and $\sin\omega t\, e^{bt} f^{\eta-1}(t)$ can be expressed in the form

$$\sin\omega t\, e^{bt} f^{\eta-1}(t) = \frac{1}{i^{2\eta}} \sum_{m=0}^{\eta} \frac{2m-\eta}{\eta} \binom{\eta}{m} e^{i(2m-\eta)\omega t} e^{\eta bt}. \quad \text{(A7)}$$

By comparison of terms $\rho_0^{(L)}$ and $\rho_1^{(L)}$ which have the same time dependence, one observes that if condition (A1) is satisfied, $\rho_1^{(L)}$ can be neglected and so can all the terms $\rho_j^{(L)}$ for which $j>0$. ψ_L can therefore be approximated by a single Houston wave function;

$$\psi_L \cong \varphi_L\left(\mathbf{K}_0 - \frac{e\mathbf{A}}{\hbar}\right) e^{i\mathbf{K}_0\mathbf{R}} \exp\left[-\frac{i}{\hbar}\int^t \epsilon_L\left(\mathbf{K}_0 - \frac{e\mathbf{A}}{\hbar}\right) d\tau\right]. \quad \text{(A8)}$$

One may also expand $\varphi_L(\mathbf{K}_0 - e\mathbf{A}/\hbar)$ and the exponential function in powers of A_0. Using similar arguments it can be shown that under condition (A1) $\varphi(\mathbf{K} - e\mathbf{A}/\hbar)$ can be considered independent of time.

PHYSICAL REVIEW VOLUME 169, NUMBER 3 15 MAY 1968

Transport Properties of Electrons in Inverted Silicon Surfaces

F. F. FANG AND A. B. FOWLER
IBM Watson Research Center, Yorktown Heights, New York
(Received 29 December 1967)

Measurements of the effective mobility, field effect mobility, Hall mobility, and carrier density of Si as a function of field perpendicular to the surface are reported. At all temperatures from 4.2 to 300°K, at least one maximum in the mobility was observed. The temperature dependence is reported for different fields. At room temperature, a single maximum in the mobility was observed close to the threshold for inversion. As the temperature was lowered, this peak increased. At temperatures near 80°K, it then decreased. Another maximum appeared at about 100°K at higher fields; it increased as the temperature was lowered. An anomalous shift in the conductance threshold between 77.3 and 4.2°K is reported and is correlated with the charge in the oxide. Effects of substrate bias are reported. Some comments are made on possible scattering mechanisms. The effect of interface states was measured and their density near the conduction band is reported.

INTRODUCTION

FOR many years, studies have been made of the transport properties of carriers in the surface of semiconductors.[1] In general, these measurements were made by varying the electric field normal to the surface of a semiconductor so that a change of the space charge near the surface occurred and with it a change in the conductivity. Before 1962 these measurements were usually limited to small ranges of field and the surfaces were either accumulated (induction of majority carriers at the surface), depleted (induction of space charge by removal of majority carriers near the surface leaving charged ions), or slightly inverted (depletion with the additional induction of minority carriers at the surface). Usually measurements were limited to the conductance as a function of normal field or to the field effect mobility—the differential change of surface conductance with respect to total induced charge. Some Hall-effect and magnetoconductance measurements were also made.[2–4] In most cases, the surface fields were controlled by means of the absorption of ambient vapors.[5,6] In general, the induced charge did not exceed $3\times10^{11}/\text{cm}^2$.

The transport properties of electrons in the potential wells at the surface were studied intensively for several years following Schrieffer's[7] initial work in which he solved Boltzmann's equation in a surface well with the assumption that the scattering at the surface was diffuse. The results of such calculations were that the mobility was expected to decrease as the surface field increased. For a linear potential well and constant bulk scattering time, Schrieffer had shown that $\mu/\mu_B \simeq (kTm)^{1/2}/(q\tau_B F_z)$ for $\tau_B \gg \tau_S$, where k is the Boltzmann constant, T the temperature, m the effective mass, τ_B and τ_S the bulk and surface scattering times, respectively, and F_z the field just inside the surface. Most of the observed reduction of the carrier mobilities has been attributed to this type of scattering. Better approximations, using wells more exactly approximating ones resulting from a classical solution of the Boltzmann-Poisson equations and taking account of energy-dependent scattering times and specular and diffuse

[1] For a general review of this subject, see A. Many, Y. Goldstein, and N. B. Grover, *Semiconductor Surfaces* (North-Holland Publishing Co., Amsterdam, 1965), p. 64.
[2] J. N. Zemel and R. L. Petritz, Phys. Rev. **110**, 1263 (1958).
[3] R. E. Coovert, J. Phys. Chem. Solids **21**, 87 (1961).
[4] P. Handler and S. Eisenhour, in *Solid Surfaces*, edited by H. Gatos (North-Holland Publishing Co., Amsterdam, 1964).
[5] W. H. Brattain and J. Bardeen, Bell System Tech. J. **32**, 1 (1953).
[6] R. H. Kingston, J. Appl. Phys. **27**, 101 (1956).
[7] J. R. Schrieffer, Phys. Rev. **97**, 641 (1955).

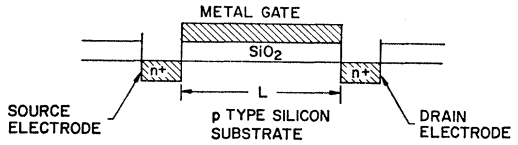

FIG. 1. An idealized cross section of a field effect device. The metal "gate" electrode is in all cases, reported in this paper, aluminum. The two n^+ regions, which serve as contacts to the channel that is induced in the silicon at the SiO_2-Si interface, are known as the source and drain. The operation of these contacts L is of the order of 10 μ. The oxide thickness δ is 1000–6000 Å. The diffused contacts are W wide perpendicular to the drawing.

scattering, resulted in somewhat modified dependence of μ on F_z.[8]

With the advent of the silicon planar technology[9] in the early 1960's it became possible to make structures in which a charge of more than $10^{13}/cm^2$ could be induced so that extreme inversion was possible.[10] A typical structure on which measurements could be made is shown in Fig. 1. Typically, thermally grown silicon oxide can sustain a field of 10^7 V/cm. When such high fields are available, it becomes possible to extend studies of transport properties over a much larger range. Furthermore, the oxide can be made so that polarization effects are negligible, can be made extremely uniform in properties, and can be made with low interface state density at least near the band edges. These factors reduce the ambiguity of the experimental results.

The extreme inversion range allows observation of quantum effects in the surface. Schrieffer[11] has discussed

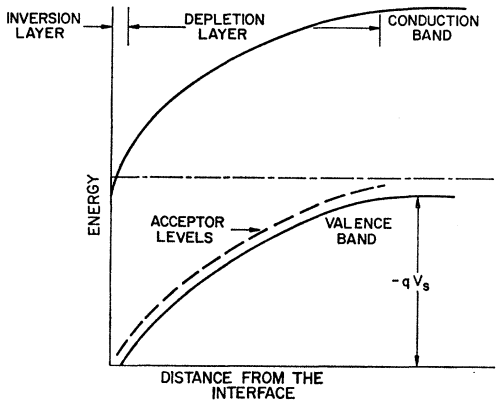

FIG. 2. The bands at the inverted surface of a semiconductor. The holes have been removed from the depleted region so that the space charge consists of negatively charged acceptors. Most of the voltage drop occurs across this region.

in 1956 the possibility that electronic states in the surface well (Fig. 2) would be quantized in the direction perpendicular to the surface but believed that the effects would be smeared out by inhomogeneity and thermal effects. Recent measurements[12,13] of oscillatory magnetoconductance have demonstrated that, in fact, the electrons are quantized in these surfaces perpendicular to the surface. These experiments revealed conductance oscillations as the Fermi level passes through successive Landau levels when the carrier density is increased by increasing the electric field. The spacing indicates that the density of states is constant, which corresponds to a case where the electron gas is two-dimensional. Measurements in the (100) surface indicated that the bands are twofold degenerate and have a mass parallel to the surface of about $0.2m_0$. This had been predicted[14] on the basis that the four bands corresponding to minima with light mass perpendicular to the surface would be raised to higher energy than the two bands with heavy mass perpendicular. In general, most aspects of a self-consistent application of the effective-mass theory in the surface well have been observed in these experiments. One germane result of these measurements is that oscillations could be observed for Landau splittings corresponding to a change of carrier density less than $5\times10^{10}/cm^2$, which sets an upper limit on the inhomogeneity.

At all of the fields corresponding to surface inversion, the conduction band is strongly quantized in the surface well. Self-consistent solutions of the Poisson equation and Schrödinger's equation[15] show that the splittings are from 10 to 100 meV between the minima of the lowest two-dimensional sub-band and the first excited sub-bands, where the splitting depends on substrate doping, surface field, surface orientation, and even temperature. For instance, for a (100) surface on a 12.5-Ω cm p-type substrate, both the fourfold degenerate ground sub-band and the first excited twofold degenerate sub-band are, respectively, approximately 22, 27, and 54 meV above the minima for the twofold degenerate ground sub-band for 10^{11}, 10^{12}, and 10^{13} electrons/cm^2 induced in the surface. The upper bands are occupied at high fields (for more than about 5×10^{12} electrons/cm^2) even at 0°K because the Fermi energy is larger than the splitting. At high temperatures, there should be a substantial population of electrons in higher sub-bands so that on this surface at 300°K (where kT is 25 meV) enough levels may be occupied so that the effects of quantization may be smeared out. For a (111) surface, where the splittings are about twice those for

[8] See, for example, R. F. Greene, J. Phys. Chem. Solids **14**, 143 (1960) (review); F. S. Ham and D. C. Mattis, IBM J. Res. Develop. **4**, 143 (1960); R. F. Greene, in *Surface Science*, edited by H. Gatos (North-Holland Publishing Co., Amsterdam, 1964), p. 101; R. F. Greene, D. R. Frankl, and J. Zemel, Phys. Rev. **118**, 967 (1960).

[9] See, for example, R. L. Petritz, Proc. IRE **50**, 1025 (1962); J. C. Haenichen, in *Integrated Circuits*, edited by R. M. Warner (McGraw-Hill Book Co., New York, 1965), Chap. 12.

[10] S. R. Hofstein and F. P. Heiman, Proc. IEEE **51**, 1190 (1963).

[11] J. R. Schrieffer, in *Semiconductor Surface Physics*, edited by R. H. Kingston (University of Pennsylvania Press, Philadelphia, Pa., 1957), p. 55.

[12] A. B. Fowler, F. F. Fang, W. E. Howard, and P. J. Stiles, Phys. Rev. Letters **16**, 901 (1966).

[13] A. B. Fowler, F. F. Fang, W. E. Howard, and P. J. Stiles, J. Phys. Soc. Japan Suppl. **21**, 331 (1966).

[14] F. F. Fang and W. E. Howard, Phys. Rev. Letters **16**, 797 (1966).

[15] F. Stern and W. E. Howard, Phys. Rev. **163**, 816 (1967); and private communication with W. E. Howard and F. Stern.

the (100) surface, most of the electrons will be in the ground state even at 300°K. At 4.2°K, it is not possible to populate the higher sub-bands even at the highest fields available because of the greater splittings and higher density of states.

In this paper, the results of a variety of measurements on the transport properties of electrons in inverted silicon surfaces are discussed. The results should properly be interpreted in terms of a quantum model but as yet only some of the possible scattering mechanisms have been studied theoretically.[15] As will be obvious from the complexity of the results, much more study will be required to understand these results.

MEASUREMENTS

Measurements of conductance and transconductance or field effect mobility were made over a large range of temperature and electric field. In addition, Hall measurements were made over a more restricted range.

The surface conductances were measured using circular coaxial field effect transistors. These samples have the advantage that the conducting channels are closed so that there are no end effects but the disadvantage that they cannot be used to observe anisotropy such as might be expected on, for instance, {110} surfaces.[16] Both ac and dc techniques were used and the results were consistent, but the ac method was generally used because of its simplicity. A small voltage signal (about 50 μV) was maintained between the diffused contacts (the source and drain). The voltage on the field plate (the gate) was varied slowly to ensure equilibrium in the semiconductor surface. The currents were measured through a small resistor or a low-impedance transformer (0.3 Ω) in the dc and ac cases, respectively. The source-drain voltages were kept small to avoid hot-electron effects and nonuniformity of the surface field across the sample. For the ac case, a lock-in amplifier was used as a detector.

The field effect mobility, which is defined as $\mu_{FE}=d\sigma_s/dQ_s$, where σ_s is the surface conductivity and Q_s is the total charge per unit area induced in the surface, is related to the transconductance G_m by

$$\mu_{FE}=\frac{L}{W}\frac{4\pi\delta}{\kappa_{ox}}\frac{G_m}{V_D}, \tag{1}$$

where W and L are the width and length of the channel path, δ is the thickness, κ_{ox} is the insulator dielectric constant, and V_D is the source-drain voltage (cgs units are used throughout). The transconductance $dI/dV_g|_{V_D}$ was measured by applying a small (50 μV to 100 mV) constant voltage V_D between the source and drain contacts and by superimposing a small ac signal on the gate or field plate voltage V_g. The signal current

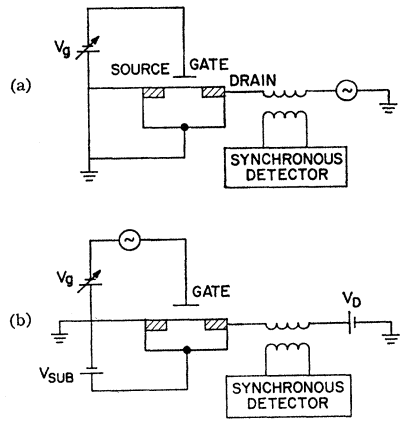

FIG. 3. (a) Circuit for measuring conductance and (b) circuit for measuring transconductance. A low-impedance (Freed 0.3 Ω) transformer is used for transconductance measurements. The resistance of the drain supply is 0.5 Ω.

was detected using a low-impedance transformer, as shown in Fig. 3. In general, the measurements were made at 500 Hz. No dependence on frequency was observed from 30 Hz to 150 kHz, however.

Hall measurements were made on samples of a van der Pauw geometry,[17] which have been described elsewhere.[18] The most important requirement for reliable Hall measurements on an inverted surface is that the surface not under the gate or field electrode should be nonconducting and that the leakage currents be small. This requirement can generally be met on n-channel structures if a proper annealing treatment either in air, nitrogen, or hydrogen[19,20] is used.

When the gate is so biased as to remove almost all of the free electrons from the surface, leakage currents (both on the surface outside the electrode and through the n^+-p junctions into the p-substrate region) can become comparable to the channel current. This problem becomes increasingly serious as the temperature increases because of the rapid increase of diode current with temperature. At 22°C, the leakage conductance below threshold in a good sample is less than 0.1 of that for a surface carrier density of about $10^{10}/cm^2$. Data are not reported for conductances between Hall probes less than 10 times the leakage conductance. Since samples showed a slight asymmetry, this was corrected using the van der Pauw correction formula[17] but if the asymmetry was large or occurred in the threshold, the data were not used. This was usually a problem only at low surface charge densities where inhomogeneity might be expected to be more serious.

[16] Such effects have been reported by Colman and Mize on p surfaces at the IEEE Solid State Device Research Conference, Santa Barbara, 1967 (unpublished).

[17] L. J. van der Pauw, Philips Res. Rept. **13**, 1 (1958).

[18] A. B. Fowler, F. F. Fang, and F. Hochberg, IBM J. Res. Develop. **8**, 427 (1964).

[19] G. Cheroff, F. Fang, and F. Hochberg, IBM J. Res. Develop. **8**, 416 (1964).

[20] E. Kooi, Phillips Res. Rept. **20**, 578 (1965); E. Kooi, IEEE Trans. Electron Devices **13**, 238 (1966); P. Balk, paper presented at the San Francisco meeting of the Electrochemical Society, Abstract No. 109, 1965 (unpublished).

The Hall measurements were made using an 180-V battery limited by resistances of from 10^{10} to 10^5 Ω. The voltages, measured with an electrometer, ranged from 10 mV to 100 mV. The usual eight measurements, with reversed currents and magnetic fields, were made for each value of carrier density and mobility at a given electric field.

In the best cases, samples with surface conductances as low as 3×10^{-6} mho for a square could be measured.

The effective mobility can be inferred from the conductance G_{SD},

$$\mu_{\text{eff}}=\frac{L}{W}\frac{4\pi\delta}{\kappa_{\text{ox}}}\frac{G_{SD}}{V_g'}, \tag{2}$$

where V_g' is the gate voltage above threshold for conduction. When Hall measurements on similarly prepared samples indicated that fast interface states were insignificant, the effective mobility was assumed to be equal to the conductivity mobility. At high temperatures, the threshold voltage is not well defined because the transition from depletion to inversion is not sharply defined. At temperatures below 80°K, Hall data were necessary to define the threshold as explained in the section on low-temperature results.

The field effect mobility is given, in general, by

$$\mu_{\text{FE}}=\frac{4\pi}{\kappa_{\text{ox}}}\frac{d\sigma}{dF}=(F-F_t)\left(\mu\frac{df}{dF}+f\frac{d\mu}{dF}\right)+f\mu, \tag{3}$$

where f is the fraction of the induced charge that is free, F is the field across the oxide, and F_t is the field at threshold. In cases of most interest, $f=1$ (no trapping and insignificant incremental charge in the depletion region) so that

$$\mu_{\text{FE}}=\mu+(F-F_t)d\mu/dF. \tag{4}$$

In general, since the mobility can vary strongly with field, the field effect mobility only approximates the conductivity mobility in field dependence. However, it is useful to measure because it is simpler to compare and because it brings out the sharp structure more easily than conductance.

EFFECTS OF FAST INTERFACE STATES

Hall data have been published elsewhere[18] for {111} surfaces that demonstrate that in properly annealed samples the incremental number of free electrons induced, dn/dF, is equal to the total incremental number of electrons induced (dN/dF); that is,

$$dn/dF=\kappa_{\text{ox}}/4\pi q=C_0\delta/q. \tag{5}$$

In most cases, this agreement is within 5% and dn/dF is constant over the measurement range, except at the lowest field at higher temperatures where the depletion region contributes to the charge.

The Hall effect can measure the surface-free carrier density n approximately. However, the density measured is equal to the true density divided by the Hall ratio r. The Hall ratio is of the Hall mobility μ_H to the conductivity mobility μ and is in the terminology of Brooks[21] $\langle\tau^2\rangle/\langle\tau\rangle^2$ where τ is the scattering time. Anisotropy factors may also play a role on {110} and {111} surfaces. The Hall ratio depends on the scattering mechanisms and has been calculated and measured for some bulk processes. The only estimate for surface scattering is for a classical case calculated by Zemel,[22] who finds that $r=0.87$. No estimates have been made for the quantized surface condition which applies over most of the temperature range studied. However, the Hall ratio seldom is more than 20% different from unity. In this work it has been assumed to be constant unless variations appeared in dn/dF which were correlated with strong mobility variations and which occurred in only a limited temperature range. These considerations result in a circular argument when the trapping effects are studied and some reservations should be made in accepting them. However, the error is probably less than 10%. In other words, it is assumed that when the observed dn/dF is constant and equal to $C_0\delta/q$ that r is constant and there is no trapping; when dn/dF is not constant in some other temperature range, the assumption is made, with reservations, that r is not constant. If dn/dF is not constant and is less than $C_0\delta/q$, the effect is assumed to be the result of trapping, again with reservations.

A deviation from the rule that dn/dF is nearly constant occurs near nitrogen temperature. There the mobility goes through two maxima as discussed below. As may be seen in Fig. 4, dn/dV increases in the range of gate voltages observed when two maxima occurred in the mobility, but not otherwise. The field effect mobility, effective mobility, and the Hall mobility are

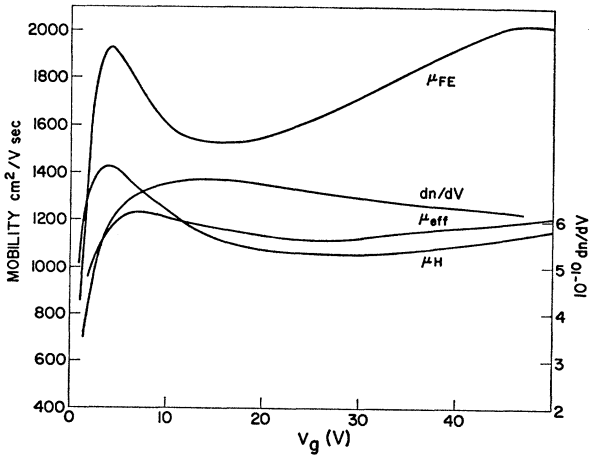

FIG. 4. Variation of Hall mobility, effective mobility, field effect mobility, and dn/dV in a 2-Ω cm sample at 77°K.

[21] H. Brooks, Advan. Electron. Electron Phys. **7** (1955).
[22] J. N. Zemel, Phys. Rev. **112**, 762 (1958).

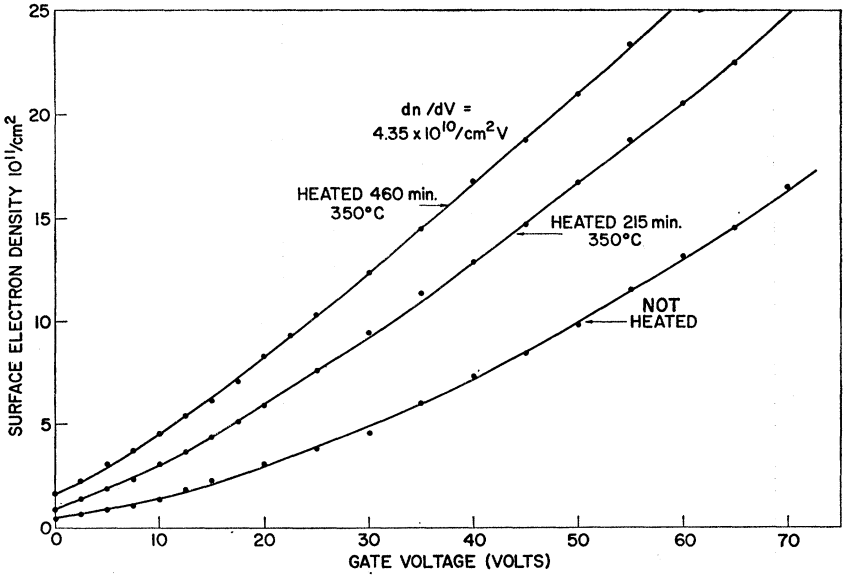

FIG. 5. Effect of annealing. This sample was made with an oxide 5400 Å thick and an aluminum gate. The carrier density is shown as a function of gate voltage for the sample as received and after 215- and 460-min anneal at 350°C in air.

also shown. This result could be interpreted as a variation in r as the scattering mechanism changes or possibly in terms of a two-band conduction model.

Many authors have reported evidence of fast states on oxidized silicon surfaces. Gray and Brown[23] have reported high densities of states near the valence band on p-type samples and near the conduction band on n-type. Workers, beginning with Terman,[24,25] have measured interface states near the center of the forbidden gap using capacitance measurements. Interface states near the conduction band edge in p-type silicon are not observable by Gray's technique and are difficult to observe with capacitance measurements. They can be observed with Hall measurements.

There have been many studies of annealing,[20] and the comments above, regarding the absence of interface states, apply to properly annealed surfaces. Annealing at 350°C for an hour, or 500°C for 5 min in air, affects only the interface state density under the aluminum. One can measure the interface state density and its change with annealing by making Hall measurements. The free carrier density n and therefore dn/dF can be measured as a function of the field F. The total charge N induced is known $[N=(\kappa_{ox}/4\pi q)F]$. The charge going into fast states dn_t/dF is given by $(\kappa_{ox}/4\pi q-dn/dF)$. Then, $dn_t/dn=(dn_t/dF)/(dn/dF)$ can be calculated as a function of n and, therefore, the surface potential ϕ_s, since ϕ_s is a known function of n. Finally, $dn_t/d\phi_s$ can be determined as a function of ϕ_s. The potential essentially measures the position of the Fermi level relative to the band edges at the surface so that the number of electrons entering interface state can be determined as a function of the position of the Fermi level.

Results of measurements of this type are shown in Fig. 5. The resulting calculated trap densities are shown in Fig. 6. A fairly flat distribution is observed before annealing between 0.175 and 0.025 eV below the conduction band edge ($kT\approx0.025$ eV) with a density of about 8×10^{13}/eV cm^2. After annealing, the density near the conduction band edge was reduced by nearly an order of magnitude for energies 0.05 eV below the conduction band edge. In other samples, the trap density was below measurable values. Suitable samples for this sort of measurement were rare because samples were seldom homogeneous before annealing. No clear pattern could be seen in the variation of mobility with the trap density near the band edge.

The apparent discrepancy between these measurements and Gray's measurements may have arisen because the latter were made on n-type silicon to determine state densities near the conduction band edge and were made on unannealed wet oxides.

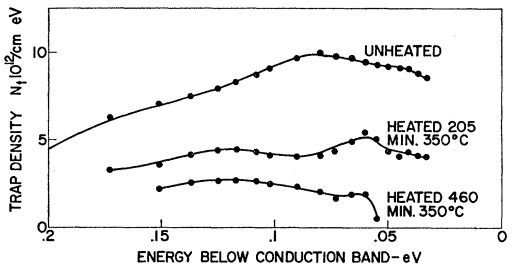

FIG. 6. Density of interface states as derived from the measurements in Fig. 5. The energy is measured from the conduction band edge.

[23] P. V. Gray and D. M. Brown, Appl. Phys. Letters 8, 31 (1966).
[24] L. M. Terman, Solid-State Electron. 5, 285 (1962).
[25] E. H. Nicollian and A. Goetzberger, Bell System Tech. J. 46, 1055 (1966); F. P. Heiman and G. Warfield, IEEE Trans. Electron. Devices 12, 167 (1965).

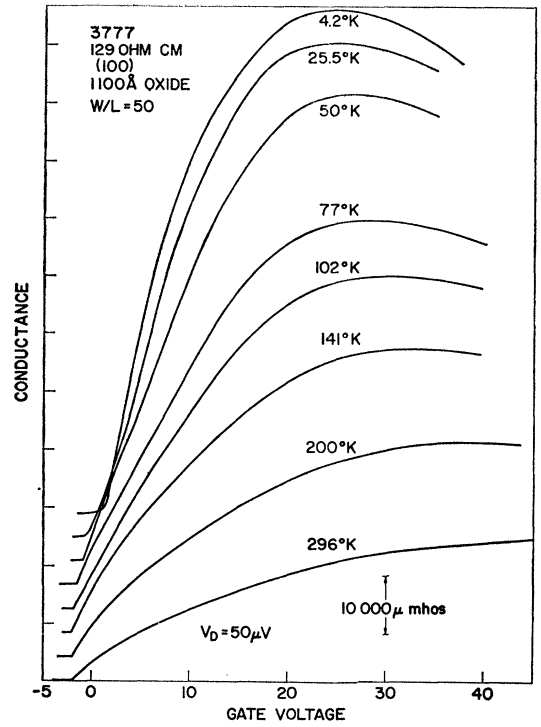

FIG. 7. Variation of conductance with temperature and electric field. In this sample, the field just inside the silicon surface is $8.7\times10^4(V-V_t)$. Other measurements were made at 6.9, 7.5, 9.1, 11.8, 13.6, 15.8, 17.3, 21.9, 30, 40, and 59°K.

The conclusion drawn from these measurements is that for many cases, after proper oxidation and annealing, the effective mobility is equal to the surface conductivity mobility. The only range where this comment does not apply is near threshold, where a significant part of the incremental charge is induced in the depletion region.

VARIATION OF MOBILITY

A typical result for the measurement of the surface conductance as a function of field is shown for a (100) surface in Fig. 7 for various temperatures. The {110} and {111} surface conductances varied in a similar way. The results of these measurements may be used to determine the dependence of the mobility on the normal electric field in the silicon, $F_\perp$, where $F_\perp$ is given by

$$F_\perp \approx \frac{\kappa_{ox}}{\kappa_{Si}}\frac{V_g'}{\delta}+\left(\frac{8\pi q N_a V_s}{\kappa_{Si}}\right)^{1/2}, \qquad (6)$$

where κ_{Si} is the silicon dielectric constant, V_g' is the gate voltage above threshold, N_a the acceptor density, and V_s is the electrostatic potential of the surface with respect to the bulk of the silicon. A general feature of the low-temperature variation from these data, as evidenced in Fig. 8, is an initial increase in the mobility, followed by a rapid decrease at higher fields which occurs at all temperatures. As pointed out, the uncertainty in assigning the threshold voltage limits the accuracy of the mobility determination in the low-field region.

The increase in mobility at low fields is shown by the Hall measurements of Figs. 9 and 10. It is present at all temperatures and had been reported earlier by Murphy.[26] Note that the samples measured at 300°K showed little difference in mobility at high carrier densities despite the variation in impurity concentration, but at low fields the impurity concentration was important.

An interesting correlation exists for the high-field scattering in {111} surfaces near room temperature. Because the energy difference between the ground state and the first excited state are larger than for the (100) surface and because the density of states is higher even at room temperature the majority of electrons are in the lowest electric sub-bands. For the data in Fig. 10, the mobility in cm²/V sec in the high-field range was found to be equal to 9.3 times the average distance in Å of the electrons from the surface z_{av}, where z_{av} was calculated by Howard and Stern using self-consistent solutions of Schrödinger's equation in the potential well. This result obtained for all of the samples which ranged in resistivity from 95 to 1 Ω cm. The temperature dependence was $T^{-1.5}$ near 22°C, which may indicate a phonon-scattering process.

The Hall mobility and carrier density are shown in Figs. 11 and 12 at several temperatures for typical {100} samples at various temperatures as a function of field. These curves established that the two maxima, discussed below, are real and not due to fast states.

Figures 13 and 14 show typical results for the temperature dependence of effective mobility on a (100) surface. At low fields, the effective mobility increased rapidly with temperature near 4.2°K. At higher fields, the temperature dependence decreased. This has been observed for all orientations as in the (110) sample in Fig. 15. This may suggest that the dominant scattering at low fields and carrier concentrations is due to im-

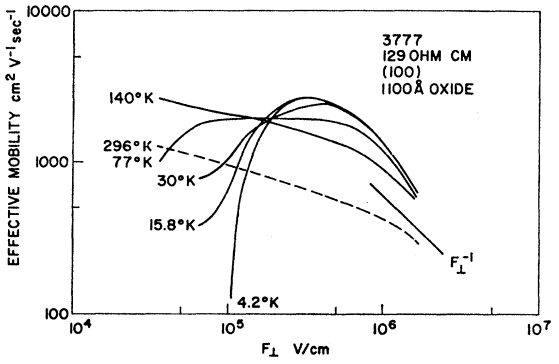

FIG. 8. Effective mobility as a function of electric field inside the silicon surface.

[26] N. Murphy, in *Solid Surfaces*, edited by H. Gatos (North-Holland Publishing Co., Amsterdam, 1964), p. 86.

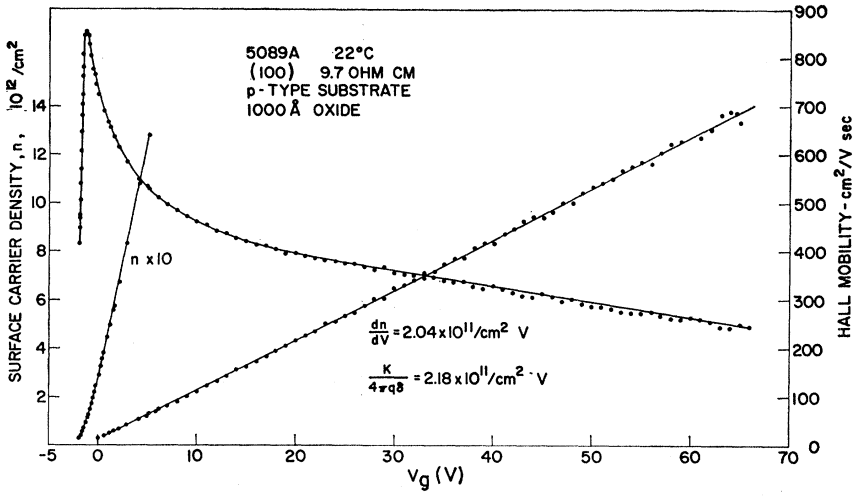

FIG. 9. The Hall mobility and carrier density for a ⟨100⟩ sample at 22°C.

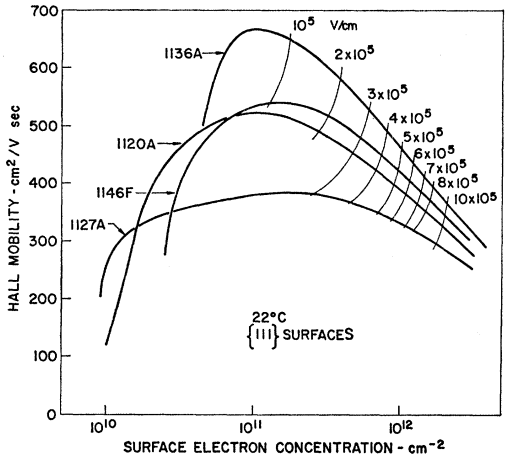

FIG. 10. The Hall mobility as a function of carrier density for a series of ⟨111⟩ samples of differing substrate resistivity. Curves are drawn through the points of equal field inside the surface as indicated in the upper right-hand corner. The resistivities of the samples were 95, 18, 6, and 1 ohm-cm, respectively, for 1136 A, 1146 F, 1120 A, and 1127 A.

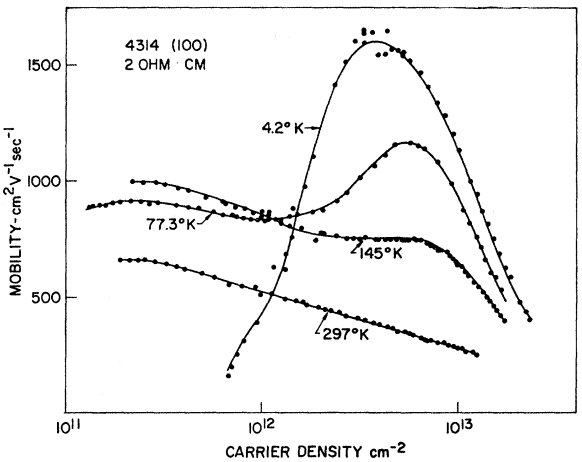

FIG. 12. Hall mobility for a 2 Ω cm sample. Both maxima are shown at 77.3°K.

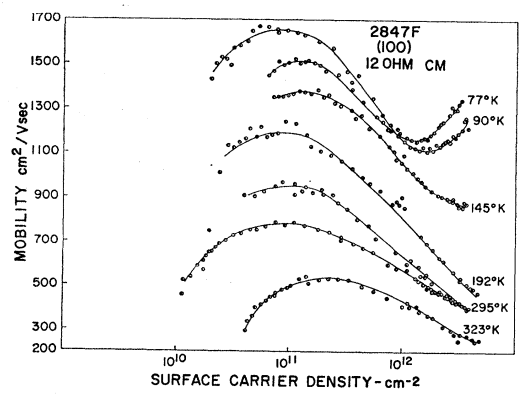

FIG. 11. Hall mobility as a function of carrier density for a 12 Ω cm (100) sample. A second maximum not shown in the data occurs somewhat below 10^{13}/cm² at 77.3°K.

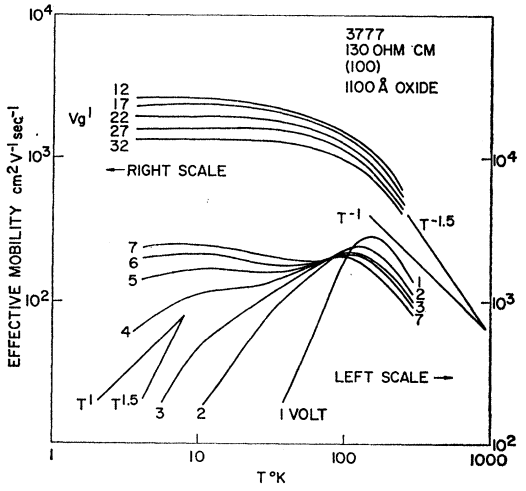

FIG. 13. Temperature dependence of effective mobility. The curves are plotted as a function of voltage above threshold, V_g'. The upper and lower sets of curves are displaced in scale.

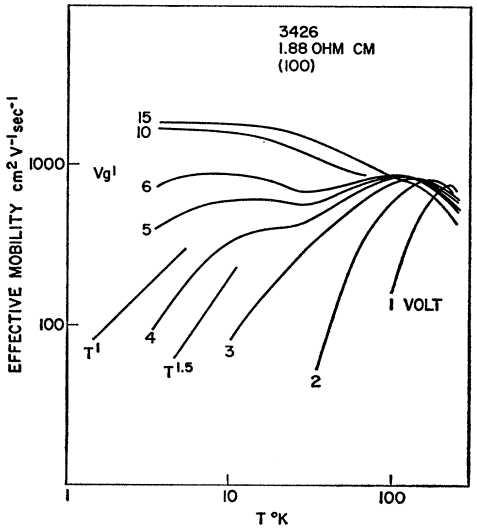

FIG. 14. Temperature dependence of effective mobility for a lower resistivity substrate.

purities. The impurity scattering cross section should decrease with increased field, carrier density, and screening, and with increasing temperature. At the highest fields, the temperature dependence should be that of a degenerate electron gas. Although some dependence of the mobility on substrate doping at low fields was found in the range of 7×10^{14} to 8×10^{15} cm^{-3}, it is much less than linear so that the major impurity scattering centers may be charged interface states or ions in the oxide near the interface. The mobility does not have a specific power-law dependence over the entire region and, in fact, increases as $\exp(-E/kT)$ in the lowest field region discussed in the low-temperature section.

Under many conditions, the finer features of the surface conductivity could be examined by the field effect mobility measurement. Figures 16 and 17 show a typical

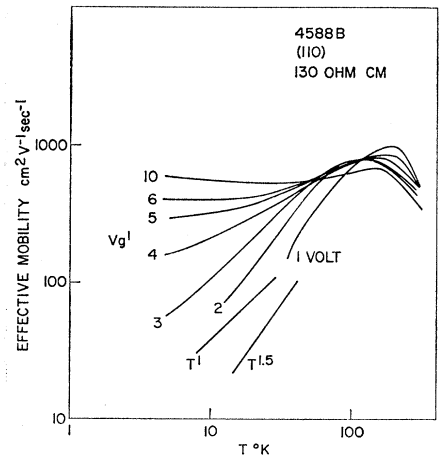

FIG. 15. Temperature dependence of effective mobility for a (110) sample. The oxide thickness is 1100 Å.

set of field effect mobility measurements on a (100) surface as a function of the field plate voltage at different temperatures. The structure near threshold is clearly shown. Between about 7.5 and 90°K, there are two main peaks in the field effect mobility which increase rapidly near threshold to the first sharp peak, decrease, and then go through a second broader peak. The first peak is very sensitive to temperature and substrate doping and is sometimes split. It always increases with temperature up to a temperature which is higher for purer substrates. It decreases above that temperature. The first peak is always more prominent for higher resistivity substrates, appearing only as a shoulder for {100} surfaces with the substrate resistivity less than 1 Ω cm. As the temperature is lowered toward 4.2°K, the peak disappears. The second peak is not observed

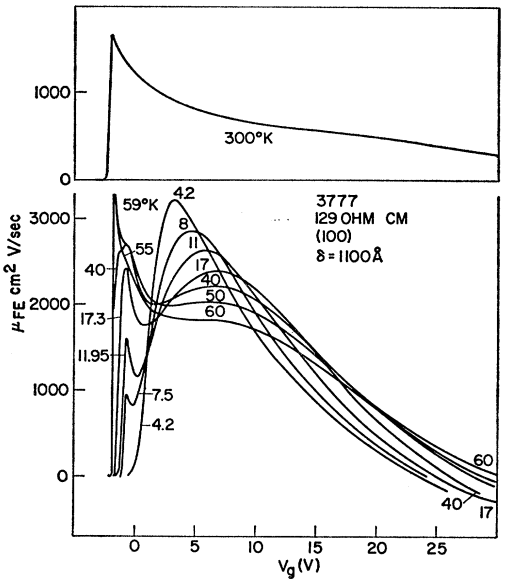

FIG. 16. Field effect mobility as a function of gate voltage for different temperatures for a high resistivity (100) sample.

until the temperature is lowered to about 90°K. The height of this peak increased with decreasing temperature below 90°K. These general features are present on the {100}, {110}, and {111} surfaces examined as well as samples of p-channel devices. However, the relative strength of the peaks and the separation depends on the orientation. It was difficult to see such structure on samples made on {111} surfaces of low-resistivity ($\rho<10$ Ω cm) substrates.

The field effect mobility decreases sharply[14] with field plate voltage after the second peak. For a (100) surface at a field of about 10^6 V/cm in the silicon (depending on the temperature), μ_{FE} reaches zero and becomes negative. That is, the conductance goes through a maximum, as seen in Fig. 7, or the mobility decreases faster than the carrier density increases. This effect has been seen on {110} surfaces at higher fields, and also on

{100} p-type surfaces. That this effect is not caused by trapping is also supported by Hall measurements.

In summary, these data demonstrate a single maximum mobility near room temperature, and an increase in this maximum as the temperature is lowered followed by a decrease to extremely low values at 4.2°K. A second maximum appears at about 90°K which increases as the temperature is lowered. A maximum in conductance is observed in most surfaces. The maxima occur at successively higher fields for {100}, {110}, and {111} surfaces.

EFFECT OF SUBSTRATE BIAS

Because the surface contacts are of opposite conductivity type to the substrate, it is possible to apply a reverse bias between the surface and the substrate. The surface field at a given gate voltage is independent of the substrate bias. However, a reverse substrate bias increases the depletion layer width so that the net effect is to reduce the number of surface electrons without changing the surface field. The number of electrons removed from the surface is approximately given by

$$\Delta n \approx (\kappa_{Si} N_a/2\pi q)^{1/2}[(V_s+V_{sub})^{1/2}-V_s^{1/2}], \quad (7)$$

where V_{sub} is the substrate bias with respect to the surface contacts and V_s is the electrostatic potential of the inversion layer in the absence of substrate bias. The normally observed threshold shift in field effect transistors is caused by this surface carrier depletion by the substrate bias.

Figure 18 shows typical results of measurements of μ_{FE} for a (100) n-channel sample at 77°K. The apparent shift in threshold is caused by just the mechanism described above and fits Eq. (7) well. The first peak decreased with reverse bias on the contact substrate diode as well as shifting. The second peak was much less affected. For forward bias, which increases the surface carrier concentration without increasing the surface field, the first peak in μ_{FE} was enhanced.

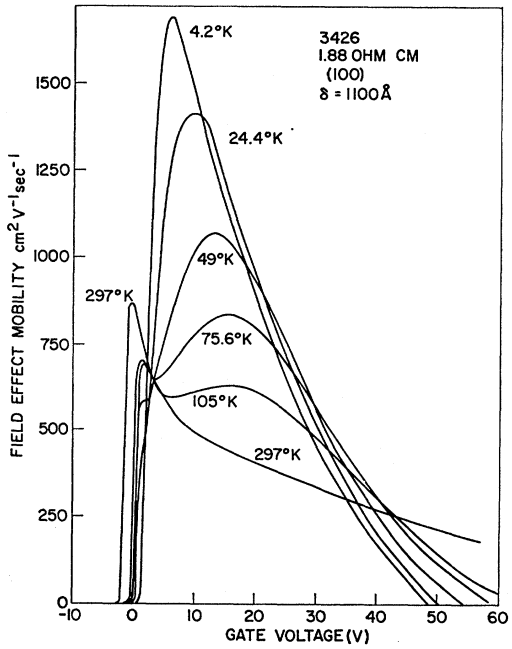

FIG. 17. Field effect mobility for a low resistivity (100) sample.

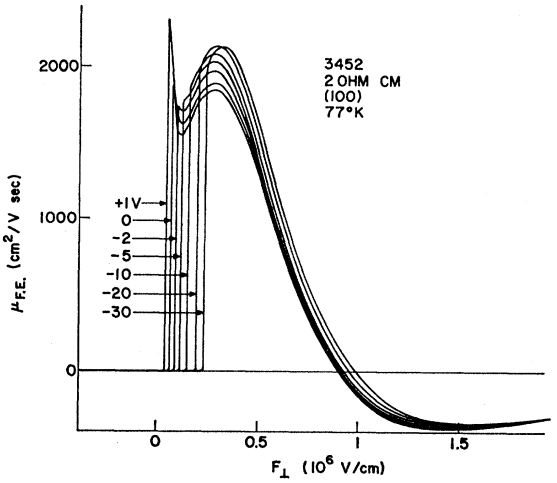

FIG. 18. Effect of substrate bias on the field effect mobility. These curves are for different voltages V_{sub} applied between the source and an Ohmic contact to the substrate.

These results seem to indicate that the mobility is primarily a function of the normal field rather than of the carrier density, at least for the region around the first peak in μ_{FE}. If only the carrier concentration were involved, a rigid shift of the curves would have been expected. In fact, the observations indicated that for equal free-carrier densities, the first peak decreased as the surface field increased. This result is consistent with the observation reported above, that the first peak is higher and more prominent for a lightly doped substrate. In that case, for $V_{sub}=0$, the normal field near threshold is proportional to $(N_a V_s)^{1/2}$. Thus, more lightly doped samples have lower surface fields at the onset of inversion and correspondingly higher initial peaks.

LOW-TEMPERATURE OBSERVATIONS

It is striking that as the temperature of a sample is lowered from 78 to 4.2°K, the current threshold voltage can change rapidly. Such a shift has also been observed by Nathanson.[27] The threshold for a field-effect device is given approximately by[28]

$$V_t=-\frac{Q_{ss}}{C_0}+\frac{Q_{sT}}{C_0}+\psi_{ms}+V_{sT}, \quad (8)$$

[27] H. I. Nathanson, C. Jund, and J. Grosvalet, in Proceedings of the International Electron Devices Meeting, Washington, D. C., 1967 (unpublished); and private communication.

[28] F. P. Heiman and H. S. Miller, IEEE Trans. Electron Devices **12**, 142 (1965); L. Vadaz and A. S. Grove, IEEE Trans. Electron Devices **13**, 863 (1966).

where V_t is the voltage at threshold between the gate and the source contact, Q_{ss} is the charge in the oxide assuming it is at the silicon-silicon oxide interface, C_0 is the oxide capacitance, $-\kappa_{ox}/4\pi\delta$, Q_{sT} is the charge in the depletion layer at threshold, $-(\kappa_{Si}N_aqV_{sT}/2\pi)^{1/2}$, V_{sT} is the voltage drop across the depletion layer at threshold, and ψ_{ms} is the contact potential difference between the source contact and the gate. The contact potential difference depends on the Fermi energy. For an aluminum electrode, it will be about[29] -1.1 V less the Fermi energy measured from the valence band. The primary contribution to its temperature dependence should be from the change in the Fermi level. However, the temperature dependence of the barrier height has not been measured.

Various authors have used different criteria for threshold. Grove and Fitzgerald[30] have used the criterion that $V_{sT}=2\psi_F$, where ψ_F is the Fermi energy in the interior measured from the center of the energy gap. This criterion is a requirement that the surface be as n-type as the bulk is p-type, and cannot be a useful definition for intrinsic substrates at low temperatures. We assume that at low temperatures the bands must be bent so that the Fermi level is at the conduction band edge at the surface. That is, $qV_{sT}=\epsilon_c-\epsilon_F$, where $\epsilon_c-\epsilon_F$ is the energy difference of the conduction band and Fermi level in the interior.

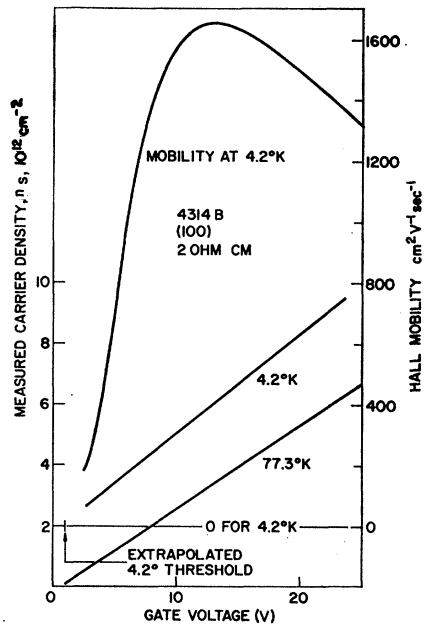

FIG. 19. Comparison of Hall data at 4.2 and 77.3°K. The data have been displaced. The extrapolated threshold for free-carrier induction is 0.7±0.2 V for both 77.3 and 4.2°K.

From this argument, the temperature variation of V_{sT} arises primarily from the change in Fermi energy in the temperature range of interest. It is in fact equal and opposite in sign to the change in ψ_{ms} so that the only variation in threshold voltage comes in through the Q_{sT} term.

Therefore,

$$\frac{dV_T}{dT}\approx\frac{dV_{sT}}{dT}\frac{dQ_{sT}}{dV_{sT}}. \tag{9}$$

Using Eq. (8), and assuming

$$V_{sT}\approx E_g+(kT/q)\ln(N_a/A_p), \tag{10}$$

where E_g is the energy gap (about 1.21 eV below 77°K), and A_p is the effective valence band density of states,[21] which is $0.92[2\pi m/h^2]^{3/2}[kT]^{3/2}$. Then,

$$\frac{dV_{sT}}{dT}\approx\left[\frac{8\pi\kappa_{Si}N_a}{2\pi}\right]^{1/2}\frac{k\delta}{{}_0\kappa_x}\ln\left(\frac{N_a}{A_p}\right).$$

This result is only approximate but gives an order-of-magnitude result between 77 and 4.2°K. For 30°K, $\delta=10^{-5}$ cm, and $N_a=10^{15}$/cm^2, then $dV_t/dT\approx 7$ mV/°K. Changes of as much as 10 V have been seen (see Table I) in this range. Thus the shift in conductance threshold is much larger than expected from the argument above.

Two measurements indicate that the threshold for onset of inversion does not change appreciably in this temperature range. The results of Hall measurements at 4.2°K, as shown in Fig. 19, demonstrate that the apparent shift is caused primarily by what seems to be a sharp decrease in mobility. Extrapolation of the data for carrier density as a function of gate voltage indicates the same threshold for both temperatures. Similarly, results for oscillatory magnetoconductance measurements[12,13] indicate that the Landau levels extrapolate to the 77°K threshold at zero magnetic field. (The Landau levels should coalesce to the conduction band edge.) The Hall measurements also indicate that this effect is not caused by contact resistance at low temperatures since they are essentially open-circuit measurements.

Table I shows representative results for the apparent low-temperature shift. The shift is not always well defined at 4.2°K even by measurements of the field effect mobility so that the error is about 10% in ΔN_s, which is taken as equal to $(\kappa_{ox}/4\pi\delta q)\Delta V_t$, is determined from Eq. (8) using $V_{sT}=1.2$ V.

One strong correlation can be made from this table. The size of the shift depends on the size of the built-in surface charge. This charge is usually ascribed, in part, to states close to the silicon-silicon oxide interface in the oxide.[31] It should be noted that the correlation obtains whether the charge is attractive or whether it is repulsive, as in the case of the gold-doped oxides, although

[29] A. M. Goodman, Phys. Rev. **152**, 785 (1966); A. M. Goodman and J. J. O'Neill, J. Appl. Phys. **37**, 3580 (1966); B. E. Deal, E. H. Snow, and C. A. Mead, J. Phys. Chem. Solids **27**, 999 (1966).

[30] A. S. Grove and D. Fitzgerald, Solid State Electron. **9**, 783 (1966).

[31] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, J. Electrochem. Soc. **114**, 266 (1967).

TABLE I. Low-temperature shift in conductance threshold.

Sample	Surface orientation	Resistivity (Ω cm)	Oxide thickness (Å)	Threshold voltage 77°K	Threshold voltage 4.2°K	N_{ss}[a] (10^{11}/cm²)	ΔN_{ss}[b] (10^{11}/cm²)	Max. field effect mobility at 4.2°K (cm²/V sec)
87-11 H_1	(100)	10	1100	0	0.9	1.45	1.81	3100
11 H_2	(100)	10		−0.5	0.8	2.50	2.6	3400
11 H_3	(100)	10	1100	−0.1	1.0	1.65	2.2	3320
87- 8 H_1	(100)	10	990	−1.4	0.9	4.57	5.12	3100
8 H_2	(100)	10	990	−1.8	0.7	5.46	5.57	
87-10 H_1	(100)	10	1100	−0.3	1.2	2.05	3.0	2700
10 H_2	(100)	10	1100	−0.4	0.7	2.25	2.2	
87- 7 H_1	(100)	10	990	0.3	0.4	0.85	0.20	2550
H_2	(100)	10	990	0	0.4	1.45	0.60	
87- 5 H_1	(100)	2	1100	0.7	1.3	1.98	1.21	1100
H_2	(100)	2	1100	0.6	1.0	2.18	0.81	1200
87- 4 H_1	(100)	2	1100	0.7	1.2	1.98	1.11	2910
H_2	(100)	2	1100	0	1.2	3.39	2.68	2540
87- 2 H_1	(100)	2	990	0.7	1.2	1.83	1.11	2910
H_2	(100)	2	990	0	1.2	3.39	2.68	2540
87- 1 H_1	(100)	2	990	−0.4	0.9	4.28	2.90	1600
H_2	(100)	2	990	0.8	1.0	1.61	0.45	2530
8279 H_1	(100)	10	1000	0.15	0.9	1.08	1.65	
H_2	(100)	10	1000	0	0.7	1.45	1.55	
8286 H_1	(100)	10	1000	0.6	1.5	0.12	1.99	
H_2	(100)	10	1000	0.45	1.5	0.99	2.32	
8282 H_1	(100)	10	1000	0.45	1.3	0.99	1.88	
H_2	(100)	10	1000	0.45	1.0	0.99	1.44	
8278 H_1	(100)	10	1000	0.7	1.2	−0.1	1.11	
H_2	(100)	10	1000	0.8	1.3	−0.32	1.11	
8292 H_1	(100)	10	1000	1.6	2.7	−2.09	2.43	
H_2	(100)	10	1000	1.4	2.5	−1.45	2.43	
3046	(100)	8.6	1100	−3.5	1.7	7.71	10.4	3300
3777	(100)	129	1100	−2.0	−0.5	4.47	5.0	5850
3426	(100)	1.88	1100	−0.5	1.0	4.44	3.01	3060
3431 A	(100)	2	1100	−0.2	2.0	3.79	4.22	
B	(100)	2	1100	−0.9	1.8	5.20	5.44	
8276 A	(100)	10	1100	−0.3	0.85	2.05	2.31	3900
B	(100)	10	1100	0.45	1.05	0.55	1.21	
5057 A	(100)	2	1100	−0.1	1.05	3.19	2.14	
B	(100)	2	1100	−0.25	1.0	2.89	2.52	
8280 A	(100)	10	975	0.40	0.90	0.55	1.13	
B	(100)	10	975	0.30	0.95	0.77	1.47	
1421	(100)	2	1200	0.6	1.0	2.29	0.74	
3715	(100)	2	1100	−6.8	−4.3	13.8	5.03	
3301	(111)	100	1100	−6.25	1.0	13.1	14.5	317
8757 H_1	(111)	10	1000	−0.45	0.6	2.44	2.34	1200
H_2	(111)	10	1000	−0.35	0.7	2.23	2.34	1225
9774 H_1[c]	(111)	10	1000	6.25	9.5	15.2	6.5	130[c]
H_2	(111)	10	1000	6.75	10.25	16.3	7.71	245
H_3	(111)	10	1000	6.75	9.25	16.3	5.52	240
4588	(110)	130	1100	−1.0	2.0	2.46	6.0	

[a] Calculated from Eq. (9) using $\phi_{ms}+V_{st}=0$, and $V_{st}=1.2$ V. The accuracy of these figures is not better than $\pm 0.2\times 10^{11}$.
[b] Calculated using $\Delta N_{ss}=C_0\Delta V_t/q$. This is the apparent shift in oxide charge and is always negative.
[c] These samples had gold diffused into the oxide from the back of the sample.

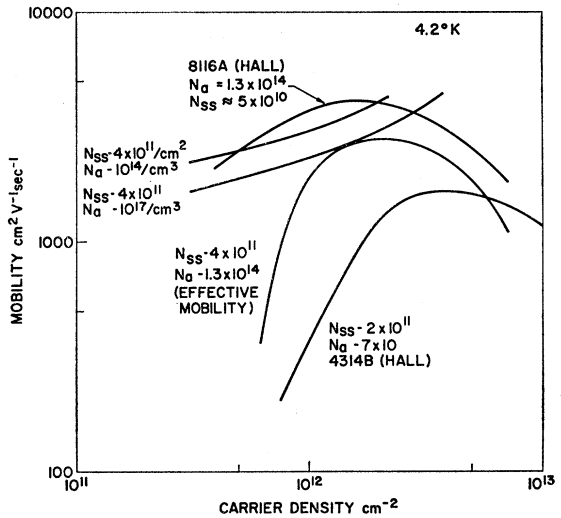

FIG. 20. Comparison of the measured mobilities and the surface impurity scattering theory of Stern and Howard. The theoretical curve corresponding to 8116 A would be off the scale.

ΔN_s is smaller than one might have expected for the gold-doped samples.

This result may be taken as strong evidence that at low temperatures the primary scattering mechanism at electron densities of the order of the oxide charge is caused by Coulombic scattering by the charges near the interface. Such scattering has been discussed by Kawaji and Kawaguchi[32] and in more detail by Stern and Howard.[15] The latter have fit some of these data at somewhat higher fields than in this region using a two-dimensional gas model. However, they[33] do not feel that their theory is applicable at the lowest fields. A comparison is shown in Fig. 20.

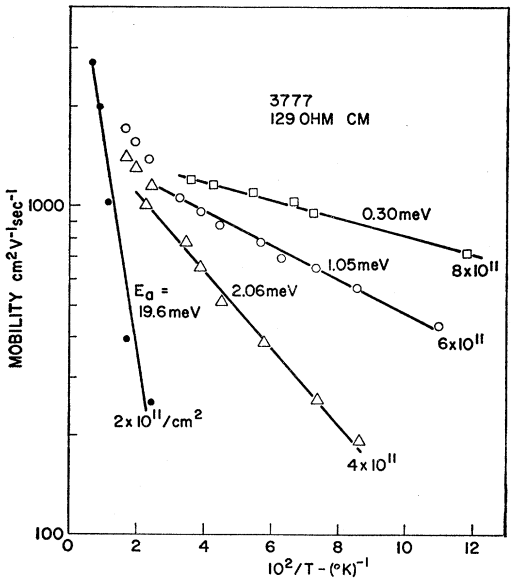

FIG. 21. Mobility as a function of temperature. The carrier densities and apparent activation energies are given for increasing gate voltage.

In their theory, Stern and Howard have suggested an explanation for the threshold shift when the potential is attractive. They have shown that positive ions near the interface result in bound states. They predict that the binding energy should decrease as the free-electron density increases because of increased screening. At some electron density, the binding energy would be small enough so that the electrons would be released. Something of this sort would qualitatively explain the conductance data.

This explanation is supported by the temperature dependence of the effective mobility in the range of the threshold shift. Figure 21 shows the dependence of the effective mobility for different fields. The effective mobility increases as $\exp(-\epsilon/kT)$. The activation energy decreases with field which would seem to be consistent with the bound-state theory of Howard and Stern.

The Hall data seem to show that the Hall mobility decreases rapidly near the threshold at 4.2°K. If the bound states overlapped slightly, a two-band model of the mobility would be necessary for interpretation. Then the measured mobility and carrier density would be approximately

$$\mu_m \cong \frac{n_1\mu_1^2+n_2\mu_2^2}{n_1\mu_1+n_2\mu_2} \cong \mu_1\left(\frac{f+(1-f)b^2}{f+(1-f)b}\right)$$

and

$$n_m \cong \frac{(n_1\mu_1+n_2\mu_2)^2}{n_1\mu_1^2+n_2\mu_2^2} \cong \frac{[f+(1-f)b]^2}{f+(1-f)b^2},$$

where b is the ratio of the mobility in the bound-state band to the free-electron mobility and f is the fraction of electrons that are free. As is well known,[34] if b is constant, the measured carrier density has a minimum when f is varied and the experimental evidence for an extremum in this case is small. There is a rise in dn/dV_g near threshold and a subsequent decrease, but it is much less sharp than would be needed to explain the conductance variation.

The results at present are not conclusive evidence for the bound-state model or for a sudden drop in mobility. Perhaps the impurity scattering is very strong at low charge densities and the bound states also play a role.

For the case of gold-doped samples, the potential is repulsive and seems much less effective in causing a shift. ($\Delta N_s < N_s$) bound electrons are not possible here but an analogous bound hole may play a role. Another way of looking at this case is that the negative ions will cause local cavities in the conducting film—a sort of

[32] S. Kawaji and Y. Kawaguchi, J. Phys. Soc. Japan Suppl. **21**, 336 (1966).

[33] F. Stern and W. E. Howard (private communication).

[34] H. Fritzsche, Phys. Rev. **99**, 106 (1955).

"Swiss cheese" conducting sheet. The screening will affect the size of the cavities and they might also disappear abruptly as free electrons are induced.

A correlation between the maximum observed field effect mobility and other properties is somewhat uncertain. It can be said that the mobility for {100} surfaces—no matter what the substrate resistivity—is greater than 1100 cm^2/V sec, whereas for {111} surfaces, the mobility was smaller than 1200 cm^2/V sec at 4.2°K. There seems to be some but not complete correlation with substrate resistivity and some with oxide charge, but the effect of these is not conclusive yet. It was to be expected that the electrons in a (111) surface with a mass of 0.38 m_0 parallel to the surface and a two-dimensional density of states[31] of $24\pi(0.38m_0)/h^2$ should be more effectively scattered by ions than those in a (100) surface with a mass of 0.2 m_0 and a density of states of $(8\pi/h^2)(0.2\ m_0)$. The Fermi velocities are inversely proportional to the square root of the density of states for equal numbers of electrons and to the square root of the effective mass parallel to the surface. Thus, the velocity for equal numbers of electrons is about 3.3 higher in a (100) surface than in a (111) surface. Hence, it is not surprising that the (100) surface demonstrates higher mobility than the (111) at low temperatures where impurity scattering dominates.

One might expect that the effect of scattering by surface phonons or by roughness of the surface should be less for a (111) surface than for a (100) surface because the lighter mass perpendicular to the surface results in the wave functions more distant from the surface on the average. No theory based on a quantized model for surface scattering exists but it does seem reasonable that when the wave function's average depth is of the order of the depth of surface irregularities, scattering should result. That should occur sooner for {100} than for {111} surfaces.

CONCLUSIONS

In this paper, we have attempted to survey some of the salient aspects of transport in the surface of silicon. While the results are incomplete in some areas, it is believed that the general picture is clear. However, many outstanding problems remain—both experimental and theoretical.

Outstanding among the experimental problems is a study of the detailed relationship between mobility and oxide charge and substrate doping for the different surfaces. The results in this paper allow general conclusions but are not complete if a unified theory of scattering is to be tested. They should serve to help decide what scattering mechanisms are important. Another question of interest is the expected anisotropy on the (110) surface. Such an effect has been reported for holes.[16]

The only scattering theory based on a quantum model of the surface that is relatively complete is for ionized impurity scattering on {100} surfaces at low temperature, and this seems to have limited range of validity. Interband scattering has also been considered at higher electron densities. However, no scattering theory exists for temperatures above 77°K.

An interesting observation is that the mobility at high fields at room temperature is a linear function of the average distance of the electrons from the surface independent of doping. This may provide a starting point for the interpretation of the high-field scattering.

One of the outstanding questions concerns the cause of the two maxima in the liquid-nitrogen temperature range. While it is a common result for mobility to have a maximum as a function of some variable, a minimum is not common.

There are other results which will not be discussed here, but will be addressed in later papers. Among these are fine structures in the first peak at 77°K and their temperature dependences and hot-electron effects at 4.2°K and below.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the help of many associates in this work. We are especially grateful to W. P. Jones and P. J. Stoller, our research assistants, for their skilled and efficient help. The samples used required participation of many members of our staff, of whom we would single out F. Hochberg, E. C. Wurst, and A. C. Brugess for particular thanks. We would also like to thank W. E. Howard, F. Stern, and P. J. Stiles for many profitable discussions. We would also like to acknowledge the careful review of the manuscript by W. E. Howard, F. Stern, and G. Cheroff.