Tunneling from Metal to Semiconductors*

PETER V. GRAY†
Electrical Engineering Research Laboratory, University of Illinois, Urbana, Illinois
(Received 23 June 1964; revised manuscript received 6 May 1965)

Conductance and differential capacitance have been measured as a function of frequency and of bias on a number of silicon–silicon-dioxide–aluminum sandwiches. The results indicate that information about the density and energy of states near the silicon–silicon-dioxide interface may be deduced from these measurements. Conductance is due to electron tunneling transitions between the aluminum and silicon. Analysis of the tunnel current is given in terms of a single-particle WKB approximation. A simplified formula is derived to explain tunneling between the metal and the intrinsic semiconductor bands. A preliminary analysis of tunneling to impurity bands is discussed. Experimental data are shown which display details of impurity-band spreading near the valence-band edge of silicon samples containing 10^{18} and 10^{19} boron atoms/cm^3. Tunneling from silicon–silicon-dioxide interface states is analyzed by defining a tunneling lifetime and using this concept to explain the observed frequency dependence of the conductance and capacitance measurements. Quantitative experimental values are obtained giving the density of interface states versus energy. These results are in close agreement with those obtained by Statz *et al.*, from inversion-layer conductance studies.

I. INTRODUCTION

In recent years tunneling techniques have come into wide use in the study of the properties of materials. Detailed analysis of the tunnel diode[1] has yielded a great deal of new and corroborative information about the structure and electrical properties of semiconductors. The use of a thin oxide film as a tunneling barrier between a metal and superconductor stimulated further interest in tunnel studies.[2] The present effort was undertaken to explore the usefulness of tunnel-current measurements between a metal and semiconductor through such an oxide film.

While all of the experimental data to be presented were obtained using the silicon–silicon-dioxide–aluminum system, the theoretical model is more general and could be applied to any metal-insulator-semiconductor system in which the principle transport limitation is the ability of the electrons to tunnel through the insulator.

Figure 1 is an electronic energy-level diagram of a sample showing the conduction and valence bands of the semiconductor and localized states at the interface between the oxide and semiconductor. A narrow band of acceptor-impurity states is shown near the semiconductor valence-band edge. The object of the experiment is to observe and interpret electron-tunnel transitions between these different semiconductor states and the metal. In particular, it is desired to study the details of impurity-band spreading at high-impurity concentrations and to measure the location and density (in energy) of the interface states. The interface states discussed here are often referred to as "fast states" in the literature.

As shown in Fig. 1, the bands of the p-type silicon used in the present measurements are bent downward at the surface forming a depletion layer.[3] The details of this nearly intrinsic region are critical to the success of the experiment. The valence band must be nearly depleted of holes near the surface to avoid a spurious current of holes tunneling to the metal beneath the Fermi level. Also, such holes could combine with electrons in the interface states; and, as will be seen, it is desired to have the interface states remain in equilibrium with the metal rather than with the semiconductor. On the other hand, the resistance of the space-charge region must be small compared with that of the oxide layer in order that nearly all of the applied bias

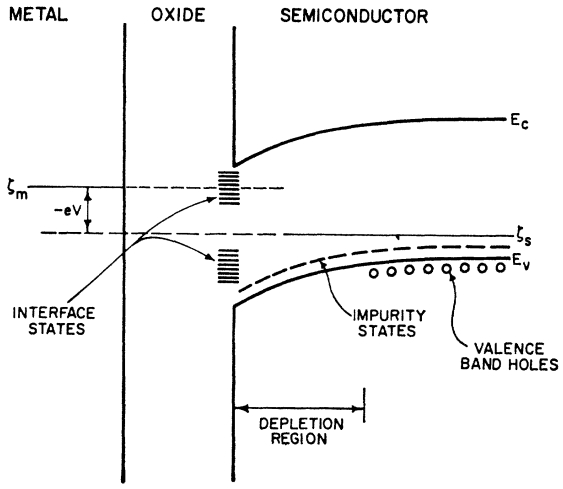

FIG. 1. Energy-level diagram of a metal-oxide-semiconductor sandwich. The locations of the impurity and interface states are shown. The conduction and valence-band edges in the semiconductor are E_c and E_v. The metal and semiconductor quasi-Fermi levels are ζ_m and ζ_s.

* Based on a thesis submitted in partial fulfillment of the requirements of the Ph.D. degree from the Graduate School of the University of Illinois, Urbana, Illinois. Sponsored by the U. S. Air Force Office of Scientific Research. A preliminary account of this work was given in Phys. Rev. Letters **9**, 302 (1962).

† Present address: General Electric Research Laboratory, Schenectady, New York.

[1] L. Esaki, Phys. Rev. **109**, 603 (1958).

[2] I. Giaever, Phys. Rev. Letters **5**, 147 (1960).

[3] For a review of the properties of semiconductor surfaces see *Semiconductor Surface Physics*, edited by R. H. Kingston (University of Pennsylvania Press, Philadelphia, Pennsylvania, 1957).

appear across the oxide rather than as "band bending" in the semiconductor. This requirement is easily met for biases at which the metal Fermi level is opposite the semiconductor forbidden band. However, when the metal Fermi level is opposite either of the intrinsic semiconductor bands, the impedance of the oxide is effectively reduced by the larger tunnel current and "band bending" may occur. As a result, the range of permissible oxide thickness is quite limited and depends upon the resistivity of the silicon. The oxide must be thick enough to be the only significant barrier to current flow, yet thin enough to allow tunneling to dominate such other processes as Ohmic conduction and Schottky emission. It is possible to compensate somewhat for oxide thickness by varying the temperature, since this changes the resistance of the depletion layer without seriously affecting the tunneling probability through the oxide.

With the semiconductor surface depleted, the lifetime of an electron before tunneling from an interface state will commonly be much shorter than its recombination lifetime. As a result, the interface states will remain filled to the energy of the Fermi level in the metal provided that this lies between the semiconductor band edges, and that variations in bias are quasistatic. Hence, little dc current flows via interface states. If a small low-frequency ac voltage is superimposed on the bias voltage, a definite charge ΔQ will tunnel back and forth between the metal and the interface states with each cycle of ac voltage of amplitude ΔV. The quantity of charge in the interface states in this voltage range will be proportional to the ac voltage, thus giving a capacitive current. The number of interface states in the range ΔV is $\Delta Q/e$. The energy density of interface states at this particular bias then is

$$\rho(E)=C(E)/e^2A\,, \tag{1}$$

where A is the sample area, E is the energy to which the interface states are filled, and C is the observed differential capacitance $\Delta Q/\Delta V$. At high frequencies very little charge will be transferred during each cycle. The *current* will then be nearly proportional to the ac voltage, giving a frequency-independent conductance which is proportional to the density of interface states. The small residual capacitive component of current corresponds to the capacitance of the oxide layer and space-charge region in series.[4] Measurement of this capacitance is important, however, since it indicates the width of the space-charge region and its variation with bias. The meaning of high and low frequency in this context is with reference to the tunneling probability per unit time of an electron in an interface state.

As indicated, the dc current is essentially unaffected by the presence of interface states. Since only electrons having energies between the metal and semiconductor quasi-Fermi levels can tunnel, little dc current is observed until the Fermi level in the metal reaches the energy of a semiconductor band edge or an impurity band. Thus, a plot of dc conductance (dI/dV) versus bias will show structure at biases corresponding to impurity bands and intrinsic band edges. The results are not affected by the band structure of the metal. These measurements are especially useful for studying the details of impurity band spreading in heavily doped semiconductors.

When samples are prepared which meet the requirements of oxide thickness and semiconductor surface potential discussed above, the possibility that conduction is due to transport processes other than tunneling is easily ruled out. The oxide films are too thin for space-charge limited current. The magnitude and bias dependence of the conductance remain virtually unchanged over a wide temperature range eliminating the temperature sensitive Schottky emission, as well as Ohmic conduction. In addition, it will be seen that the experimental results are readily interpreted in terms of established tunneling theory and the known properties of silicon.

The next section is devoted to the theoretical considerations necessary for interpretation of the measurements. In Sec. III, a convenient method of sample preparation is described in detail along with the instrumentation and equipment used for fabrication and measurement of samples. In Sec. IV, representative experimental data are presented to illustrate the sort of information which can be obtained from silicon–silicon-dioxide–aluminum samples by measurements of conductance and differential capacitance versus bias at different frequencies and temperatures.

Finally, it should be emphasized that the present study is preliminary in nature and exploratory in purpose. An established technique has been put to use in a new configuration which promises to be useful for probing the details of semiconductor surface structure.

II. THEORY OF TUNNELING THROUGH AN INSULATING FILM

It has been shown[5] that the probability per unit time that an electron will make a tunnel transition from state a on one side of a barrier to state b on the other side may be written as

$$w_{ab}=(2\pi/\hbar)\,|M_{ab}|^2\rho_b f_a(1-f_b)\,, \tag{2}$$

where M_{ab} is the matrix element for the transition, ρ_b is the density of final states, and f_a and f_b are initial- and final-state occupation probabilities. This is a direct application of the "golden rule" of time-dependent perturbation theory. From this it follows that the net tunnel current is given by

$$j_x=\frac{4\pi e}{\hbar}\sum_{k_t}\int_{-\infty}^{\infty}|M_{ab}|^2\rho_a\rho_b(f_a-f_b)dE\,, \tag{3}$$

[4] J. Bardeen, Bell System Tech. J. **28**, 428 (1949).

[5] J. Bardeen, Phys. Rev. Letters **6**, 57 (1961).

where the sum is taken over all transverse wave numbers k_t. It will be assumed throughout that k_t is conserved in each transition, although it is clear that scattering by phonons and defects will permit violations of this restriction. The sum over k_t may be converted to an integral

$$\sum_{k_t} \rightarrow \frac{1}{(2\pi)^2}\int dk_y dk_z \rightarrow \frac{1}{2\pi} \times \int k_t dk_t \rightarrow \frac{1}{2\pi}\frac{m_t}{\hbar^2}\int dE_t, \quad (4)$$

where $E_t=(\hbar^2 k_t^2/2m_t)$ and m_t is the cyclotron effective mass for a magnetic field in the x direction. Thus

$$j_x=\frac{2m_t e}{\hbar^3}\int_{-\infty}^{\infty} dE \int dE_t |M_{ab}|^2 \rho_a \rho_b (f_a - f_b) \quad (5)$$

and the limits on E_t must depend on the configuration.

The first case to be considered is that of transitions between a metal and an intrinsic band in a semiconductor. The two materials are separated by a thin insulating film. Because of the lack of detailed knowledge of the structure near the two interfaces to be crossed and of the insulator itself, we shall make a WKB approximation and attempt to describe a plausible potential barrier. Figure 1 shows the sample energy configuration.

We take the electron wave function in the metal to be

$$\psi_a = c_a k_x^{-1/2} \cos\left(\int_x^{x_a} k_x dx + \delta\right) \quad (6)$$

and in the oxide film

$$\psi_a = \tfrac{1}{2} c_a k_x^{-1/2} \exp\left(-\int_{x_a}^{x} |k_x| dx\right). \quad (7)$$

The wave function is normalized in a region of width L_a to the left of the metal-oxide interface x_a, over which the structure of the metal does not vary. Thus,

$$c_a^2 = 2(k_x)_a / L_a. \quad (8)$$

Following an analogous procedure with the semiconductor yields the matrix element for the transition

$$|M_{ab}|^2 = \left(\frac{\hbar^2}{2m}\right)\frac{2(k_x)_a}{L_a}\frac{(k_x)_b}{L_b} e^{-2K}, \quad (9)$$

where $K=\int |k_x| dx$ taken between classical turning points and the subscripts a and b refer to metal and semiconductor, respectively. The one-dimensional density of states factors of Eq. (5) may be written

$$\rho_i = m L_i / \pi \hbar^2 (k_x)_i. \quad (10)$$

Substituting Eqs. (9) and (10) into (5) gives for the current

$$j_x = \frac{m_t e}{2\pi\hbar^3}\int_{-\infty}^{\infty} dE \int dE_t e^{-2K}(f_a - f_b). \quad (11)$$

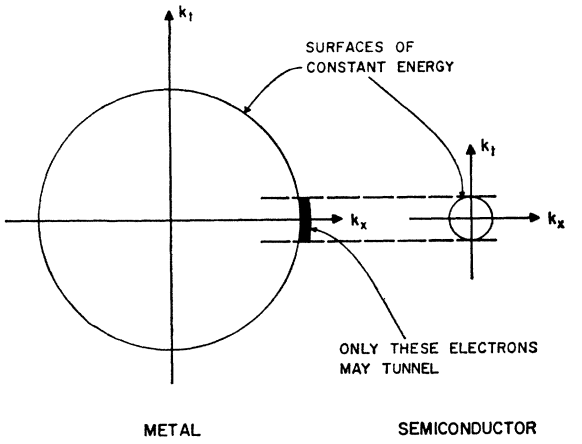

FIG. 2. Two-dimensional constant-energy loci in k space. The energy is close to a normal band edge in the semiconductor. Relatively few metal electrons are available for k_t-conserving tunnel transitions.

Of particular interest is the lack of explicit dependence upon density of states. It is for this reason that the results are independent of the band structure of the metal.

Even without precise knowledge of K, it is possible to make a useful analysis by considering only the dependence of the tunnel current on the semiconductor band structure. Suppose that when a voltage $-V_b$ is applied between metal and semiconductor, the Fermi level in the metal is at the same energy as the conduction-band edge in the semiconductor. We now examine conductance (dj_x/dV) as the voltage is varied a small amount. It is assumed that the semiconductor conduction band is parabolic with minimum at $k=0$. We neglect the dependence of K on V and E. Figure 2 shows the allowed transitions in k space. The limitations on E_t may readily be seen. For simplicity, the calculation is done at $T=0°$K. Using the notation of Fig. 1, Eq. (11) may now be written as

$$j_x = \frac{m_t e \exp(-2K)}{2\pi^2\hbar^3}\int_{-e(V-V_b)}^{0} dE \times \int_0^{e(V-V_b)} dE_t\, e^{-2K' E_t}, \quad (12)$$

where K has been expanded to first order in E_t and $K'=(\partial K/\partial E_t)$. Both K and K' are evaluated at $V=V_b$ and $E_t=0$. Thus, for $|V|>|V_b|$

$$j_x \cong \frac{m_t e^2 \exp(-2K)}{4\pi^2\hbar^3 K'}(V-V_b)[1-e^{-2K'e(V-V_b)}] \quad (13)$$

and

$$dj_x/dV \cong G_0[1+(2K'e(V-V_b)-1)e^{-2K'e(V-V_b)}], \quad (14)$$

where

$$G_0 \equiv \frac{m_t e^2 \exp(-2K)}{4\pi^2\hbar^3 K'}. \quad (15)$$

Equation (14) gives the conductance which is the quantity usually measured. It rises linearly at biases close to the band edge, but at large values of $(V-V_b)$ approaches a constant value G_0, aside from variations in K.

The conductance in the approximation of Eq. (14) is shown in Fig. 3 plotted in units of G_0. This expression is equally valid for an empty conduction band or a filled valence band. In a real material, G_0 would vary with changes in K and K'. For tunneling between two metals, the conductance would simply be G_0. For tunneling between a metal and semiconductor, the conductance is best thought of as the product of a "transmission factor" G_0 and the modulating influence of the semiconductor forbidden band given by the square bracket in Eq. (14).

The value of the WKB integral K will profoundly affect the magnitude of current expected. Some previous treatments[6] of oxide layer tunneling have considered the barrier to be the equivalent of a vacuum layer or used an effective mass approach.[7] Such a treatment is not valid for the present problem. If such were the case, the value of K would always increase with increasing positive bias (i.e., metal positive) which is to say that G_0 would always decrease rapidly with increasing bias. Such behavior is not observed. Actually, it will be seen that G_0 may increase, decrease, or even pass through a minimum with bias changes. It is much more slowly varying than would be expected from a "vacuum layer" WKB approximation. When proper account is taken of the finite width of the oxide forbidden band, it is seen that k_x must vanish at both oxide band edges and, thus, reach a maximum value somewhere between them.[8] In this light, then, it is not surprising that G_0 may go through a minimum at some value of bias. The qualitative behavior of G_0 is shown in Fig. 4 along with the modulating influence of the semiconductor forbidden band.

Tunnel current between the metal and interface states is characterized by a frequency dependence. The origin of this dependence is best shown by noting that these are localized states, and that an electron occupying one of these states will have a definite lifetime before tunneling. If all final states in the metal are empty, the transition probability per unit time will be

$$w_a = (2\pi/\hbar) \sum_{k_t} |M_{ab}|^2 \rho_b. \quad (16)$$

Thus, the current will be frequency-dependent in the manner described in the previous section for frequencies of order w_a. If we consider the high-frequency case in which conductance is not limited by charging of the interface states, it will be seen that the conductance will be proportional to the density of interface states, but independent of the density of states in the metal. The conductance will also depend strongly on oxide thickness, as will the frequency w_a because of the WKB exponential factor in the matrix element.

The impurity states communicate readily with the valence band and so permit dc conductance measurements. The conductance is expected to be proportional to the density-of-impurity states except possibly in the degenerate case.

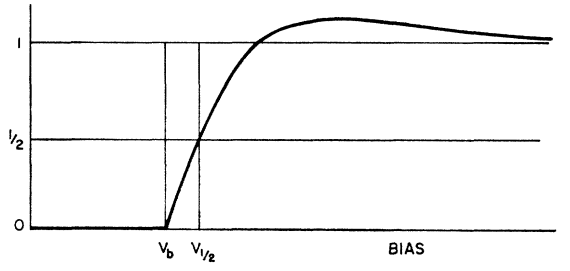

FIG. 3. Conductance (dj_x/dV) per unit area in units of G_0 as a function of bias. The rate of rise of the conductance depends on K' and this quantity may be estimated from an experimental estimate of $V_{1/2}$.

$$K' \approx 0.158 |e(V_{1/2}-V_b)|^{-1}.$$

III. EXPERIMENTAL PROCEDURE

1. Sample Preparation

Boron-doped silicon single crystals having resistivities of 0.014, 0.063, 0.25, and 2.2 Ω-cm were obtained from Merck and Company. The crystals were optically oriented and samples were cut to the approximate dimensions 5×3×0.5 mm³, with the 5-mm length being along a [111] direction. The samples were polished and etched in a small amount of CP-4A at room temperature. The etching process was quenched by quickly

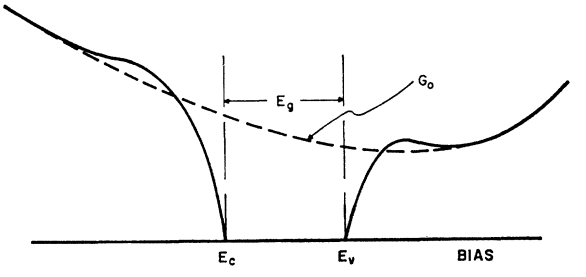

FIG. 4. Theoretical dc conductance versus bias for a sample. The transmission coefficient G_0 (dashed curve) has been drawn arbitrarily to illustrate its expected qualitative behavior. The bias values corresponding to the semiconductor band edges are E_c and E_v. The modulating effect on G_0 of a semiconductor having band gap E_g is apparent.

[6] R. Holm, J. Appl. Phys. **22**, 569 (1951).
[7] J. C. Fisher and I. Giaever, J. Appl. Phys. **32**, 172 (1961).
[8] Such behavior is exhibited by a two-band **k·p** perturbation analysis valid for some narrow band-gap semiconductors. See E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959).

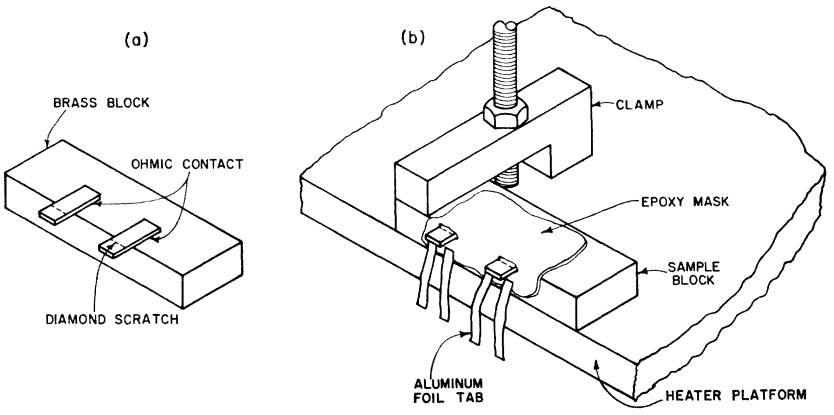

FIG. 5. Sample mounting. Figure 5(a) shows two samples soldered to the brass block. In Fig. 5(b), the epoxy mask and aluminum foil contacts are in place and the block is clamped to the heater platform in the vacuum jar (Fig. 6).

flooding with deionized water. The samples were rinsed with deionized water and dried in air.

Large-area Ohmic contacts were made by heating the samples to about 500°C and "tinning" half of one face with indium using an ultrasonic soldering tool. Small brass (or copper) blocks were similarly tinned and the silicon samples were soldered to the blocks by their Ohmic contacts as shown in Fig. 5(a). A light diamond scratch was made to aid cleavage. A low-vapor-pressure epoxy resin was used to mask the sample (up to the diamond scratch) and the area of the mounting block near the sample. Aluminum tabs were imbedded in the epoxy for electrical contacts.

Cleaving and oxidation were performed in the vacuum jar shown in Fig. 6. The sample block was clamped in good thermal contact to a thermostatically controlled, heated platform [Figs. 5(b) and 6]. The temperature was initially raised to approximately 130°C to outgas the sample area at a mechanical pump pressure of 10^{-3} mmHg. The temperature was then adjusted to a suitable value for oxidation (typically 100°C) and tank oxygen was admitted to a pressure of approximately 150 mmHg. The sample was then cleaved by means of a glass knife edge controlled from outside the system. Oxidation of the cleaved surface took place over a period of about 10 h. In a few instances, water vapor was introduced. No particular difference was noted in the sample properties. After oxidation, the sample was removed from the system and exposed to air for about 5 min while being masked and mounted in the evaporator. Aluminum was then evaporated onto the oxide layer, epoxy mask, and aluminum tabs to form a circuit for ease of electrical contact. No special care was taken to use high-purity aluminum although the tungsten filament and aluminum charge were heated to incandescence in vacuum prior to mounting the sample in the evaporator. Evaporation was carried out at pressures varying from 10^{-4} to 2×10^{-3} mmHg. No cold trap was used. The resistance of the aluminum film was typically 0.5 Ω/square. Final preparation of the sample involved scraping off parts of the aluminum film to assure adequate separation of the evaporated circuits. Usually two separate contacts were made on each cleaved surface.

2. Measuring Apparatus

The majority of the samples were measured at a single high frequency (14 kc/sec) and at dc. A few were measured at a number of frequencies using an impedance bridge. In all cases, it is the parallel conductance and capacitance which are plotted.

For the high-frequency measurements an automatic device was constructed which would plot either conductance or differential capacitance as a function of bias on an X-Y recorder. This device makes use of a phase discriminator to take advantage of the orthogonality of the resistive and reactive components of the ac current in the sample.

The dc current-voltage curves were also plotted automatically on an X-Y recorder with a Hewlett-Packard

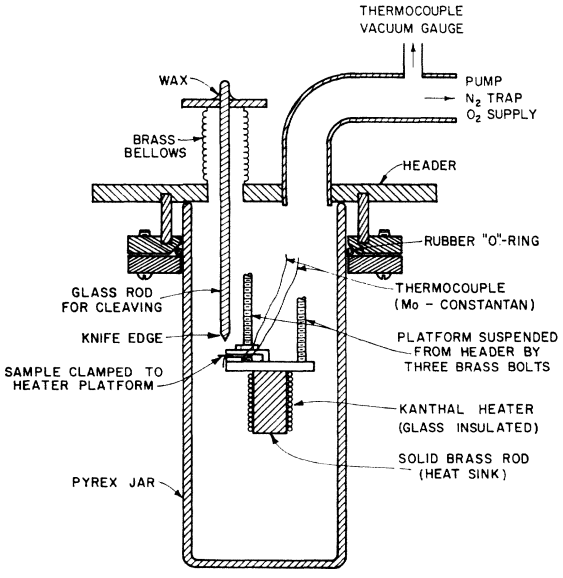

FIG. 6. The vacuum jar. Details of the sample mounting on the heater platform are shown in Fig. 5(b).

model 425A dc microvoltammeter used as a current amplifier. In all cases, the bias was changed very slowly to preserve steady-state conditions. The automatically plotted dc curves were graphically differentiated, and the derivative labeled "zero-frequency" conductance for comparison with the ac measurements.

IV. EXPERIMENTAL RESULTS AND INTERPRETATION

1. Tunneling to Intrinsic and Impurity Bands

Three typical plots of dc conductance as a function of bias are shown as the solid curves in Fig. 7. While these curves were taken at room temperature, similar results have been obtained over the temperature range of 77° to 373°K. The curves were selected to show how the transmission factor G_0 may increase, decrease, or pass through a minimum with increasing bias. The dashed lines are merely sketches of G_0 included to illustrate the interpretation. A minimum in G_0 can be seen in Fig. 7(b). The strong resemblance to the curve of Fig. 4 is evident. The approximate biases corresponding to silicon band edges have been labeled E_c and E_v in Fig. 7. Note that the value of the silicon band gap (1.1 eV) is given with reasonable accuracy except in 7(b) where flattening of the bands has resulted in excess current at negative bias. A similar effect is observed in Figs. 10 and 11. In all of the data presented the observed structure will be smeared in voltage by several kT/e.

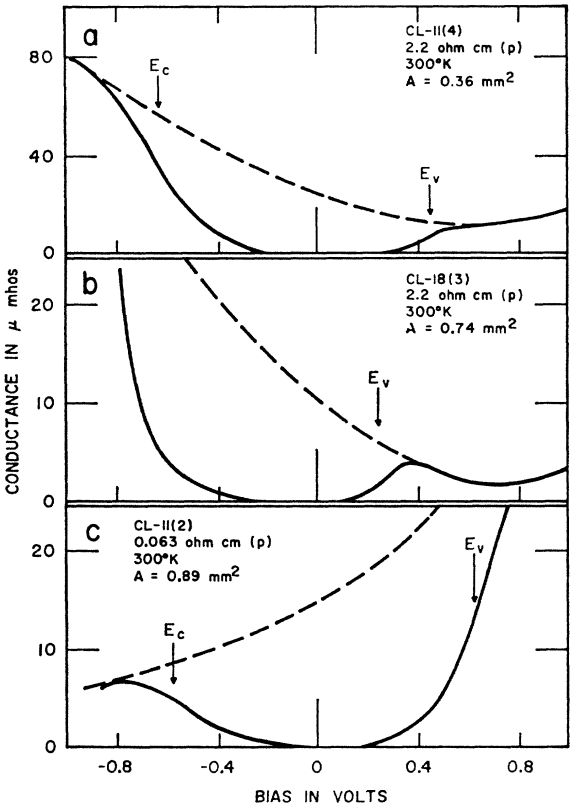

FIG. 7. Experimental dc conductance versus bias. The dashed curves represent G_0 and have been included to aid comparison with Fig. 4. Sample areas (A) are given on each graph.

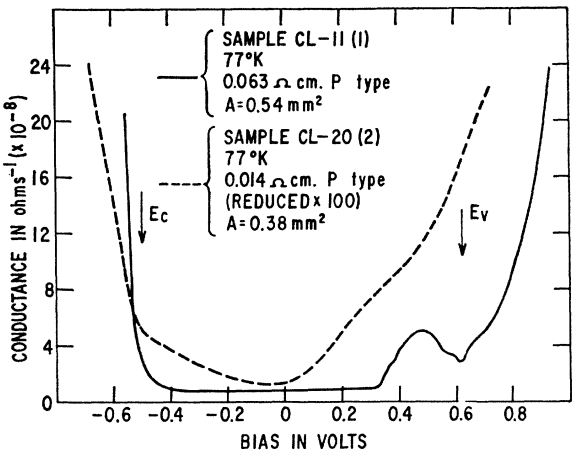

FIG. 8. dc conductance versus bias showing the effect of a boron-impurity band at two different doping levels. The conductance of the more heavily doped sample is reduced ×100 for comparison.

At sufficiently high doping levels, a peak in the dc conductance curve resulting from tunnel transitions to the impurity band is observed as shown in Fig. 8 (solid curve). The impurity band is clearly shown at this impurity concentration of $\sim 10^{18}/cm^3$. The apparent excessive width of the impurity band is due to band bending at the onset of conduction. The effect is much less pronounced with thicker oxides (and lower current densities) as in Fig. 9. At lower doping levels ($\sim 10^{17}/cm^3$), a faint trace of a bump was observed during one run. At still higher doping levels ($\sim 10^{19}/cm^3$) the impurity band has broadened and merged with the

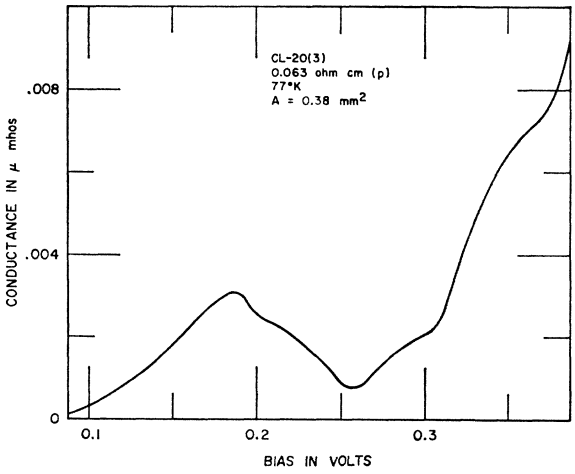

FIG. 9. Detailed dc conductance near an impurity band (0.19 V). This sample was cleaved in air before mounting in the vacuum jar.

valence band as shown by the dashed curve of Fig. 8. This effect has been observed as a disappearance of impurity ionization energy,[9,10] and more recently as a mechanism to explain excess current in tunnel diodes.[11] In Fig. 9 an enlarged view of the region near the impurity band is shown. The structure in this curve is reproducible among samples cut from this crystal. At this point it is not possible to identify the dip in the curve at 0.31 V, although it may be associated with the spin-orbit split-off band.

The curves of Figs. 8 and 9 were made at 77°K because at this low temperature the leakage current via interface states is small.

2. Tunneling to Interface States

The interface states are best studied in rather lightly doped samples in which the intrinsic band edges are not excessively disturbed by impurity states. Figure 10 shows the differential capacitance versus bias at various frequencies when the impurity content is $\sim 10^{16}/cm^3$. The peak in the low-frequency curves is due to a band of interface states. A similar peak occurs at negative bias (see Fig. 12), but is obscured in Fig. 10 by diffusion capacitance. At the highest frequency no appreciable charging of the interface states can take place. The capacitance then is that of the oxide layer and space-charge region in series. The decrease in capacitance with increasing bias of the 8-kc/sec curve of Fig. 10 is due to the increase in width of the space-charge region as the semiconductor bands are bent downward.

The ac conductance, measured simultaneously with the capacitance of Fig. 10 is shown in Fig. 11. At high frequency it is proportional to the density-of-interface

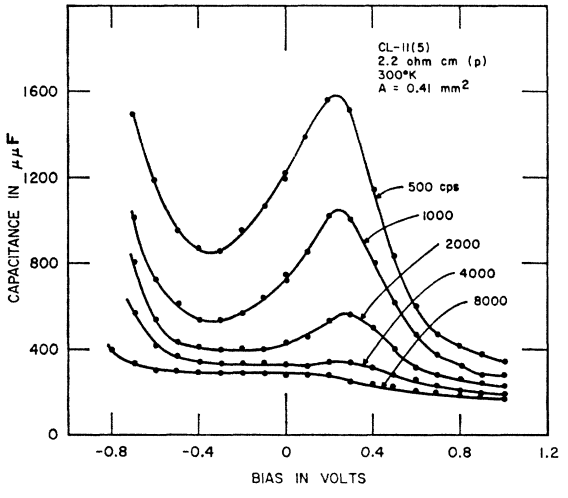

FIG. 10. Differential capacitance versus bias at several frequencies.

[9] G. L. Pearson and J. Bardeen, Phys. Rev. **75**, 865 (1949).
[10] R. K. Ray and H. Y. Fan, Phys. Rev. **121**, 768 (1961).
[11] T. P. Brody, J. Appl. Phys. **33**, 100 (1962).

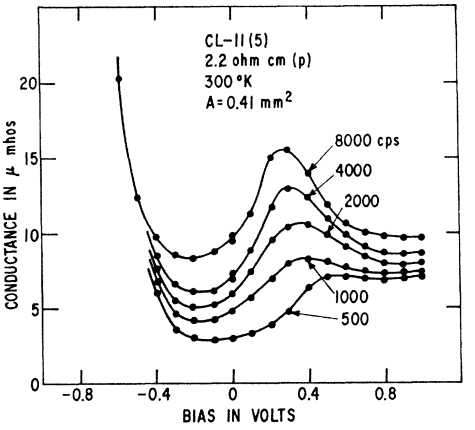

FIG. 11. Conductance versus bias at several frequencies. These data were obtained simultaneously with those of Fig. 10.

states, but also shows the effect of transitions to the intrinsic bands. At low frequencies conductance to interface states is greatly reduced due to charging. Automatically plotted conductance measurements at 14 kc/sec were nearly identical to the 8-kc/sec data shown. A few points were obtained at 50 kc/sec and these also fell close to the 8-kc/sec data.

The data shown in Figs. 10 and 11 were taken soon after the sample was prepared. After aging several weeks, it was possible to obtain useful data in the region of the conduction-band edge. Before aging, it is believed that tunneling of holes from the valence band resulting from flattening of the semiconductor bands obscured measurements in this region. Figure 12(a) shows the

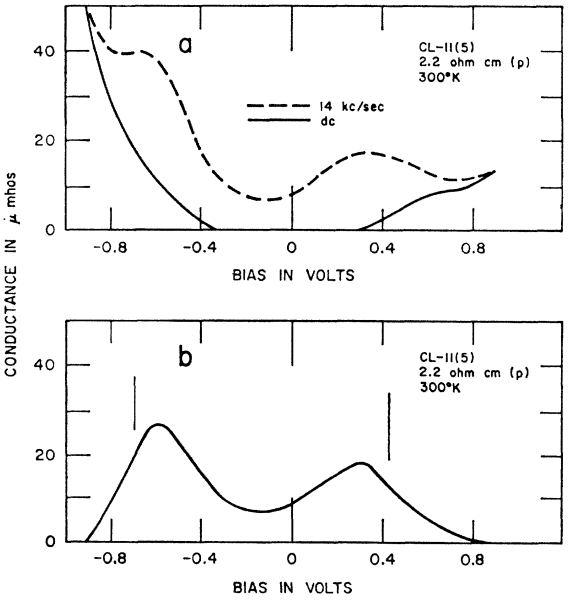

FIG. 12(a). High-frequency and dc conductance versus bias. (b). The difference between the two curves in (a). Two bands of interface states are seen.

high-frequency (dashed line) and zero-frequency (solid line) conductance plotted on the same graph. Since the zero-frequency conductance is almost completely unaffected by interface states while the conductance due to intrinsic states is independent of frequency, the two curves have been subtracted and the difference has been plotted in Fig. 12(b). This difference curve should be proportional to the density-of-interface states and to the transmission factor G_0. The constant of proportionality is obtained from the capacitance data of Fig. 10 using Eq. (1) of Sec. I. These data then not only provide a description of the interface states, but also because the true density of states is found experimentally permit analysis of the tunnel parameters involved.

Measurements of the density and energy of such interface states have been made by Statz *et al.*,[12] by means of inversion layer conductance studies. They find a band of about 1.4×10^{12} states/cm^2 located 0.455 eV below the intrinsic Fermi level, and another band with 10^{11} to 10^{12} states/cm^2 located approximately 0.44 to 0.48 eV above the intrinsic Fermi level. The peak density of states of the curve shown in Fig. 12(b) at 0.3 V bias is 2.4×10^{12} states/eV cm^2. Integration of this peak from -0.1 to 0.8 V bias gives a total of 1.5×10^{12} states/cm^2 below the intrinsic Fermi level in excellent agreement with Statz. The fact that the left-hand peak is larger is undoubtedly due to reduced barrier height. The energy separation of the peaks is quite close to Statz' estimate of 0.92 ± 0.02 eV.

V. DISCUSSION

The properties of the oxide film are most critical to the success of the experiment. A very high resistivity, particularly at high electric fields, is essential and this fact may severely limit the number of different semiconductors which may be investigated this way. In addition, the rapid variation of tunneling probability with thickness requires close control over the oxide growth rate.

The oxide thickness of the samples discussed here is probably quite uniform. The possibility that current is flowing through a few thin spots or even pin holes in the oxide has been considered. If this were so, there would be a wide distribution of time constants (w_a^{-1}) associated with the frequency-dependent measurements; a situation which is not observed. The fairly close quantitative agreement with the known density-of-interface states indicates that at least an appreciable fraction of the surface is involved. The oxide uniformity is achieved, as indicated, by slow oxidation. Another important advantage of low-temperature oxidation is that the bulk-impurity distribution is not upset by possible "surface-gettering" effects. This added to the fact that oxidation occurs by oxygen ion migration through the silicon-dioxide film[13] means that the bulk-impurity concentration is probably preserved in the oxide and at the interface.

The oxide thickness is estimated to be 40–70 Å. The basis of this estimate is the observation that for heavily doped ($\gtrsim10^{18}$/cm^3) samples, the capacitance at biases corresponding to the silicon forbidden band is large and nearly independent of frequency. It is believed that the large number of final states available for tunneling permits the true geometrical capacitance of the oxide layer to be measured in these cases. These estimates assume that the rate of oxide growth and dielectric constant are independent of impurity concentration, and that the bulk silica dielectric constant is appropriate.

All of the room-temperature curves show marked hysteresis effects, particularly when the bias excursions are large. At 77°K these effects are greatly reduced. The precise mechanism responsible for this polarization of the oxide is not known, nor is it possible to state whether it is an intrinsic property of the oxide or due to impurities. It is apparent, however, in view of the very small oxide thickness and the long time constants associated with recovery ($\sim\frac{1}{2}$ h) that these "slow states" are not electron traps. They are most likely mobile ions in the oxide.

VI. CONCLUSIONS

The technique of oxide-layer tunneling into semiconductors has been shown to yield information about the energy-band structure of the semiconductor and about the nature of the oxide-semiconductor interface in agreement with other experiments. A satisfactory method of sample fabrication has been described and the nature of the problems involved in performing and interpreting these experiments has been discussed.

There are certain measurements for which the present method appears especially well suited. They are (1) determination of the density and energy of interface states, (2) examination of the details of impurity banding at high doping levels, and (3) analysis of the properties of the insulating oxide film. Further refinements in sample preparation and measuring technique will be necessary before these ends can be fully realized.

ACKNOWLEDGMENTS

The author wishes to express his gratitute to Professor John Bardeen for suggesting this problem and for his guidance throughout the work, and to Professor Paul Handler for his continued interest and helpful discussions. In addition, he wishes to thank Henry Pao and Johnie Driver for their aid in preparing some of the samples; and the General Electric Company for a fellowship granted during a portion of the work.

[12] H. Statz, G. A. deMars, L. Davis, Jr., and A. Adams, Jr., Phys. Rev. **101**, 1272 (1956); *ibid.* **106**, 455 (1957).

[13] P. J. Jorgensen, J. Chem. Phys. **37**, 874 (1962).