high gate voltages, where the ferromagnetic solution ultimately becomes stable. For this system, the existing theories<sup>3,4</sup> apply as do the normal many-body effects.<sup>20</sup>

## VI. DISCUSSION

We have presented a symmetry-breaking many-electron theory that accounts for all the occupied-valley-degeneracies for the Si-SiO<sub>2</sub>  $n$ -type inversion layers. The cyclotron effective masses have been explained for the Si(111)-SiO<sub>2</sub> system, but not the Si(110)-SiO<sub>2</sub> or the Si(100)-SiO<sub>2</sub> systems. In the latter cases, where the solutions are paramagnetic in our model, a conventional many-body approach<sup>20</sup> must be adopted to account for the (10–20)% increase in cyclotron effective mass. Our explanation for the Si(111)-SiO<sub>2</sub> system does

not require us to make very large random stresses<sup>10</sup> to account for the reduced valley occupation. Our CDW $\beta$  solution reduces the  $C_{3v}$  symmetry of this surface to mirror symmetry. Domains of CDW $\beta$  solutions, each with just one of the possible  $q$  vectors, should appear and, for random distribution of such domains, the transport properties will be isotropic.<sup>10</sup> The observation of anisotropic conductivity effects under uniaxial stress<sup>10</sup> is completely in accord with a CDW $\beta$  solution as just one density-wave domain is preferentially driven. We are at present calculating the optical absorption as electrons are excited from the bonding to antibonding CDW $\beta$  bands. Fuller details of the anisotropy of the absorption, its line shape and its dependence on gate voltage will be published elsewhere. Such transitions would not occur in a random stress model of the Si(111)-SiO<sub>2</sub> system.

The weakest point of our theory is the evaluation of the intervalley exchange terms. Should any evidence for phase transitions in the inversion layers be seen experimentally we would be able to estimate both  $U_{2\alpha}$  and  $U_{2\beta}$  by comparison with our phase diagrams.

Finally, reduced "valley" occupation and the higher effective masses observed in inversion layers of tellurium<sup>11</sup> probably owe their origin to density-wave ground states, but the electronic structure of Te is not so well characterized, either theoretically or experimentally, as to attempt a further analysis at this stage. The recent observation<sup>21</sup> of Shubnikov-de Haas oscillations and cyclotron resonance at a germanium-insulator interface offers further possibilities of testing our theory. The anisotropy of the valleys is much greater than in silicon, there are fewer valleys and at different parts of the surface Brillouin zone.<sup>3</sup> Quite different phase diagrams are to be expected.

\*Supported by the IBM Corporation through its postdoctoral Fellowship program. Permanent address: Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, England.

†Work supported in part by the NSF through Grant No. DMR72-03106-A02.

<sup>1</sup>J. R. Schrieffer, in *Semiconductor Surface Physics*, edited by R. H. Kingston (University of Pennsylvania, Philadelphia, 1957), p. 57.

<sup>2</sup>J. M. Luttinger and W. Kohn, *Phys. Rev.* **97**, 869 (1955).

<sup>3</sup>F. Stern and W. E. Howard, *Phys. Rev.* **163**, 816 (1967).

<sup>4</sup>F. Stern, *Phys. Rev. B* **5**, 4891 (1972).

<sup>5</sup>G. Landwehr, *Festkörperprobleme XV* (Pergamon, New York, 1975), p. 49.

<sup>6</sup>F. F. Fang and P. J. Stiles, *Phys. Rev.* **174**, 823

(1968).

<sup>7</sup>D. C. Tsui and G. Kaminsky, *Phys. Rev. Lett.* **35**, 1468 (1975).

<sup>8</sup>T. Neugebauer, K. von Klitzing, G. Landwehr, and G. Dorda, *Solid State Commun.* **17**, 295 (1975).

<sup>9</sup>A. A. Lakhani and P. J. Stiles, *Phys. Lett. A* **51**, 117 (1975).

<sup>10</sup>D. C. Tsui and G. Kaminsky, *Solid State Commun.* **20**, 93 (1976).

<sup>11</sup>K. von Klitzing and G. Landwehr, *Solid State Commun.* **9**, 2201 (1971); and R. Silberman, G. Landwehr, J. C. Thuillier, and J. Bouat, *Jpn. J. Appl. Phys. Suppl.* **2**, part 2, 359 (1974).

<sup>12</sup>J. A. Wilson, F. J. di Salvo, and S. Mahajan, *Adv. Phys.* **24**, 117 (1975).

<sup>13</sup>A. I. Gubanov and S. Yu Davydov, *Fiz. Tekh. Polu-*

provodn. 6, 1677 (1972) [Sov. Phys.-Semicond. 6, 1448 (1973)].

<sup>14</sup>M. L. Cohen, Phys. Rev. 134, A511 (1964).

<sup>15</sup>Terms which involve four valleys, of the form  $a^\dagger b d^\dagger e$  for example, could be easily included. We have not considered them explicitly.

<sup>16</sup>D. R. Penn, Phys. Rev. 14, 350 (1966).

<sup>17</sup>F. Stern, Phys. Rev. Lett. 18, 546 (1967); P. B.

Visscher and L. M. Falicov, Phys. Rev. B 3, 2541 (1971).

<sup>18</sup>M. Lax and J. J. Hopfield, Phys. Rev. 124, 115 (1961).

<sup>19</sup>R. M. Martin, Phys. Rev. 186, 871 (1969).

<sup>20</sup>B. Vinter, Phys. Rev. B 13, 4447 (1976), and references therein.

<sup>21</sup>W. Weber, G. Abstreiter, and J. F. Koch, Solid State Commun. 18, 1397 (1976).