THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XLVI JULY–AUGUST 1967 NUMBER 6

The Si-SiO$_2$ Interface – Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique

By E. H. NICOLLIAN and A. GOETZBERGER

(Manuscript received December 28, 1966)

Measurements of the equivalent parallel conductance of metal-insulator-semiconductor (MIS) capacitors are shown to give more detailed and accurate information about interface states than capacitance measurements. Experimental techniques and methods of analysis are described. From the results of the conductance technique, a realistic characterization of the Si–SiO$_2$ interface is developed. Salient features are: A continuum of states is found across the band gap of the silicon. Capture cross sections for holes and electrons are independent of energy over large portions of the band gap. The surface potential is subject to statistical fluctuations arising from various sources. The dominant contribution in the samples measured arises from a random distribution of surface charge. The fluctuating surface potential causes a dispersion of interface state time constants in the depletion region. In the weak inversion region the dispersion is eliminated by interaction between interface states and the minority carrier band. A single time constant results. From the experimentally established facts, equivalent circuits accurately describing the measurements are constructed.

I. INTRODUCTION

The electrical properties of semiconductor surfaces have been studied for a long time. Until recently most investigators were concerned with either etched surfaces covered by a thin natural oxide layer or with ultraclean surfaces that exist only in high vacuum. Development of improved oxidation techniques for silicon have made it now possible to study the electrical phenomena occurring at the interface between silicon and silicon dioxide. These studies resulted in the recognition that a semiconductor-insulator interface behaves differently in many ways from a "bare" surface. Of practical importance is the fact that an oxide-covered surface is more stable and can be made electrically more perfect than an unprotected surface.

The difference between oxidized and bare surfaces will be listed here briefly.

(*i*) There is no charge exchange with states at the air oxide interface. This type of state, the "slow surface state," has a long time constant and is encountered on etched surfaces. For thick oxide layers, only states at the oxide-semiconductor interface have to be considered. They are equivalent to "fast states" and will be called interface states in this paper. In addition to interface states, there may be traps within the oxide.[1, 2] Not much is known about these traps and they will not be a major topic of this paper.

(*ii*) All oxidized silicon surfaces contain a so-called surface charge. It consists of positive charges residing close to the interface. The surface charge originates from two sources, alkali ions[3] and built-in charges.[4, 5] Alkali ions can drift through the oxide at fairly low temperatures and thus be distinguished from built-in or residual charges which are fixed and have a total density which depends on oxidation rate and crystal orientation.[6] The present paper deals with interface states in alkali-free systems. Reference to residual surface charge will be necessary in connection with their influence on the uniformity of surface potential.

(*iii*) Thermally oxidized Si–Si0_2 interfaces are characterized by a much lower density of states than "bare" surfaces. The density of states over most of the silicon band gap in the oxidized systems is usually not much greater than 10^{12} cm^{-2}-eV^{-1} or smaller than 10^{10} cm^{-2}-eV^{-1} depending on the method of oxide preparation. For "bare" surfaces, density of states can be as high as 10^{14} cm^{-2}-eV^{-1}.

The electrical properties of interface states are characterized by their density, their position in the energy gap of the silicon, and their capture cross section. In addition, it is necessary to know whether they are of the donor or acceptor type. The most widely used tool for investigating these interface state properties is the metal-insulator-semiconductor (MIS) capacitor.

Dispersion of the capacitance can be used to obtain information about the energy distribution and density of interface states as has been shown by Terman.[7] The capacitance technique, however, has severe limitations.[8] Essentially, the difficulty is that interface state capacitance must be extracted from measured capacitance which consists of oxide capacitance, depletion layer capacitance, and interface state capacitance. This difficulty does not apply to the equivalent parallel conductance because conductance arises solely from the steady-state loss due to the capture and emission of carriers by interface states and is thus, a more direct measure of these properties.[9] Conductance measurements yield more accurate and reliable results, particularly when the density of interface states is low as in the thermally oxidized system because only directly measured quantities are used with no matching to a theoretical curve required. Both the capacitance and equivalent parallel conductance as functions of voltage and frequency contain identical information about interface states. Greater inaccuracies arise in extracting this information from the capacitance. This is illustrated in Fig. 1 which shows capacitance

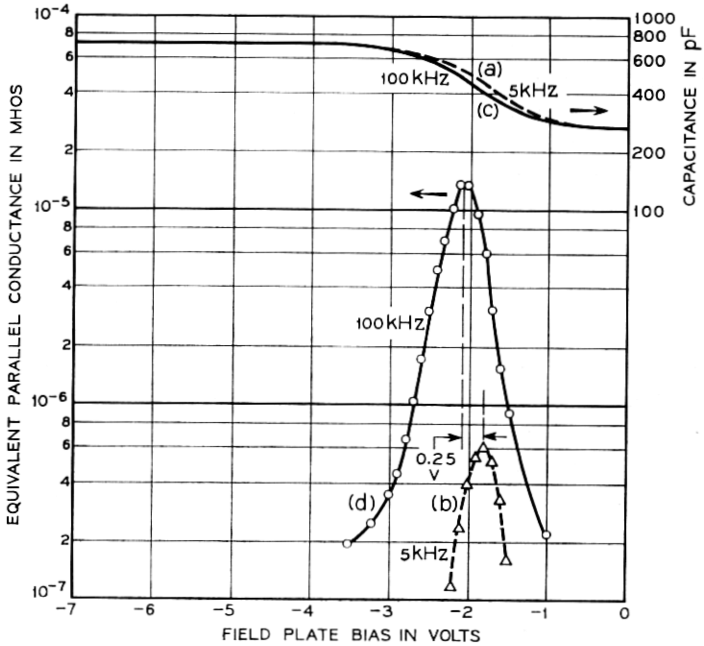

Fig. 1 — Capacitance and equivalent parallel conductance measured at 5 kHz and 100 kHz on a p-type sample having an acceptor density of 2.08×10^{16} cm^{-3} and interface state density in the 10^{11}cm^{-2}-eV^{-1} range. Interface state time constants and densities are given in Fig. 14 and Fig. 15 for this sample.

and equivalent parallel conductance measured at 5 kHz and 100 kHz. The largest capacitance spread is 14 percent while the magnitude of the conductance peak increases by over one order of magnitude in this frequency range.

In order to evaluate MIS measurements with regard to interface states, it is desirable to have one-dimensional current flow perpendicular to the interface. This condition does not apply when the silicon underneath the field plate and beyond is inverted. Then, the capacitance and conductance are dominated by the lateral ac current flow that has been explained elsewhere.[10, 11] Because of the positive polarity of surface charge, p-type silicon is normally inverted and is therefore most easily investigated in the depletion-accumulation region.

In the MIS capacitor, the density of states is normally low so that the largest silicon band bending possible is not limited by their influence. Maximum possible band bending in the MIS capacitor is slightly less than the band gap. In this paper, interface state properties are described for energies in the forbidden gap within the range $2u_B$ [see Fig. 5(a)] only at room temperature.

Lehovec and Slobodskoy[12] theoretically treat losses due to steady state capture and emission of both minority* and majority* carriers by a hypothetical single level interface state for the cases of zero and infinite recombination-generation rate through states in the silicon space-charge region. No direct experimental measurement and evaluation of these losses has been made until recently.[9] The theory and equivalent circuits developed in Refs. 12 and 13 constitute a very general description of all the losses and charge storage processes possible in an MIS structure. The gap between this general theory and what is actually measured in a sodium-free MIS structure is bridged in this work. It can be shown by conductance measurements that:

(*i*) loss in the oxide measured at room temperature is negligible,

(*ii*) in the samples studied, the dominant loss is by capture and emission of carriers by interface states rather than states in the silicon space charge region, and

(*iii*) the majority carrier time constant is always the dominant one.

These three results make it possible to study the capture and emission properties of the interface states for each type of carrier independently. It is this fact which makes the MIS capacitor a powerful tool. We shall divide the range of surface potential from accumulation to

---

* The term minority carrier denotes those carriers which are in a minority in the bulk of the semiconductor crystal. The term majority carrier denotes those in a majority in the bulk of the crystal.

inversion into four regions each of which is characterized by a different equivalent circuit. These regions are: depletion-accumulation, mid-gap, weak inversion, and strong inversion. The depletion-accumulation region includes those values of surface potential less than u_B where u_B is shown in Fig. 5(a). Mid-gap is the region a few kT/q volts wide about u_B. Weak inversion is between u_B and $2u_B$. Strong inversion includes those values of surface potential greater than $2u_B$. The following results have been obtained using the MIS conductance technique.

(*i*) A continuum of states across the bandgap of the silicon appears to be characteristic of the Si–SiO_2 interface.

(*ii*) Capture cross sections for holes and electrons are independent of energy over large portions of the bandgap.

(*iii*) Time constant dispersion at each bias in the depletion-accumulation region is dominated by statistical fluctuations of surface potential. In the samples measured, these fluctuations are caused primarily by a random distribution of built-in residual oxide charges and charged interface states over the plane of the interface.

(*iv*) A single time constant at each bias is observed in the region of weak inversion because the minority carrier capture resistance becomes negligibly small and the inversion layer provides a large lateral conductance across the silicon surface.

(*v*) Equivalent circuits are derived for the depletion-accumulation, mid-gap, and weak inversion regions. (See Fig. 31) The equivalent circuit which applies in strong inversion has been worked out previously.[11, 14] This circuit was proposed in Ref. 11 for the case where lateral ac current flow is absent and has been verified by experiment using n-type samples.[14]

II. SAMPLE PREPARATION

The sample preparation processing described in this section is intended to yield oxides which remain stable only during admittance-voltage measurements made at room temperature or lower. The stability requirements for device applications would be more stringent than this and would require somewhat different processing. No attempt is made to minimize built-in charge density. Both high and low interface state densities are purposely produced.

Measurements were made on both boron and arsenic doped silicon oriented in the [111] and [100] directions. Samples were prepared by first growing a 10-μ thick epitaxial layer having a nominal resistivity of 1 ohm-cm on a 0.005 ohm-cm substrate. The substrate always incorporates the same dopant as the epitaxial layer. The purpose of using an epitaxial layer is to minimize bulk series resistance to the point where it can be neglected—an important consideration when making conductance measurements.

The slices about 1 cm square were removed from the epitaxial grower and immediately placed in the oxidation apparatus described in Ref. 4. An oxide typically between 500 Å and 700 Å thick was then grown at 1000°C in steam at atmospheric pressure by the bias technique of Ref. 4.

Another group of samples was prepared by growing about the same thickness of oxide the same way except in dry oxygen at 1000°C and atmospheric pressure. Trace quantities of H_2 were removed from the O_2 stream by first passing it through a Deoxo* unit. The O_2 was then dried by passing it through a dry ice acetone trap. The oxide should be as thin as possible without being in the tunnelling range to make the measured admittance correspond more closely to the admittance of the interface. Oxidizing the sample immediately after the epitaxial layer is grown insures that the silicon surface has not had much opportunity to gather contamination from its surroundings. The oxide was grown by the bias technique to obtain an oxide in which ionic contamination is a minimum. The reason for taking pains to minimize ionic contamination in the oxide is that the presence of an ion near an interface state may perturb its energy and capture probability although such ions do not themselves act as interface states.[3]

Immediately after oxidation, the slice was placed in an oil diffusion pump evaporator and a number of chrome-gold field plates were evaporated through a metal mask held in intimate contact with the oxide surface. First, a few hundred Å of chromium was deposited to make a strong bond to the oxide. This was followed by 2000 Å of gold to provide good electrical contact.

The chromium and gold sources were heated to evaporation temperature in adjacent tungsten baskets. The baskets are 15 cm above the surface of the sample to reduce penumbra effects.[8] Since not only capacitance but also equivalent parallel conductance depends on the area of the field plate, a large field-plate diameter is used to get a large admittance thereby increasing detection sensitivity. However, making the field-plate diameter large increases the likelihood of encompassing defects and gross nonuniformities under it. A satisfactory diameter with these considerations in mind can be as large as

---

* This unit contains a catalyst which enables H_2 and O_2 to react at room temperature producing H_2O.

1500 μ. The oxide on the substrate side was then etched off in HF. Cr–Au in the same proportions as in the field plate was evaporated all over this side to make the back contact. The slice was placed in good thermal contact with a massive copper block during all the evaporation steps to keep it always at room temperature. This minimizes sodium migration into the oxide during evaporation. Minimizing sodium contamination in this step of the processing also reduces the likelihood of a patchy charge distribution underneath the field plate.

III. ORIGIN OF THE EQUIVALENT PARALLEL CONDUCTANCE

Two experiments are described which show that the dominant process causing the measured equivalent parallel conductance is capture and emission of carriers by interface states. Series resistance is not important because it can be made negligible by using epitaxial samples or it can be measured in strong accumulation and substracted from the total impedance. Normally, small-signal measurements are made so that harmonics of the signal frequency arising from the nonlinearity of the charge-voltage characteristic are unimportant (see Section 5.1).

3.1 *Loss in the Oxide*

The first experiment shows that equivalent parallel conductance measured at room temperature does not arise from loss in the oxide. Capacitance and equivalent parallel conductance for a p-type sample are plotted as functions of field plate bias in Fig. 2 before and after aging an unstable oxide in room air at 150°C. This aging produces an increased positive charge density in the oxide as seen by the shift of the capacitance curve to higher negative bias. This means that the electric field in the oxide resulting from the applied bias is greater for a given silicon surface potential after aging. However, the conductance curve has not changed appreciably in shape or magnitude. The conductance curve is shifted to higher negative bias by the same amount as the capacitance curve. This is seen by the fact that the peak of the conductance occurs at the same value of capacitance before and after aging. Therefore, the conductance is a function of silicon surface potential and not of field in the oxide. Therefore, it is not related to loss in the oxide.

3.2 *Loss in the Silicon Space-Charge Region*

The most important processes which can give rise to the measured conductance are capture and emission of carriers by interface states or by states in the silicon space-charge region. It is interesting to note that the MIS capacitor can be used to measure the properties of either type of state provided one or the other dominates the loss processes.* Processing determines the densities and positions in the bandgap of levels of each type of state. In all samples made as described in Section II, interface state loss is dominant over the entire measurable range. In all these samples, no impurities other than arsenic or boron

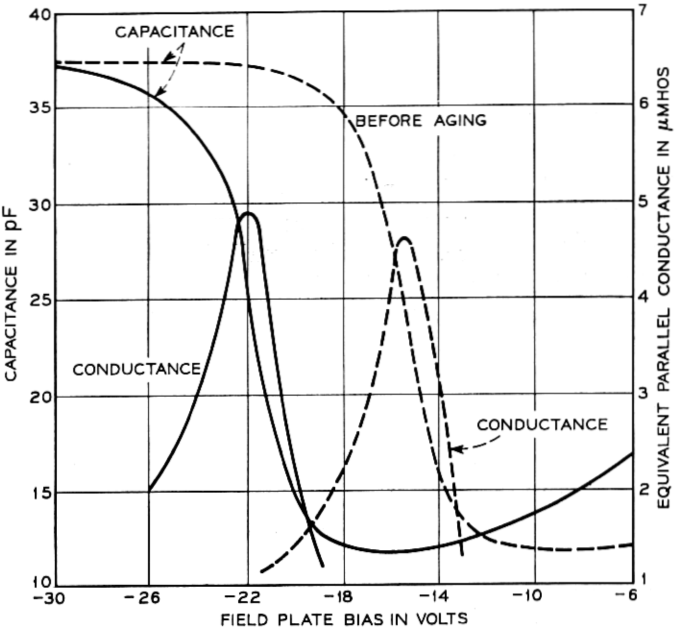

Fig. 2 — Measured capacitance and equivalent parallel conductance as functions of field plate bias before and after aging at 150°C in room air. Field plate diameter is 3.8 × 10^{-2}cm, silicon resistivity 50 Ω-cm, p-type, crystal orientation [111], oxide thickness 1200 Å, and measurement frequency 100 kHz.

are purposely introduced into the silicon. The densities and energy levels of other impurities or defects inadvertently introduced into the silicon during fabrication of the MIS capacitor are unknown. The same is true for interface states. It is best therefore, to establish by experiment which effect dominates. This is described next.

The magnitude of the equivalent parallel conductance measured is always much smaller when the oxide is grown in steam than when the

---

* Sah[15] has used the MIS capacitor to study states introduced into the silicon by heavy gold doping.

oxide is grown in dry oxygen, other parameters being the same.[4] In addition, the oxide can be reversibly cycled from the wet to the dry condition by heat treatment between 200°C and 400°C in various ambients. To enable a quantitative comparison, it is necessary that the resistivity of the silicon, oxide thickness, and field plate area be the same for both conditions. The easiest way to accomplish this is to cycle the same capacitor from the wet to a drier condition. Curves (a) and (b) in Fig. 3 are plots of capacitance and equivalent parallel

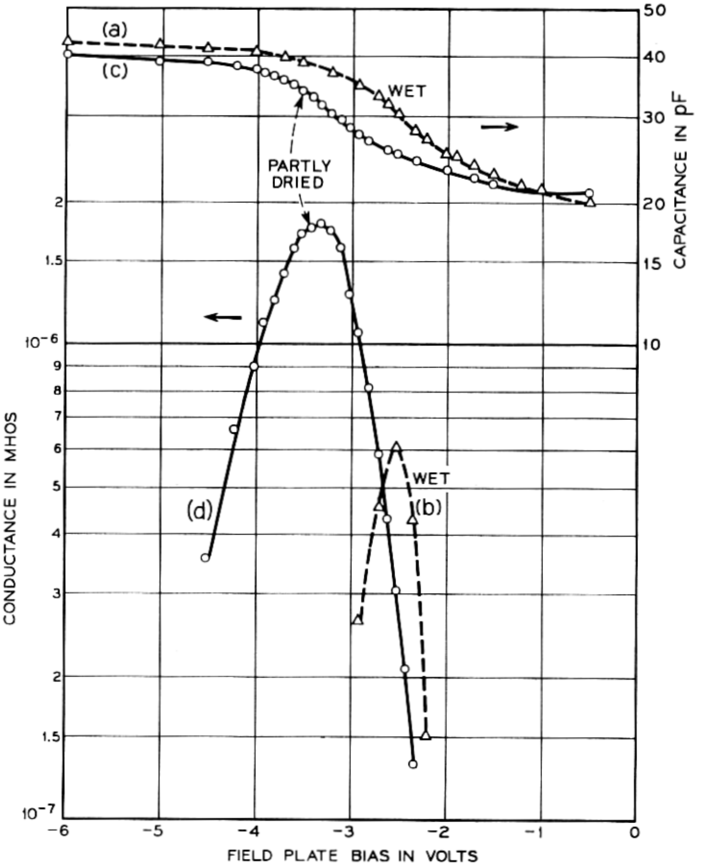

Fig. 3 — (a) capacitance and (b) equivalent parallel conductance vs field plate bias before drying; (c) capacitance and (d) equivalent parallel conductance vs field plate bias of the same capacitor after heating in dry nitrogen for 17 hours at 350°C. Field plate diameter is 3.8×10^{-2} cm, silicon resistivity 1 Ω-cm, p-type, crystal orientation [111], oxide thickness 910 Å grown in stream, and measurement frequency 100 kHz.

conductance for a steam grown oxide on p-type silicon. Curves (c) and (d) in this figure are measured after heating the capacitor in dry nitrogen (dried by passing through a liquid nitrogen trap) for 17 hours at 350°C. A similar set of curves for n-type is given in Fig. 4. All the curves in these two figures are measured at 100 kHz. Curves similar to these can be obtained at all frequencies in the range investigated (50Hz to 500 kHz). It is important to note only that the capacitance in Fig. 3 from curve (c) associated with the conductance peak of curve

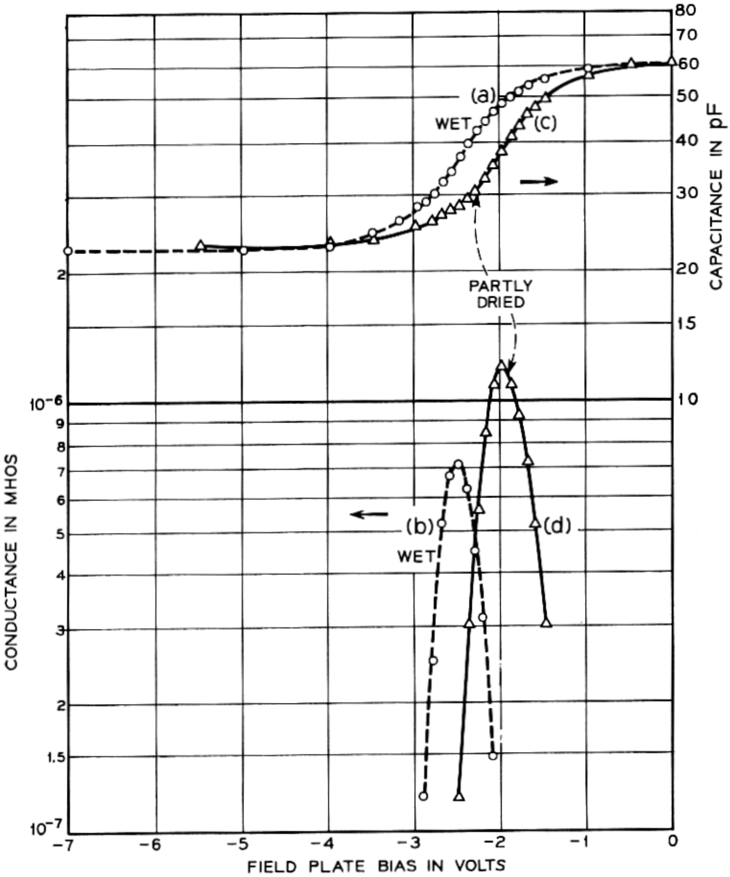

Fig. 4 — (a) capacitance and (b) equivalent parallel conductance measured vs field plate bias before drying; (c) capacitance and (d) equivalent parallel conductance vs field plate bias of the same capacitor after heating in dry nitrogen for 17 hours at 350°C. Field plate diameter is 3.8×10^{-2}cm, silicon resistivity 0.75 Ω-cm, n-type, crystal orientation [111], oxide thickness 600 Å grown in stream, and measurement frequency 100 kHz.

(d) is nearly the same as the capacitance from curve (a) associated with the conductance peak of curve (b). The same is true in Fig. 4. Therefore, all the relevant parameters before and after drying are the same except the magnitudes of the conductance peaks. The peak conductance of curve (b) in Fig. 3 is about one-third the peak conductance of curve (d). This result cannot be due to sodium contamination which might have been introduced into the oxide during the drying cycle, the initial wet oxide being sodium free (see Section II). The reasons for this are: (*i*) Sodium at the interface always remains charged in the range between accumulation and strong inversion.[8] Therefore, it does not become an interface state which could influence the conductance. (*ii*) A nonuniform charge distribution in the oxide caused by the sodium would spread the conductance curve over a larger range of bias which would decrease rather than increase the magnitude of the conductance peak if this effect were dominant. (*iii*) The p-type curves in Fig. 3 are shifted towards more negative bias while the n-type curves in Fig. 4 are shifted towards less negative bias with drying. Introduction of sodium during drying would shift both curves towards more negative bias. The increase of the conductance peak in both p and n-type is then a true drying effect.

The fact that the oxide can be reversibly cycled from the wet to the dry state at low temperatures rules out the possibility that a large density of structural defects or impurities which act as bulk states is introduced into the silicon during one method of oxide preparation and not the other. All of this strongly suggests that the measured equivalent parallel conductance is due mainly to losses in states which reside on the oxide side of the interface and exchange charge at 50 Hz to 500 kHz with the silicon because only the properties of such states could be so markedly altered by low temperature wet-dry cycling.

The most convincing evidence, however, is the dependence of interface state time constants on surface potential to be discussed later in this paper.

3.3 *Loss Due to Minority Carrier Transitions*

Minority carrier density can respond to the applied signal frequency by diffusion from the bulk to the interface and by recombination-generation processes through interface states and states in the silicon space charge region. These processes in the samples measured are too slow to allow the minority carrier density to follow 50 Hz to 500 kHz.[11, 14] It was indicated in the preceding part of this section that in this frequency range, interface state loss is solely responsible for the measured equivalent parallel conductance. In the depletion region where minority carrier density is orders of magnitude less than majority carrier density, loss due to minority carrier transitions will be negligible compared to loss due to majority carrier transitions. Thus, in this region, the conductance measured arises simply from capture and emission of majority carriers by those interface states within a few kT/q of the Fermi level. In the region of weak inversion where minority carrier density is orders of magnitude greater than majority carrier density, minority carrier time constant becomes shorter than majority carrier time constant. Minority carriers will follow the signal frequency. However, current can flow and produce a loss only by transitions through the interface states from the minority carrier band to the majority carrier band. Therefore, this process will be controlled by the majority carrier time constant. This case will be treated in detail in Section IV. The important point is that the loss is determined entirely by the majority carrier time constant in both the depletion and weak inversion regions.

IV. THEORY

4.1 *Phenomenological Description of Loss*

A continuum of interface states over the bandgap of the silicon appears to be characteristic of the Si–SiO$_2$ interface. Fig. 5(a) is a schematic of an MIS capacitor showing the band bending in depletion for n-type silicon with an arbitrary distribution of interface states in the silicon bandgap. Initially, let the interface states be in equilibrium with the silicon. Then let a small ac signal between 50 Hz and 500 kHz be applied. Consider the first half of the cycle which moves the conduction band towards the Fermi level. This immediately increases the average energy of the electrons in the silicon. Because a conductance is observed, it is evident that the interface states do not respond immediately but lag behind. Therefore, electrons at a higher average energy in the silicon will be captured by interface states at a lower average energy. This results in an energy loss. On the other half of the cycle, the signal moves the valence band towards the Fermi level. Electrons in filled interface states now will be at a higher average energy than electrons in the silicon. As the electrons are emitted by the interface states into the silicon they will lose energy again. Thus, there will be an energy loss on both halves of the cycle which must be supplied by the signal source. The energy required for transitions between the band edge and an interface state is much higher than this and is

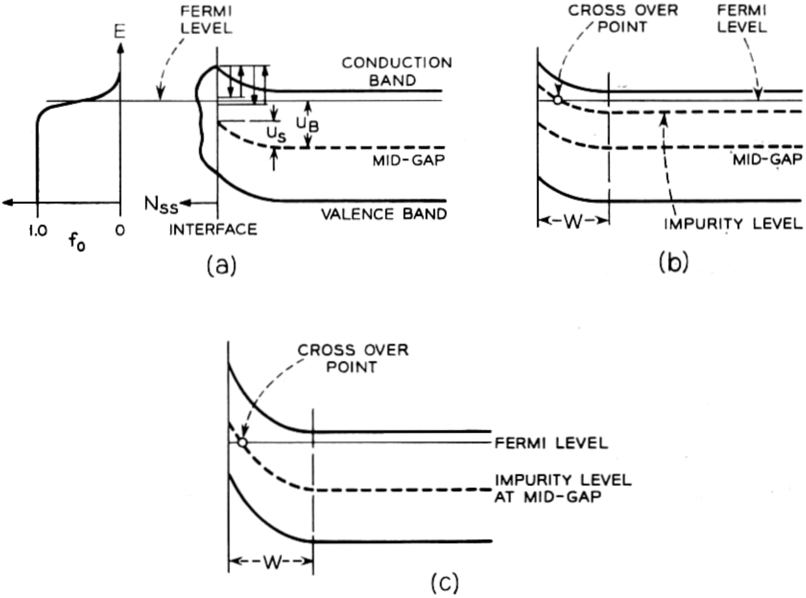

Fig. 5 — (a) Schematic showing band bending in depletion at silicon surface in an n-type MIS capacitor. An arbitrary distribution of interface states is shown for illustrative purposes. Majority carrier transitions to two levels within kT/q of the Fermi level are also shown. f_o is the Fermi function. (b) Band bending in depletion showing point at which Fermi level crosses a shallow impurity level. This impurity level is hypothetical. W is space charge width. (c) Band bending in weak inversion showing point at which Fermi level crosses a hypothetical impurity level located at mid-gap.

supplied by phonons. Only majority carrier capture and emission is considered here because this is the dominant loss mechanism for reasons given in Section III.

After a disturbance has been applied, the interface states will reach equilibrium with the silicon exponentially with a characteristic time, τ. The interface states are also capable of storing charge. Thus, there will be a capacitance, C_s, associated with them which is proportional to their density. The conductance related to the loss is $G_s = C_s/\tau$. It should be noted that G_s is independent of signal amplitude as long as C_s and τ are independent of it. This will be the case for small-signal amplitude. Interface state τ and C_s can then be extracted directly from a measurement of equivalent parallel conductance. Capture cross section is extracted from τ and density of states from C_s as will be shown in Section VII.

4.2 *Admittance of a Single Level State*

Our purpose is to derive expressions for the admittance of the MIS capacitor as a function of bias and frequency, and to obtain the corresponding equivalent circuits which fit in detail to experimental observations. The capacitance-voltage characteristics for the MIS capacitor without interface states has been described.[15, 16, 17] Capacitance dispersion caused by interface state loss was measured by Terman[7] to obtain interface state densities and time constants. Lehovec et al,[12] Bertz,[13] and Sah[15] developed equivalent circuits for a hypothetical single-level interface state from Shockley-Read statistics.[18] The equivalent circuits given in this previous work are either too general[12, 13, 15] or the measurement technique too inaccurate[7] to give a real and complete description of interface state properties. To describe the MIS conductance technique and interpret the measurements in the depletion and weak inversion regions, we shall rederive the theory for the admittance of the MIS capacitor. The starting point will be the theory for majority carrier capture and emission by a single-level interface state first calculated by Lehovec and Slobodskoy.[12]

Experimental evidence shows that only capture and emission of majority carriers are important when measuring in the depletion region. The quasi Fermi level and Fermi level are identical for majority carriers. Thus, application of an ac signal simply results in a time varying Fermi function. From Shockley and Read,[18] the capture rate of electrons, taken as majority carriers, by a single-level interface state is

$$R_n(t) = N_s c_n [1 - f(t)] n_s(t), \tag{1}$$

and the emission rate is

$$G_n(t) = N_s e_n f(t), \tag{2}$$

where N_s is the density of states cm^{-2}; c_n electron capture probability,* cm^3/sec, e_n electron emission constant, sec^{-1}, $f(t)$ the Fermi function at time t, and $n_s(t)$ electron density at the silicon surface at time t, cm^{-3}.

Net current density flowing is

$$i_s(t) = qN_s c_n [1 - f(t)] n_s(t) - qN_s e_n f(t), \tag{3}$$

where q is electronic charge in coulombs. Making the small-signal approximation, the admittance is, as is shown in Appendix A,

---

* This capture probability is the average over all states in the conduction band *near the silicon surface.*

$$Y_s = j\omega \frac{q^2}{kT} \frac{N_s f_o(1 - f_o)}{(1 + j\omega f_o/c_n n_{so})}, \tag{4}$$

where $j = \sqrt{-1}$, ω the angular frequency of the ac signal, sec^{-1}, k Boltzman's constant, eV × coul/°K, T the absolute temperature, °K, f_o the Fermi function established by the bias, $f_o = [1 + \exp(u - u_s)]^{-1}$, and n_{so} the electron density at the silicon surface established by the bias, cm^{-3}.

Equation (4) is the admittance of a series RC network with capacitance $C_s = q^2 N_s f_o(1 - f_o)/kT$ and time constant $\tau = f_o/c_n n_{so}$. Separating (4) into its real and imaginary parts, the equivalent parallel capacitance is

$$C = \frac{C_s}{1 + \omega^2\tau^2}, \tag{5}$$

and the equivalent parallel conductance is

$$G_p = \frac{C_s\omega^2\tau}{1 + \omega^2\tau^2}. \tag{6}$$

4.3 *Equivalent Circuit of MIS Capacitor*

Let Q_T be the total charge density at a given bias. Then

$$Q_T = Q_{sc} + Q_s + Q_f, \tag{7}$$

where Q_{sc} is the silicon space-charge density, coul/cm^2, Q_s the interface state charge density, coul/cm^2, and Q_f the fixed-charge density in the oxide, coul/cm^2. The ac current density, $i_T(t)$, obtained by differentiating (7) with respect to time is

$$i_T(t) = i_{sc}(t) + i_s(t), \tag{8}$$

where $i_{sc}(t)$ and $i_s(t)$ are the ac current densities charging the silicon space-charge layer and the interface states, respectively, To obtain $i_{sc}(t)$ we have

$$i_{sc}(t) = \left(\frac{dQ_{sc}}{d\psi_s}\right)\left(\frac{d\psi_s}{dt}\right), \tag{9}$$

where $\psi_s(t) = \psi_{so} + \delta\psi_s$ is the silicon band bending in volts at time, t, ψ_{so} the silicon band bending established by the bias, and $\delta\psi_s = a \exp(j\omega t)$. From this, $d\psi_s/dt = j\omega\, \delta\psi_s$. Also, $dQ_{sc}/d\psi_s = C_D$ the depletion layer capacitance per cm^2. Substituting these into (9) we get

$$i_{sc} = j\omega C_D\, \delta\psi_s. \tag{10}$$

It is shown in Appendix A that $i_s(t)$ given by (3) can be written

$$i_s(t) = Y_s \, \delta\psi_s \, , \tag{11}$$

where Y_s is defined by (4). Substituting (10) and (11) into (8), we have

$$i_T(t) = (j\omega C_D + Y_s) \, \delta\psi_s \, . \tag{12}$$

From (12), it is seen that C_D appears in parallel with the series RC network of the interface states. The voltage, $v_a = v_o + \delta v_a$, applied to the capacitor divides between the silicon and the oxide film so that

$$v_a(t) = v_o + \delta v_a = \psi_s(t) + \frac{Q_T}{C_{ox}} , \tag{13}$$

where v_o is the dc bias, $\delta v_a = b \exp (j\omega t)$, and C_{ox} is the oxide capacitance per cm^2. Differentiating (13) with respect to t to get the ac terms only

$$j\omega \, \delta v_a = j\omega \, \delta\psi_s + \frac{i_T(t)}{C_{ox}} . \tag{14}$$

Substituting for $\delta\psi_s$ from (12)

$$\delta v_a = i_T(t)\left(Z_s + \frac{1}{j\omega C_{ox}}\right) , \tag{15}$$

where

$$Z_s = (j\omega C_D + Y_s)^{-1} .$$

The bracketed term in (15) is the impedance of a circuit in which C_{ox} is in series with Z_s. The equivalent circuit for the MIS capacitor from (4), (12), and (15) is given in Fig. 6.

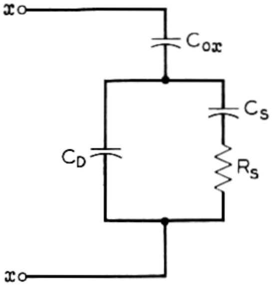

Fig. 6 — Equivalent circuit of MIS capacitor for capture and emission of majority carriers by a single level interface state. C_{ox} is the oxide capacitance per unit area, C_D the depletion layer capacitance per unit area, C_s the interface state capacitance per unit area, and R_s the majority carrier capture resistance. Bulk generation-recombination is zero.

The principle of the MIS conductance technique is easily illustrated by this simplified equivalent circuit. The admittance of the capacitor is measured by a bridge across the terminals x–x. The oxide capacitance is measured in the region of strong accumulation. The admittance of the network is then converted into an impedance. The reactance of the oxide capacitance is subtracted from this impedance and the resulting impedance converted back into an admittance. This leaves C_D in parallel with the series RC network of the interface states. The capacitance from (5) and (12) is

$$C_p = C_D + \frac{C_s}{1 + \omega^2\tau^2}, \tag{16}$$

and the equivalent parallel conductance divided by ω from (6) is

$$\frac{G_p}{\omega} = \frac{C_s\omega\tau}{1 + \omega^2\tau^2}. \tag{17}$$

G_p is divided by ω to make (17) symmetrical in $\omega\tau$. It should be noted that (16) and (17) are equivalent to the Debye equations,[19, 20] for a dielectric in which the polarization can relax exponentially with characteristic time, τ, after a change in applied electric field occurs.

Equation (16) describes the capacitance dispersion and is the basis of Terman's method. To extract C_s and τ from C_p using (16), C_D must be known. C_D can be calculated[16] using an estimated doping density. The doping density is not accurately known near the silicon surface because of pile up or depletion of the dopant during oxidation. Equation (17) on the other hand depends *only* on the interface state branch of the equivalent circuit. G_p/ω goes through a maximum when $\omega\tau = 1$ which gives τ directly. The value of G_p/ω at the maximum is $C_s/2$. Thus, *equivalent parallel conductance* corrected for C_{ox} gives C_s and τ *directly* from the *measured* conductance.

4.4 *Admittance of an Interface State Continuum*

The interface states are observed to be comprised of many levels so closely spaced in energy that they cannot be distinguished as separate levels. Thus, they appear as a continuum over the bandgap of the silicon. These observations are documented in Section VII and discussed in Section VIII and agreed with those of other investigators.[7, 21] A continuum of interface states appears to be characteristic of the Si–SiO$_2$ interface. The admittance given by (4) for a single-level state and the equivalent circuit in Fig. 6 must therefore be modified.

For a continuum of states at a finite absolute temperature, capture and emission of majority carriers can occur by states located within a few kT/q on either side of the Fermi level. This results in a time constant dispersion. The admittance of the continuum first calculated by Lehovec[22] is obtained by integrating (4) over the bandgap

$$Y_{ss} = j\omega\left(\frac{q^2}{kT}\right) \int \frac{N_{ss} f_o (1 - f_o)\, d\psi}{(1 + j\omega f_o / c_n n_{so})}. \tag{18}$$

Here N_{ss} is the density of interface states, $\text{cm}^{-2}\ \text{eV}^{-1}$, and ψ is energy in ev. The integrand of (18) is sharply peaked about the Fermi level with a width of about kT/q. Thus, (18) can be easily integrated if both N_{ss} and c_n the capture probability* do not vary very much with ψ over a range kT/q. Experimentally, it is found (see Section VII) that neither vary appreciably over kT/q. Making the substitution $f_o(1-f_o) = (kT/q)\ (df_o/d\psi)$ transforms (18) into an integral over f_o. Integrating from zero to unity yields.

$$Y_{ss} = \frac{qN_{ss}}{2\tau_m} \ln\,(1 + \omega^2\tau_m^2) + jq\,\frac{N_{ss}}{\tau_m} \text{arc tan}\,(\omega\tau_m) \tag{19a}$$

where

$$\tau_m = 1/c_n n_{so}. \tag{19b}$$

Concentrating on the real part of (19a), we have for the continuum

$$\frac{G_p}{\omega} = \frac{qN_{ss}}{2\omega\tau_m} \ln\,(1 + \omega^2\tau_m^2). \tag{20}$$

At no value of bias does (20) even closely fit experimentally measured G_p/ω vs ω curves. As will be shown in the next section, (20) is just the first step in deriving an expression for G_p/ω which fits the experiment in depletion. To proceed further, it is necessary to consider separately the depletion region and the weak inversion region. In depletion, the time constant dispersion observed is far broader than given by (20). In weak inversion, the time constant dispersion disappears and the experimental G_p/ω vs ω curves are characterized by a single time constant at each bias.

4.5 *Depletion Region*

The model suggested here for the increased broadening of the time constant dispersion observed experimentally (see Section VII) in the depletion region assumes statistical fluctuations of surface potential in the plane of the interface. Majority carrier density at the silicon

---

* This capture probability is now the average over all the interface states in the range of the integral and all states in the conduction band near the silicon surface.

surface is related to surface potential by

$$\begin{aligned} n_{so} &= N_D \exp u_s = n_i \exp (u_s - u_B) && \text{for n-type} \\ \text{and} \quad p_{so} &= N_A \exp (-u_s) = n_i \exp -(u_s - u_B) && \text{for p-type,} \end{aligned} \tag{21}$$

where n_i is intrinsic carrier density, cm^{-3}, $u_B = \ln n_i/N_D$ for n-type and $u_B = \ln N_A/n_i$ for p-type. u_B is the potential difference between mid-gap* and Fermi level in the quasi neutral region of the silicon in units of kT/q, N_D is the ionized donor and N_A the ionized acceptor density in the silicon, cm^{-3}, u_s is the silicon band bending or surface potential in units of kT/q. $u_s - u_B$ is the potential difference between mid-gap and Fermi level at the silicon surface.

The relation between time constant and surface potential is from (19b) and (21).

$$\begin{aligned} \tau_m &= \frac{1}{c_n n_i} \exp -(u_s - u_B) && \text{for n-type} \\ \text{and} \quad \tau_m &= \frac{1}{c_p n_i} \exp (u_s - u_B) && \text{for p-type.} \end{aligned} \tag{22}$$

Equation (22) shows that small fluctuations in u_s will cause large fluctuations in τ_m and thus broadened time constant dispersion. If we assume that the built-in charges and charged interface states are randomly distributed in the plane of the interface, the electric field at the silicon surface will fluctuate over the plane of the interface. This is shown schematically in Fig. 7. Fluctuations in electric field will cause corresponding fluctuations in surface potential. To express this quantitatively,† we conceptually divide the plane of the interface into a number of squares of equal area. The area of each square is the largest area within which the surface potential is uniform. We shall call this particular area a characteristic area. Now, the admittance of the continuum given by (19) can be regarded as the result of integrating the admittance of a single level (4) over all the levels located within a characteristic area rather than within the entire area of the field plate. The total admittance is obtained by integrating the contribution from each characteristic area over all the characteristic areas within the area of the field plate. This must be done for each bias and frequency to generate the complete family of curves. The calculated

---

* Although the intrinsic Fermi level is slightly above mid-gap in silicon, this is neglected and intrinsic Fermi level is taken to be at mid-gap.

† This treatment is similar to the one used by W. Shockley in Ref. 23 for calculations of the influence of fluctuations of doping on avalanche breakdown.

G_p/ω curves will be compared with the measured G_p/ω curves to see whether the model agrees with observation. For simplicity, we shall derive in detail only an expression for the real part of the interface state admittance. The steps in deriving the imaginary part are similar.

Let $P(N)$ be the probability that there are N built-in charges and charged interface states in a characteristic area. The number of characteristic areas which contain between N and $N+dN$ surface charges is

$$d\nu = P(N)\, dN. \tag{23}$$

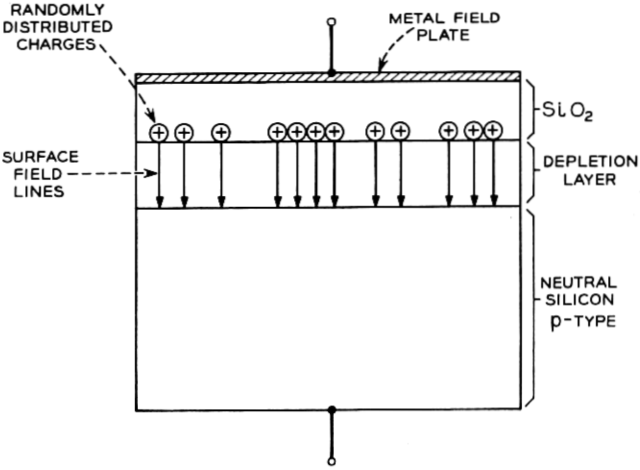

Fig. 7 — Cross section of MIS structure illustrating the random distribution of built-in charges and charged interface states over plane of the interface. Field lines are all terminated at edge of depletion layer rather than on ionized acceptor impurities in the layer for simplicity.

The total G_p/ω is G_p/ω for the continuum of levels located in a characteristic area (20) times the number of characteristic areas containing between N and $N+dN$ surface charges (23) integrated over all the characteristic areas. To integrate this product over all the characteristic areas, (20) is expressed as a function of u_s using (22) and $P(N)$ transformed to $P(u_s)$, the probability that the surface potential in a characteristic area is u_s. This transformation will be performed in two steps. The first transformation is

$$P(Q) = P(N)\, dN/dQ\ , \tag{24}^*$$

---

* See Ref. 24 for a discussion of transforming the probability density as a function of one random variable to another.

where Q is the density of built-in charges and charged interface states, coul/cm². The second transformation is

$$P(u_s) = P(Q)\, dQ/du_s\,. \tag{25}$$

The number of characteristic areas in which the surface potential is between u_s and $u_s + du_s$ is $P(u_s)du_s$. Multiplying this by (20) and integrating, the total G_p/ω is

$$G_p/\omega = \frac{qN_{ss}}{2}\int_{-\infty}^{\infty}\frac{1}{\omega\tau_m}\ln\,(1 + \omega^2\tau_m^2)P(u_s)\, du_s\,. \tag{26}$$

It is assumed in (26) that both interface state density and capture probability are uniform over several kT of bandgap energy. That these are reasonable assumptions can be seen from Fig. 14 and 15.

Substituting (22) for p-type* into (26) to get (26) as a function of u_s

$$G_p/\omega = \tfrac{1}{2}qN_{ss}\int_{-\infty}^{\infty}\exp\,(u_B - u_s - u_0)$$
$$\cdot\ln\,\{1 + \exp\,[2(u_0 + u_s - u_B)]\}P(u_s)\, du_s\,, \tag{27}$$

where $u_0 = \ln\,\omega/c_p n_i$.

The next task is to obtain $P(u_s)$ from $P(N)$. Let $\bar{N}$ be the mean number of built-in charges and charged interface states in a characteristic area. When $\bar{N}$ is large, $P(N)$ is given by the Gaussian approximation of a Poisson distribution

$$P(N) = (2\pi\bar{N})^{-\frac{1}{2}}\exp\,[-(N - \bar{N})^2/2\bar{N}]. \tag{28}$$

The characteristic area α is the ratio between the mean number of surface charges $\bar{N}$ in α and their mean density $\bar{n}$: $\alpha = \bar{N}/\bar{n}$. From this, the relation between N and Q required for the first transformation is

$$N = \alpha Q/q. \tag{29}$$

Combining (24), (28), and (29), we get

$$P(Q) = (2\pi\alpha\bar{Q}/q)^{-\frac{1}{2}}(\alpha/q)\exp\,[-\alpha(Q - \bar{Q})^2/2q\bar{Q}], \tag{30}$$

where $\bar{Q}$ is the mean of Q.

The transformation of $P(Q)$ to $P(u_s)$ consists of transforming an area in the space made up of all the points Q into an area in the space made up of all the points u_s. This requires a single-valued relation between

---

* This part of the derivation is carried out for p-type because the sign of the variables for n-type is negative making the analysis harder to follow.

Q and u_s. A relation between Q and u_s which is not single valued is obtained from (7) and (13). Retaining only the dc terms in (13) and expressing surface potential in units of $kT/q = \beta^{-1}$, we get

$$Q = Q_s + Q_f = C_{ox}(v_o + u_s/\beta) + Q_{sc}(u_s). \tag{31*}$$

$Q_{sc}(u_s)$ is the space-charge density which from Ref. 16 is

$$Q_{sc}(u_s) = (2q\epsilon_{si}N_A/\beta)^{\frac{1}{2}}[\exp(-u_s) + u_s - 1]^{\frac{1}{2}}, \tag{32}$$

where ϵ_{si} is the dielectric permittivity of silicon, farads/cm and N_A the ionized acceptor density, cm^{-3}. Equation (31) is not a single-valued relationship between Q and u_s because of (32). The difficulty in making the transformation from Q space to u_s space imposed by (31) and (32) is avoided by restricting the problem to the case where the fluctuations $Q - \bar{Q}$ are very small. For small fluctuations, (31) can be differentiated assuming N_A and oxide thickness to be uniform to get

$$dQ = \frac{C_{ox}}{\beta}\, du_s + dQ_{sc}. \tag{33}$$

dQ_{sc} is eliminated from (33) by using the relation $dQ_{sc} = (C_D/\beta)\, du_s = (\epsilon_{si}/W\beta)\, du_s$ where W is the space-charge width. Because $\bar{Q}$ is given by (31) when $u_s = \bar{u}_s$, dQ can be evaluated about $\bar{Q}$ at each bias by setting $W = W(\bar{u}_s)$. Doing this, (33) becomes

$$dQ = \frac{1}{W(\bar{u}_s)\beta}[W(\bar{u}_s)C_{ox} + \epsilon_{si}]\, du_s. \tag{34a}$$

Equation (34a) is the transformation equation we seek from an element of area in Q space to the corresponding element in u_s space. Replacing dQ and du_s in (34a) by the small fluctuations $Q - \bar{Q}$ and $u_s - \bar{u}_s$, respectively, we get

$$Q - \bar{Q} = \frac{1}{W(\bar{u}_s)\beta}[W(\bar{u}_s)C_{ox} + \epsilon_{si}](u_s - \bar{u}_s). \tag{34b}$$

Combining (25), (30), (34a), and (34b), we get

$$P(u_s) = (2\pi\sigma_s^2)^{-\frac{1}{2}} \exp[-(u_s - \bar{u}_s)^2/2\sigma_s^2], \tag{35}$$

where the standard deviation σ_s is

$$\sigma_s = \frac{W(\bar{u}_s)\beta}{[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}\left(\frac{q\bar{Q}}{\alpha}\right)^{\frac{1}{2}}. \tag{36}$$

---

* In this paper, $Q_f + Q_s$ will be calculated from flat band voltage shift neglecting work function differences. Work function differences are neglected because flat band voltage shifts of several volts independent of the field plate metal are found in the samples used.

$\bar{Q}$ can be calculated from (31) and (32) at $u_s = \bar{u}_s$. In the samples measured, the total number of built-in charges was more than an order of magnitude greater than the total number of charged interface states. Therefore, the bias dependence and the variation during each cycle of the signal of the number of charged interface states is negligible.

Another cause of surface potential fluctuations which shall be taken into account is the random distribution of ionized acceptors in the space charge region. This is calculated in Appendix B using a model which has previously been used to determine the fluctuations of breakdown voltage.[23,25] Essentially, the procedure is similar to the one for the surface charges except a characteristic cube having a side equal to W is used instead of a characteristic area and Q rather than N_A is assumed to be uniformly distributed. The standard deviation from Appendix B, paragraph B.1 is

$$\sigma_B = \frac{q\beta[\bar{N}_A W(\bar{u}_s)]^{\frac{1}{2}}[1 - \exp(-\bar{u}_s)]}{2[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}, \tag{37}$$

where $\bar{N}_A$ is the mean ionized acceptor density.

A theorem in statistics states[26] that if two or more independent variables are normally distributed their sum is also normally distributed with a mean and variance which is the sum of the mean and variance of each variable. Using this theorem, $P(u_s)$, which includes the influence of both the randomly distributed surface charges and ionized acceptors on the fluctuations of surface potential is

$$P(u_s) = [2\pi(\sigma_s^2 + \sigma_B^2)]^{-\frac{1}{2}} \exp[-(u_s - \bar{u}_s)^2/2(\sigma_s^2 + \sigma_B^2)]. \tag{38}$$

Substituting (38) into (27)

$$G_p/\omega = \tfrac{1}{2}qN_{ss}[2\pi(\sigma_s^2 + \sigma_B^2)]^{-\frac{1}{2}} \cdot \int_{-\infty}^{\infty} \exp[-(z + y)] \ln(1 + e^{2y})\, du_s, \tag{39}$$

where $y = \ln \omega\tau_m = u_0 + u_s - u_B$ and $z = (u_s - \bar{u}_s)^2/2(\sigma_s^2 + \sigma_B^2)$.

By similar arguments using (19), the equivalent parallel capacitance is

$$C_p = C_D(\bar{u}_s) + qN_{ss}[2\pi(\sigma_s^2 + \sigma_B^2)]^{-\frac{1}{2}} \cdot \int_{-\infty}^{\infty} \exp[-(z + y)] \arctan(e^{y})\, du_s. \tag{40}$$

Depletion layer capacitance C_D now becomes a function of $\bar{u}_s$. The function $C_D(\bar{u}_s)$ has the same form as in Ref. 16 as long as the fluctuations $Q - \bar{Q}$ are small.

It will be shown in Section VIII that the influence of the randomly distributed ionized acceptors on the fluctuations of surface potential are small in the samples measured compared to the influence of built-in charges.

In summarizing the theory for the depletion region, it is shown how the interface state branch of the equivalent circuit is developed. The interface state branch of the equivalent circuit of Fig. 6 applies to a single level in the continuum. In each characteristic area where the surface potential is uniform, each level contributes a series RC network. Integrating the admittance of each level given by (4) over all the levels located in a characteristic area gives (19) and (20). Integrating again but this time the contribution to the total admittance from each characteristic area over all the characteristic areas gives (39) and (40). The interface state branch of the equivalent circuit therefore, can be represented by an infinite number of series RC networks connected in parallel as illustrated in Fig. 8(a).

4.6 *Weak Inversion*

For n-type, the silicon surface will be in weak inversion when the Fermi level is more than a few kT/q below mid-gap in the lower half of the bandgap. The minority carrier density at the silicon surface will now be orders of magnitude greater than the majority carrier density

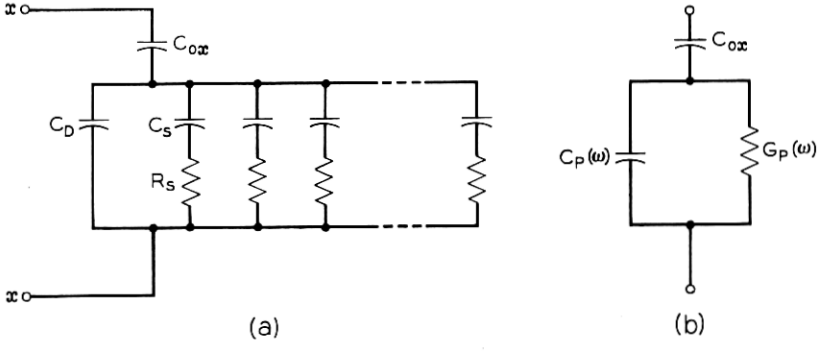

Fig. 8 — (a) Equivalent circuit for depletion region showing time constant dispersion caused primarily by statistical fluctuations of surface potential. C_{ox} is the oxide capacitance per unit area, and C_D the depletion layer capacitance per unit area. Each subnetwork consisting of C_s and R_s in series represents time constant of the continuum of interface states in a characteristic area. Bulk generation-recombination is zero. (b) Simplified version of (a) useful when extracting C_{ox} from the measured admittance. $C_p(\omega)$ is the capacitance per unit area at a given bias and frequency of the distributed network in parallel with C_D. $G_p(\omega)$ is the equivalent parallel conductance per unit area.

at the silicon surface. The minority carriers can no longer be ignored as in depletion and the equivalent circuit of Fig. 8(a) does not apply. To derive expressions for the interface state admittance we start again by considering a single level state. The capacitance associated with this state is $C_s = q^2 N_s f_o(1 - f_o)/kT$ and its time constant for majority carrier transitions (electrons) is $\tau_{n,s} = f_o/c_n n_{so}$ as derived previously in paragraph 4.2. The majority carrier capture resistance from this is

$$R_{n,s} = \tau_{n,s}/C_s = (kT/q^2)[c_n n_{so} N_s (1 - f_o)]^{-1}. \tag{41}$$

By similar reasoning, it can be shown that the minority carrier (holes) capture resistance is

$$R_{p,s} = (kT/q^2)[c_p p_{so} N_s f_0]^{-1}, \tag{42}$$

where c_p is the capture probability for holes cm^3/sec and p_{so} is the hole density cm^{-3} at the silicon surface when the Fermi level is at the interface state level.

The equivalent circuit for a single level is given in Fig. 9(a). From (41) and (42)

$$R_{n,s}/R_{p,s} = (c_p/c_n) \exp (u_F - 3u). \tag{43}$$

Equation (43) has been obtained by substituting $p_{so} = n_i \exp(-u)$, $n_{so} = n_i \exp(u)$, and $f_o = [1 + \exp(u - u_F)]^{-1}$ into (41) and (42). u is the potential difference between mid-gap and the interface state level in units of kT/q and u_F is the potential difference between mid-gap and the Fermi level at the silicon surface in units of kT/q. There will be no detectable loss due to transitions of electrons and holes to and from interface states unless $u \approx u_F$ because $f_o(1 - f_o)$ is so sharply peaked around the Fermi level. Equation (43) becomes

$$R_{n,s}/R_{p,s} = (c_p/c_n) \exp (-2u_F). \tag{44}$$

It is seen from (44) that $R_{n,s}$ will rapidly become orders of magnitude greater than $R_{p,s}$ when the Fermi level gets more than a few kT/q below mid-gap even if c_p is one or two orders of magnitude smaller than c_n. Time constant for majority carrier transitions is $\tau_{n,s} = C_s R_{n,s}$, and for minority carrier transitions it is $\tau_{p,s} = C_s R_{p,s}$. Thus, $\tau_{n,s}$ will be much greater than $\tau_{p,s}$ in weak inversion. If the period of the applied signal frequency is comparable to $\tau_{n,s}$ as it is in these experiments, then there will be virtually no loss associated with minority carrier transitions. Each interface state near the Fermi level is charged and discharged instantaneously by minority carriers so that $R_{p,s} \approx 0$.

Because $R_{p,s}$ is negligibly small, it can be replaced by a short circuit

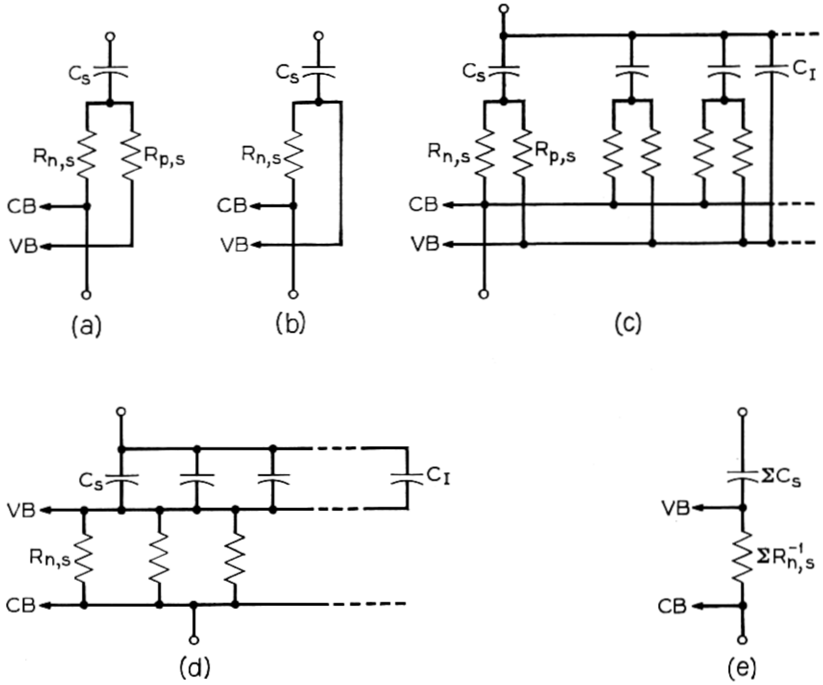

Fig. 9 — (a) Equivalent circuit for a single level interface state in region of weak inversion. C_s is the interface state capacitance per unit area, $R_{n,s}$ the capture resistance for electrons, and $R_{p,s}$ the capture resistance for holes. *CB* represents the conduction band and *VB* the valence band. (b) Equivalent circuit for a single level interface state for the case in which minority carriers are holes which respond instantaneously to the applied signal. For this representation, it is also necessary that the lateral conductance due to minority carriers along the surface is large compared to $R_{n,s}^{-1}$. (c) Circuit (a) generalized to include the fact that there is a continuum of states across the bandgap of silicon. C_I is the inversion layer capacitance per unit area. (d) Circuit (c) simplified by fast response of minority carriers. C_I is negligible in the region of weak inversion. (e) Simplification of circuit (d).

which connects the interface state capacitance directly to the minority carrier (valence) band as shown in Fig. 9(b). Those interface states in the continuum which are located in a given characteristic area each have a different time constant. These interface states must be represented by a distributed network such as Fig. 9(c). Only those levels in a narrow range of energy about a kT wide centered about the Fermi level participate in the capture and emission process. Therefore, the corresponding range of time constants is narrow. The large minority carrier density at the surface then means that the minority carrier capture resistance $R_{p,s}$ of every one of these levels is negligibly small. Thus, the capacitance corresponding to each of these levels can be connected directly to the minority carrier band. The large inversion layer conductance shorts all the characteristic areas together. The net result is represented in Fig. 9(d) which can be further simplified to Fig. 9(e). G_p/ω vs ω now will be characterized by a single time constant at each bias. The expression for G_p/ω is (17) and for C_p it is (16). The value of C_s in (16) and (17) is the sum of all the individual interface state capacitances in parallel as shown in Fig. 9(c). The inversion layer capacitance C_I calculated in weak inversion using the relation given in Ref. 12 is negligible compared to ΣC_s obtained from measurement.

If interface state density N_{ss} is slowly varying with the position of the Fermi level, the statistical fluctuations of surface potential will have negligible effect on the total C_s. However, in analyzing the measurements on n-type samples in weak inversion using (17), it is found that C_s varies quite rapidly with position of the Fermi level (see Fig. 27). This implies that N_{ss} is also rapidly varying so that the statistical fluctuations of surface potential must be taken into account when calculating the relation between N_{ss} and total C_s. Total capture conductance is the sum of the capture conductances $R_{n,s}^{-1}$ of the individual interface states in parallel as shown in Fig. 9(e).

The statistical fluctuations of surface potential must be taken into account in calculating the total capture conductance because capture conductance depends on n_{so} as well as the rapidly varying N_{ss}. If desired, C_s can be calculated by integrating C_s for a single level over bandgap energy and then over surface potential using the statistical model developed earlier. Similarly, total capture conductance can be calculated by integrating capture conductance for a single level obtained from (41) over bandgap energy and then over surface potential. These integrals cannot be evaluated without knowing the functional dependence of N_{ss} on the position of the Fermi level. This dependence can be found by numerically fitting the calculations to the measurements. Although N_{ss} and majority carrier capture probability can be obtained in this way from measured C_s and τ_m, such an analysis will not be carried out here.

A comparison between the depletion and weak inversion regions will illustrate the processes occurring. In depletion except very near mid-gap, majority carrier density at the silicon surface is several orders of magnitude greater than minority carrier density at the silicon surface. Therefore, the time constant for minority carrier transitions is much greater than the time constant for majority carrier transitions. The time constant for majority carrier transitions is comparable to the period of the applied signal frequency. Therefore, minority carrier density cannot follow the signal at all and ac current flows simply by transitions between interface states near the Fermi level and the majority carrier band. Thus, the interface state branch of the equivalent circuit in Fig. 8(a) can be obtained by open circuiting $R_{p,s}$ in Fig. 9(c).

In the weak inversion region except very near mid-gap, again the time constant for majority carrier transitions is comparable to the period of the signal frequency and several orders of magnitude greater than the time constant for minority carrier transitions. Now *both* majority and minority carrier densities respond to the signal so that ac current flows by generation and recombination through interface states near the Fermi level. This current flow is controlled by the majority carrier capture resistance as it is the largest. In this range, the interface states can easily communicate with the minority carrier band but the minority carrier band has no connection with the bulk. Thus, the main effect minority carriers have in the equivalent circuit is to tie all $R_{n,s}C_s$ together as shown in Fig. 9(d), thereby eliminating the time constant dispersion.

In the region around mid-gap where majority and minority carrier densities at the silicon surface are comparable, no simplification is possible and the equivalent circuit of Fig. 9(c) holds. This region is only a few kT/q wide and no analysis was attempted of measurements in this region.

V. MEASUREMENT APPARATUS, REPRODUCIBILITY AND SAMPLE DRIFT

5.1 *Measurement Apparatus*

The measurement of the equivalent parallel conductance of an MIS capacitor for the purpose of extracting interface state properties requires consideration of the following points which are much more important in conductance measurements than in capacitance measurements alone.

(*i*) Harmonics of the signal frequency due to the non-linearity of the charge-voltage characteristic of the MIS capacitor can give rise to a conductance. To insure that this is negligible and the conductance measured is due to interface state transitions only, the maximum swing of surface potential caused by the applied ac signal should be less than kT/q volts.

(*ii*) The conductance measurements reported here are in the range between a nmho and about 10 μ mhos which corresponds to very small loss angles. Such small conductance values require that current leakage paths along the air-oxide interface be minimized.

(*iii*) To verify the theory developed in the previous section, capacitance, equivalent parallel conductance, and bias because of the exponential dependence in (22) must be measured very accurately.

The slices with the completed MIS capacitors were placed on a brass pedestal capped with platinum in a light-tight grounded box for measurement. A steady stream of dry nitrogen was supplied to the box. This keeps the oxide surface dry thereby preventing the formation of a current path for the ac signal along the oxide surface from the field plate to the back contact. It is especially important when making conductance measurements that such current paths be eliminated.

To avoid scratching the field plate and thereby reducing its area, a gold wire, 125 μ in diameter, mounted on a micromanipulator was used to make contact to the field plate.

For high sensitivity, it is necessary to measure accurately conductance when the loss angle is very small. The most suitable instrument for doing this over a wide range of frequencies is a capacitance bridge. The General Radio 1615-A capacitance bridge was found satisfactory over the frequency range from 50 Hz to 20 kHz and the Boonton 75-C capacitance bridge from 5 kHz to 500 kHz. Three terminal capacitance measurements were made with these bridges. Fig. 10 is a block diagram showing the arrangement.

There are three parameters which require special attention in these measurements. They are (*i*) frequency, (*ii*) signal amplitude, and (*iii*) dc bias.

5.1.1 *Frequency*

Frequency must be measured with a counter so that the precision in the calculated admittance of interface states and silicon is limited only by the precision of the bridges. The calibration marks on the oscillator dial in either the Boonton bridge or the hp signal generator are not accurate enough to obtain frequency (see Section VI).

5.1.2 *Signal Amplitude*

To keep harmonics of the signal frequency from giving rise to a spurious conductance, only signals of small amplitude can be applied. The small signal range is determined by experiment. Fig. 11 is a plot of measured capacitance and equivalent parallel conductance

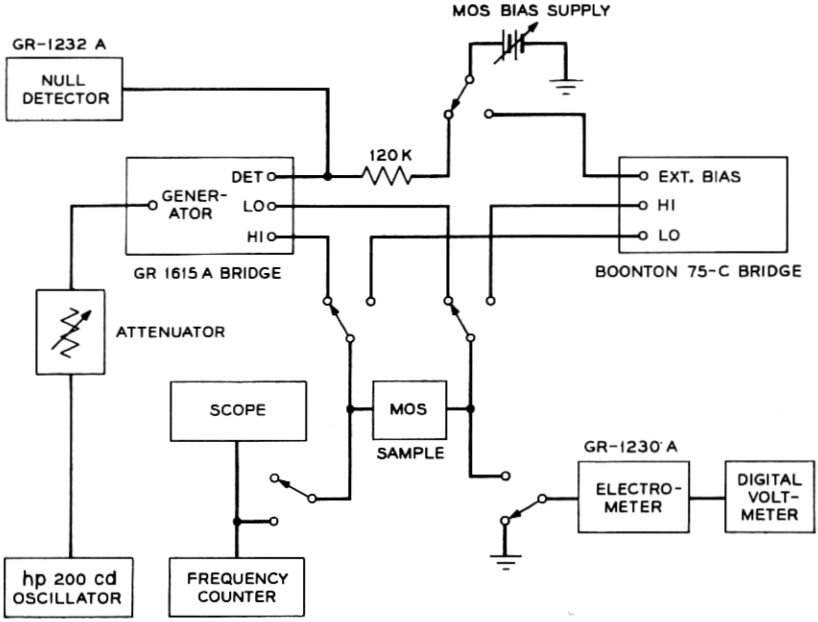

Fig. 10 — Block diagram of measuring apparatus.

normalized to their largest values as functions of peak-to-peak signal amplitude at a bias corresponding to peak conductance at 100 Hz, measured at a point where the slope of the MIS capacitance is large. There is a range of low signal amplitude at which both capacitance and equivalent parallel conductance are independent of signal amplitude. This is the experimentally determined small signal range. The maximum amplitude of the ac silicon surface potential produced by this signal is about two fifths the applied signal amplitude.

To see this, note that the applied signal divides between the oxide layer and the silicon. Expressing this mathematically, substitute (12) into (15):

$$\delta v_a = \delta\psi_s\left[1 + \frac{j\omega C_D + Y_s}{j\omega C_{ox}}\right]. \tag{45}$$

For a fixed signal amplitude, the amplitude of the silicon surface potential varies with bias because C_D and Y_s are functions of bias. Maximum silicon surface potential amplitude occurs when the Fermi level is near mid-gap because that is where C_D and $|Y_s|$ are smallest. The reason $|Y_s|$ is small is that the majority carrier density is small making the interface state time constant long. Thus, to a good approximation $|\omega(C_{ox}+C_D)| > |Y_s|$ so that (45) becomes

$$\frac{\delta\psi_s}{\delta v_a} = \frac{C_{ox}}{C_{ox} + C_D}. \tag{46}$$

Typical values are: $C_{ox} = 5.73 \times 10^{-8}$ farads/cm^2 and $C_D = 8.60 \times 10^{-8}$ farads/cm^2 which gives $\delta\psi_s/\delta v_a = 2/5$.

At the maximum sensitivity setting of the null amplifier, null is found in the small signal range at a null meter reading of $\approx 1/2$ μA on the Boonton bridge and $\approx 3/4$ μV on the General Radio bridge. These are the null readings determined essentially by the noise in the system. Therefore, harmonics of the signal frequency arising from the nonlinearity of the charge-voltage characteristics of the MIS capacitor are unimportant. However, as signal amplitude is increased beyond this range, nulls are obtained at successively higher null meter readings. This means that the amplitude of the harmonics has become large enough to get through the tuned null amplifier. In this range of signal amplitude, the conductance peak decreases in magnitude and the conductance curve is broadened.

Operating in the small signal range entails a reduction in bridge sensitivity. Therefore, optimum signal amplitude is at the high end of the small signal range. The small signal range is determined by the

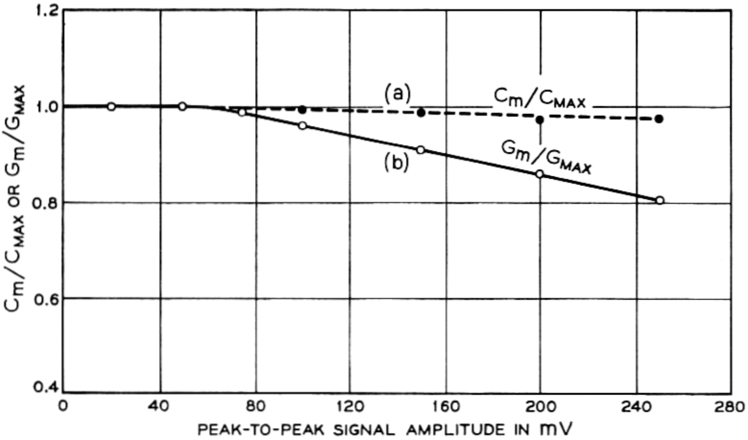

Fig. 11 — (a) Normalized capacitance and (b) normalized equivalent parallel conductance at fixed bias. The maximum value of capacitance is 367.8 pF and conductance 58.8 n mhos. Frequency is 100 Hz. Bias corresponds to peak conductance at this frequency.

oxide thickness and silicon resistivity. This range extends to higher signal amplitudes as the oxide is made thicker or silicon resistivity reduced. For oxide thicknesses between 500 Å and 700 Å and silicon resistivities around 1 ohm cm which is typical of the samples measured in this experiment, optimum peak-to-peak signal amplitude is found to be 50 mV. For other oxide thicknesses and silicon resistivities, optimum peak-to-peak signal amplitude can be found by repeating the measurements shown in Fig. 11. An external attenuator is needed on the hp signal generator to get an amplitude of 50 mV. Signal amplitude is measured with an oscillosocope and adjusted whenever frequency is changed. The maximum peak to peak silicon surface potential with this signal is $2/5 \times 50 = 20$ mV which is less than kT/q at 300°K and therefore small signal.

5.1.3 *DC Bias*

Because interface state time constants are an exponential function of silicon surface potential, dc bias must be measured to at least three decimal places. The Cimron 6900A digital voltmeter is more than adequate for this purpose.

DC bias is supplied to the MIS capacitor through a voltage divider in each bridge. Therefore, it must be measured across the sample rather than at the supply terminals. The voltmeter puts an additional capacitance and conductance across the sample which are read by the bridge. To avoid the correction this entails, the voltmeter is switched out of the circuit before the bridge is balanced. Because the input resistance of the voltmeter no longer shunts the MIS capacitor, the current through the voltage divider in the bridge decreases. Therefore, the voltage across the MIS capacitor is slightly larger than that measured. For the Cimron 6900A digital voltmeter which has an input impedance when balanced of 10^9 ohms, this is not a problem. However, for a digital voltmeter having an input impedance of a few megohms, this correction becomes apparent whenever the resistance values of the internal bridge voltage divider are changed. These changes occur when switching the conductance multiplier on the Boonton bridge, the conductance to loss factors switch on the General Radio bridge or in switching from one bridge to another. To eliminate this error, a General Radio 1230A electrometer which has an input resistance of 10^{15} ohms with the resistance selector set at ∞ can be used to measure bias voltage across the sample. The output of the electrometer is then read by the digital voltmeter.

To permit control of the bias voltage to within a fraction of a millivolt, a voltage divider consisting of a bank of fixed resistors in series with a 10-turn helipot is used in conjunction with a heavy duty 45-volt battery in a shielded box.

5.2 *Reproducibility of Data and Sample Drift*

Sample preparation is designed to eliminate drift under negative bias at room temperature only. The admittance-voltage characteristics in the range of interest for both n- and p-type samples are located in the negative bias range because of the presence of a fixed residual positive charge in the oxide. Therefore, n-type samples are in strong accumulation and p-type in strong inversion at positive bias. No conductance measurements are made at positive bias because conductance due to interface states is too low to detect in strong accumulation or inversion.

Reproducibility of the measurements on a given MIS capacitor was assured as follows. Measurements are repeated after room temperature bias aging at high negative bias, after the lapse of several days, and by hysteresis measurements where the admittance-voltage characteristics are measured first with increasing bias and then with decreasing bias. Only those samples where the measurements were found to be reproducible within experimental error are used. Interface state properties for such capacitors on a few different slices are found to be similar.

VI. TECHNIQUE FOR ANALYZING THE DATA

6.1 *Properties of Equivalent Parallel Conductance*

The variation of G_p/ω with bias or frequency can be most easily seen from (17) and (22) for a single time constant. In (39) derived from the statistical model which fits the measurements exactly in the depletion-region, (see Fig. 21) the variation of G_p/ω with bias and frequency is obscured by the mathematical complexity. Equation (17) is adequate for illustrative purposes because of its simplicity and the fact that its variation with bias and frequency is qualitatively the same as (39) in the depletion region. Equation (17) fits the data exactly in the weak inversion region (see Fig. 25). G_p or G_p/ω can be obtained from measurements as a function of bias with frequency as parameter or as a function of frequency with bias as parameter. G_p or G_p/ω will go through a peak as a function of bias at each frequency as seen from (17) and (22). For a continuum of states, this peak occurs when the response time of the interface states which is being varied by the bias equals the period of the applied signal frequency multiplied by a constant. When the frequency is increased, this peak shifts to a bias value corresponding to a surface potential nearer to flat bands. Fig. 1 shows capacitance and equivalent parallel conductance measured at 5 kHz and 100 kHz in a wet oxide. This figure illustrates how measured equivalent parallel conductance varies with bias and frequency. Fig. 1 also shows how inaccurate the capacitance method for extracting information about interface states is compared to the conductance method. It can be noted from Fig. 1 that

(*i*) Maximum capacitance dispersion in this frequency range is about 14 percent while equivalent parallel conductance changes more than an order of magnitude.

(*ii*) The conductance peak moves towards a bias voltage closer to flat band bias which in Fig. 1 is −2.6 volts as frequency is increased. Between 50 Hz and 500 kHz, conductance does not peak at flat band bias in the samples measured.* The peak shifts over the bias range from weak inversion to depletion in this frequency range.

(*iii*) Capacitance and equivalent parallel conductance are related to each other by the Kronig-Kramers relations. Rather than going through a mathematical proof of this, we note that (16), (17), (39), and (40) have the same form as the Debye equations.[19, 20] Because the Debye equations satisfy the Kronig-Kramers relations,[27] we infer that (16), (17), (39), and (40) do also.

G_p/ω as a function of frequency will go through a peak at each value of bias but in this case G_p will not. It can be seen from (17) that only G_p/ω is a symmetric in $\omega\tau$ when frequency is the variable while both G_p and G_p/ω are symmetric in $\omega\tau$ when bias is the variable.

G_p/ω vs *bias* (with frequency as parameter) will be spread over a bias range determined by the time constant dispersion and the density of interface states. The density of states spreads the G_p/ω vs bias curve in the same way that it spreads the high-frequency capacitance vs bias curve. However, G_p/ω vs ω (with bias as parameter) will be spread over a frequency range determined only by the time constanct dispersion. Because time constant dispersion determines the interface state admittance, G_p/ω vs ω curves are the more useful of the two for the purposes of this work.

6.2 *Extraction of G_p/ω*

It is necessary first to get G_p/ω and C_p from the measurements. C_p and G_p are shown in Fig. 8(b) to be the capacitance and equivalent

---

* To get the conductance to peak at flat band bias, frequencies higher than 500 kHz have to be applied.

parallel conductance of the portion of the equivalent circuit consisting of C_D in parallel with the distributed network representing the interface states. The admittance measured across the terminals x–x in Fig. 8(a) is $G_m + j\omega C_m$. Converting this to an impedance, subtracting away the reactance of C_{ox}, and converting back to an admittance, we get

$$\frac{G_p}{\omega} = \frac{\omega C_{ox}^2 G_m (G_m^2 + \omega^2 C_m^2)}{\omega^2 C_{ox}^2 G_m^2 + [\omega^2 C_m (C_{ox} - C_m) - G_m^2]^2} \tag{47}^*$$

and:

$$C_p = \frac{C_{ox}(G_m^2 + \omega^2 C_m^2)[\omega^2 C_m (C_{ox} - C_m) - G_m^2]}{\omega^2 C_{ox}^2 G_m^2 + [\omega^2 C_m (C_{ox} - C_m) - G_m^2]^2}. \tag{48}^*$$

The measured G_p/ω from (47) is related to N_{ss} and τ_m by (39) and C_p from (48) is related to C_D, N_{ss} and τ_m by (40) in the depletion region.

It should be noted that the circuit across the terminals x–x in Fig. 8(a) has a shorter time constant than that of just the interface state branch of the circuit because of C_{ox}. Therefore, G_p/ω from (47) will peak at a lower frequency than G_p/ω measured across terminals x–x when equivalent parallel conductance is measured as a function of frequency with bias as parameter. If equivalent parallel conductance is measured as a function of bias with frequency as parameter, G_p/ω from (47) will peak at a bias close to flat bands for the same reason as can be seen from (22).

To minimize errors introduced by correcting for C_{ox}, the oxide should be made as thin as practicable and frequency measured accurately with a counter as described in Section V.

6.3 *Bias vs Surface Potential*

From (22), there is an exponential relation between interface state time constant and surface potential. Therefore, bias is measured with a digital voltmeter as described in Section V and the relation between bias and surface potential is determined as accurately as possible. It is necessary to do this to prove that (22) fits the measurements and to extract c_n and c_p accurately. One of the most accurate methods for obtaining the relation between bias and surface potential has been developed by C. N. Berglund.[28] In this method, capacitance is measured as a function of bias at such a low frequency that the interface

---

* These equations were evaluated using the measured values of ω, C_{ox}, C_m, and G_m on an IBM 1620 computer.

states are in equilibrium with the silicon. This means that the loss associated with the capture and emission of carriers by the interface states is negligible.

Bias vs surface potential is obtained in two steps from the measured data. First, surface potential is found to within an additive constant at each bias. Then, the additive constant is found.

6.3.1 *Surface Potential within an Additive Constant*

Low-frequency capacitance C_{om} measured at bias v_o is

$$C_{om} = dQ_T/dv_o \,. \tag{49}$$

Differentiating (13) with respect to v_o

$$1 = d\psi_s/dv_o + (1/C_{ox})\, dQ_T/dv_o \,. \tag{50}$$

From (49) and (50)

$$d\psi_s/dv_o = 1 - C_{om}/C_{ox} \,. \tag{51}$$

The surface potential at any bias v_{o1} is found by integrating (51)

$$\psi_s(v_{o1}) = \int_{v_{o2}}^{v_{o1}} (1 - C_{om}/C_{ox})\, dv_o + \Delta, \tag{52}$$

Where Δ is an additive constant and v_{o2} is a bias in strong accumulation.

Thus, surface potential at each bias can be obtained to within an additive constant directly from measured data by an integration. The fact that an integration is carried out using directly measured quantities makes this method accurate.

It is found that in samples oxidized in steam as described in Section II, 50 Hz is a low enough frequency to measure C_{om} vs v_o and use (52) with negligible error. This is seen by the observation that the loss tangent is negligibly small in the bias range investigated here below about 80 Hz so that dispersion in the capacitance-voltage characteristic virtually disappears in this frequency range.

Another important consideration is that (52) will not give an accurate result if there are any *gross* nonuniformities. Gross nonuniformities occur when there are macroscopic areas under the field plate in which the charge density is significantly different from adjacent areas. Whether or not there are such nonuniformities present in a sample can be determined by evaluating (52) using C_{om} measured from strong accumulation to strong inversion. The resulting value of

ψ_s will be slightly less than the bandgap,* if there are no gross nonuniformities and considerably higher if there are.

The random distribution of built-in charges and charged interface states does not affect the accuracy of (52). Because of the statistical fluctuations of surface potential caused by this random charge distribution, bias is really found as a function of mean surface potential.

6.3.2 *Determination of Additive Constant*

The additive constant Δ in (52) and the doping density is calculated next. For this purpose, C_D must be found as a function of bias in the depletion region. Then C_D^{-2} is plotted as a function of $\psi_{se} - kT/q$ where

$$\psi_{se} = \int_{v_{o2}}^{v_{o1}} (1 - C_{om}/C_{ox})\, dv_a \ .$$

G_p/ω from (47) peaks at bias values within the depletion region at frequencies above a few hundred Hz. C_D is found from admittance measured at these frequencies using (48) to get first C_p . Then from (40)

$$C_p = C_D + C_s(\omega, \bar{u}_s), \tag{53}$$

where $C_s(\omega, \bar{u}_s)$ is the equivalent capacitance of the distributed network representing the interface states in Fig. 8(a) at the frequency and bias corresponding to $(G_p/\omega)_{\max}$ of the measured curve obtained using (47). It is shown in Appendix C that $C_s(\omega, \bar{u}_s) = 2(G_p/\omega)_{\max}$ where $(G_p/\omega)_{\max}$ is the measured peak. Then, using (53) we get

$$C_D = C_p - 2(G_p/\omega)_{\max} \ . \tag{54}$$

In the depletion region except near flat bands, C_D can be written in the following form using the approximate relation between C_D and $\bar{u}_s$ [16]

$$C_D^{-2} = 2(\bar{u}_s - 1)/q\epsilon_{si}N_A\beta, \tag{55}$$

where ϵ_{si} is the dielectric permittivity of silicon. Thus, a plot of C_D^{-2} measured far from flat bands in the depletion region vs $\bar{\psi}_{se} - kT/q$ measured in the same region will yield a straight line as shown in Fig. 12. The intercept of this line with the abcissa when it is extrapolated to $C_D^{-2} = 0$ gives Δ. From this, bias is found as a function of surface potential. Finally, the slope of the line in Fig. 12 gives the doping density from (55).

---

* This is particularly the case for samples having high doping density such as those used in this work.

6.4 N_{ss} *and* τ_m *in Depletion*

Equation (39) for G_p/ω from the statistical model is found to fit the measured G_p/ω vs frequency curves accurately at each bias (see Section VII). To prove this, (39) is numerically integrated on an IBM 7094 computer using the trapezoidal rule.* $\bar{u}_s$ corresponding to the particular bias at which G_p/ω vs frequency is measured is obtained from the relation between bias and surface potential just found. The characteristic area α is the only parameter which is varied to obtain the best fit to the measured curve. N_{ss} can be obtained independently.

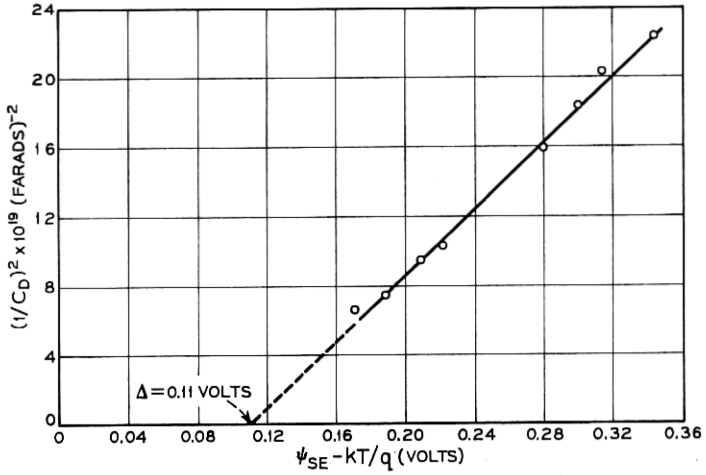

Fig. 12—$1/C_D{}^2$ vs $\psi_{se}-kT/q$ where C_D in farads is obtained from measurements in the depletion region of an n-type sample ($N_D = 1.15 \times 10^{15}$ cm^{-3}) and ψ_{se} from the integral in (52). Field plate diameter is 3.8×10^{-2} cm.

The parameter α is not arbitrary. It is found that $\alpha^{\frac{1}{2}}$ has about twice the value and the same bias dependence as the silicon depletion layer width. (see Fig. 22).

Fig. 13(a) shows G_p/ω found by integrating (39) plotted as a function of $\omega\tau_m$ with N_{ss} arbitrarily chosen as 1×10^{11} cm^{-2} eV^{-1} and α at a bias midway between flat bands and mid-gap. For comparison, Fig. 13(b) is a plot of (17) and Fig. 13(c) is a plot of (20) vs $\omega\tau_m$ for the same value of N_{ss}. From Fig. 13(a), it is seen that the maximum value of G_p/ω from (39) occurs when $\omega\tau_m = 2.5$. This was found to hold

---

* To do this, limits of integration in (39) must be finite. Because the fluctuations of surface charge are small, integrating from $u_s = -2u_B$ to $u_s = 2u_B$ will give a good approximation to the integral from $-\infty$ to ∞.

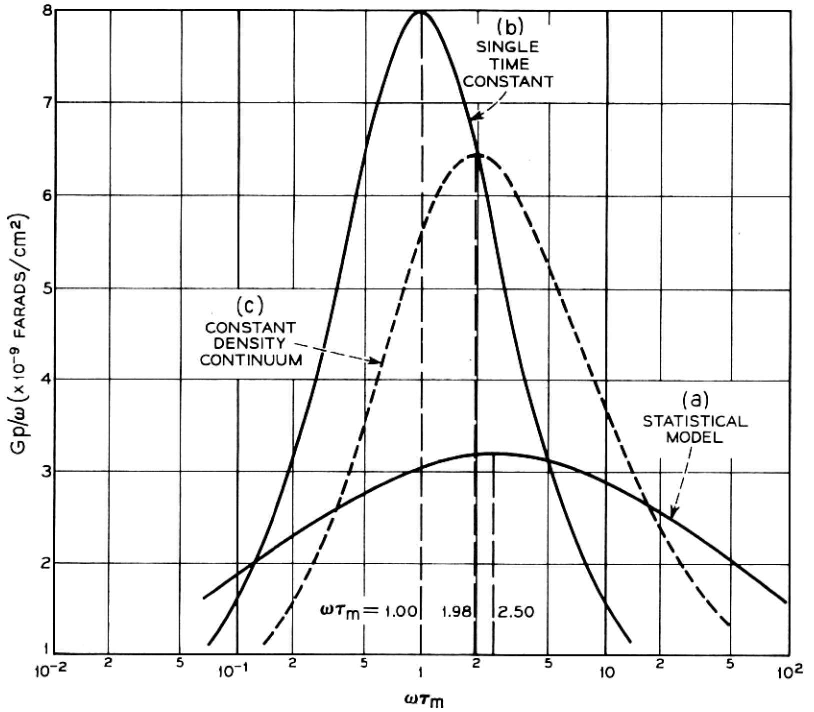

Fig. 13 — Calculated G_p/ω vs log $\omega\tau_m$. (a) Plot of (39) incorporating time constant broadening from statistical fluctuations of surface potential as well as continuum of states. $\alpha \doteq 2.5 \times 10^{-10}$ cm² and impurity density $= 2.1 \times 10^{16}$ cm^{-3} are used. (b) Plot of (20) calculated by Lehovec in Ref. (19) for a continuum of constant density and capture cross section. The same interface state density arbitrarily chosen as 1×10^{11} cm^{-2}-eV^{-1} is used in calculating all the curves.

at all values of bias measured. The condition $\omega\tau_m = 2.5$ is used to find τ_m from the frequency at which the measured G_p/ω vs frequency curve goes through a maximum at each bias. The τ_m found this way corresponds to the mean surface potential $\bar{u}_s$ at each bias. The condition $\omega\tau_m = 2.5$ can be used to get τ_m when G_p/ω is measured as a function of bias with frequency as parameter. In this case, ω is given by the fixed frequency but the τ_m obtained now corresponds to the bias at peak G_p/ω.

Interface state density vs bias is found by numerically integrating (39) with $\omega\tau_m = 2.5$, equating the result to the measured peak G_p/ω value, and solving for N_{ss} at each bias. The condition $\omega\tau_m = 2.5$ is equivalent to setting $u_0 = u_B - \bar{u}_s + \ln 2.5$ in (39). Doing this, (39) becomes

$$\left(\frac{G_p}{\omega}\right)_{\max} = \tfrac{1}{2} q N_{ss} [2\pi(\sigma_s^2 + \sigma_B^2)]^{-\frac{1}{2}} I_{\max} \tag{56}$$

and

$$I_{\max} = \int_{-2u_B}^{2u_B} \exp\,[-(z + y_m)]\ \ln\,[1 + \exp\,(2y_m)]\ du_s\ , \tag{57}$$

where

$$y_m = u_s - \bar{u}_s + \ln 2.5 \quad \text{and} \quad z = (u_s - \bar{u}_s)^2 / 2(\sigma_s^2 + \sigma_B^2).$$

Equation (56) and (57) can be evaluated once α is known. The simplest way to find an approximate value for α is to use the experimentally established relation between $\alpha^{\frac{1}{2}}$ and W where W is the depletion layer width. W can be obtained from the relation

$$W = \frac{\epsilon_{si}}{C_D}\ , \tag{58}$$

where C_D is obtained from (54).

Solving (56) for N_{ss}

$$N_{ss} = 2\left(\frac{G_p}{\omega}\right)_{\max} [2\pi(\sigma_s^2 + \sigma_B^2)]^{\frac{1}{2}} (I_{\max})^{-1} q^{-1}. \tag{59}$$

Substituting (57) into (59) and using values of α from Fig. 22, (59) was evaluated to find N_{ss} as a function of the position of the Fermi level in the bandgap of the silicon.

6.5 *N_{ss} and τ_m in Weak Inversion*

This is a particularly simple case because G_p/ω vs frequency is characterized by a single time constant at each bias. G_p/ω vs frequency or bias curves are obtained using (47) as before. Fig. 6 and (17) apply so that the condition for maximum G_p/ω is $\omega\tau_m = 1$ and the value of $(G_p/\omega)_{\max}$ is $C_s/2$.

6.6 *Capture Probabilities and Cross Sections*

Log τ_m is plotted as a function of $u_B - \bar{u}_s$ from measurements in the depletion region to see if the data fit (22). It is found this way that the measurements fit (22) quite well for both holes and electrons (see Section VII). The fact that the experimental data fit (22) means that c_n and c_p are independent of bandgap energy. Therefore, c_n and c_p are not affected by statistical fluctuations of surface potential and are the true capture probabilities of the interface states. Capture cross section is related to capture probability by the relation $c_n = \bar{v}\sigma_n$ for electrons where $\bar{v}$ is the average thermal velocity of electrons. c_n and c_p are calculated from experimental plots of (22) and σ_n and σ_p found using $v = 10^7$ cm/sec for both holes and electrons.

VII. EXPERIMENTAL RESULTS

7.1 *Depletion Region*

7.1.1 *Steam-Grown Oxides on [111] Orientation*

Fig. 14 shows log τ_m vs $u_B - \bar{u}_s$ calculated from the measurements as explained in Section VI. The values plotted correspond to peaks

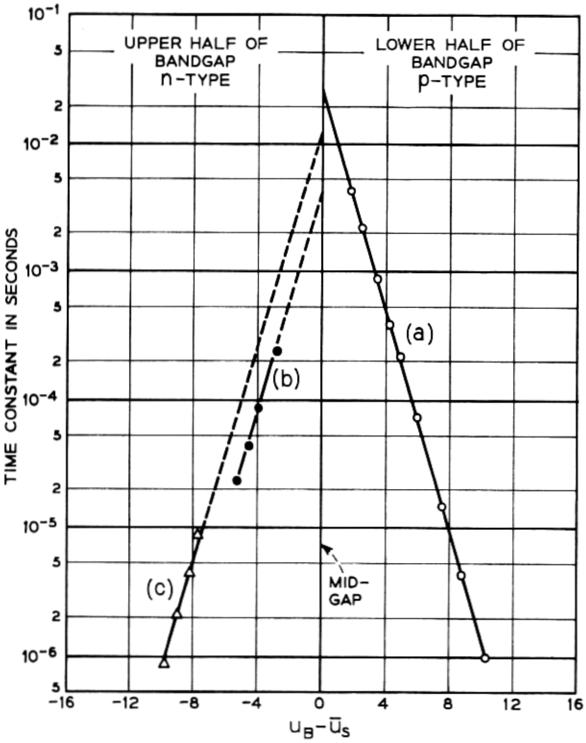

Fig. 14 — Log τ_m vs $u_B - \bar{u}_s$—showing fit of experimentally obtained points to (22) for both holes and electrons using steam grown oxide and [111] orientation. Points in lower half of the gap were measured on a p-type sample. Field plate diameter is 1.2×10^{-1} cm, acceptor density 2.08×10^{16} cm^{-3}, $u_B = 14.1$ and $C_{ox} = 5.74 \times 10^{-8}$ farads/cm^2. Points in upper half of the gap were measured on an n-type sample. Field plate diameter 3.8×10^{-2} cm, donor density 1.15×10^{16} cm^{-3}, $u_B = 13.5$, and $C_{ox} = 5.08 \times 10^{-8}$farads/cm^2.

of the G_p/ω curves obtained from measurements made at 300°K. The condition $\omega\tau_m = 2.5$ is used to get τ_m. Because only majority carrier transitions give rise to measured conductance, curves in the lower half of the silicon bandgap were obtained from measurements on p-type and in the upper half on n-type. It is seen that these curves fit (22) so that capture cross sections are obtained simply by extrapolating the curves to $u_B - \bar{u}_s = 0$ where (22) becomes $\tau_{mo} = (\bar{v}n_i\sigma)^{-1}$. This is then solved for σ to get capture cross section. Capture cross sections obtained this way from Fig. 14 are

σ_p (holes):	2.2×10^{-16} cm^2	from curve (a)
σ_n (electrons):	1.7×10^{-15} cm^2	from curve (b)
σ_n (electrons):	5.9×10^{-16} cm^2	from curve (c).

In all cases, $\bar{v} = 10^7$ cm/sec and $n_i = 1.55 \times 10^{10}$ cm^{-3} have been used.

Fig. 15 shows the corresponding values of N_{ss} as a function of the average position of the Fermi level E_F with respect to mid-gap potential at the silicon surface. Measurements in the lower half of the gap were made on p-type and in the upper half on n-type.

Fig. 16 illustrates the temperature dependence of τ. Curve (a) is calculated from measurements made at 204°K while curve (b) is just a replot of curve (a) from Fig. 14 for comparison.

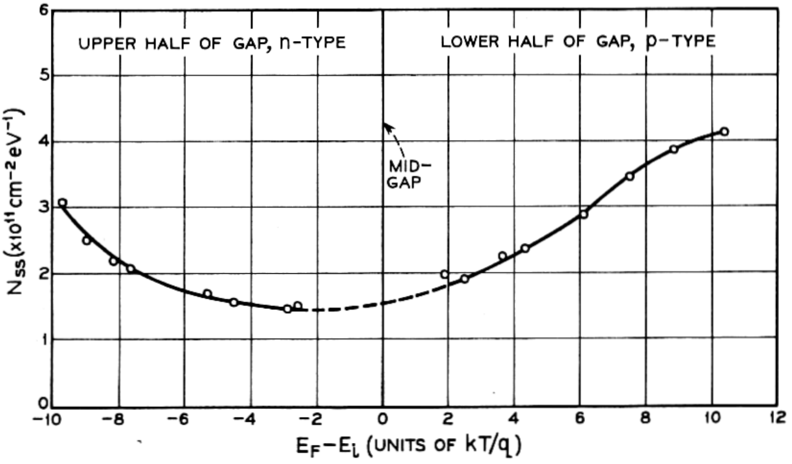

Fig. 15 — N_{ss} from (59) vs $E_F - E_i$ measured in units of kT/q for same two samples in Fig. 14. Flat band is 14.1 in lower half of the gap and 13.5 in upper half. E_F is position of the Fermi level and E_i is mid-gap.

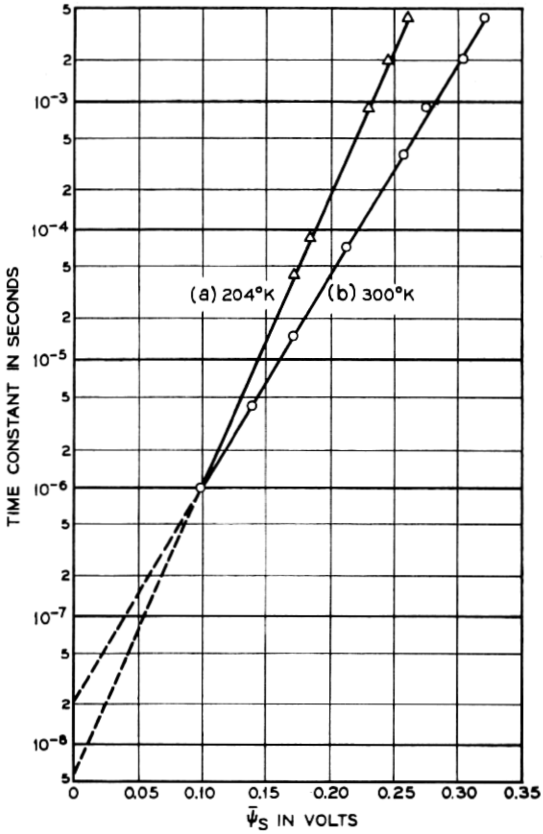

Fig. 16—Log τ_m vs $\bar{\psi}_s$ for p-type sample of Fig. 14: (a) measured at 204°K and (b) at 300°K. (b) Replot of (a) in Fig. 14. Curves show that experimentally obtained points satisfy temperature dependence of (22). (a) is the only one in this experiment measured at a temperature other than 300°K.

7.1.2 *Steam-Grown Oxides on [100] Orientation*

Fig. 17 shows log τ_m vs $u_B - \bar{u}_s$ for the [100] orientation calculated from measurements made at 300°K. Capture cross sections are

σ_p (holes):	2.0×10^{-16} cm^2	from curve (a)
σ_p (holes):	4.0×10^{-16} cm^2	from curve (b)
σ_n (electrons):	1.2×10^{-15} cm^2	from curve (c).

Fig. 18 shows the corresponding values of N_{ss} in a plot similar to Fig. 15.

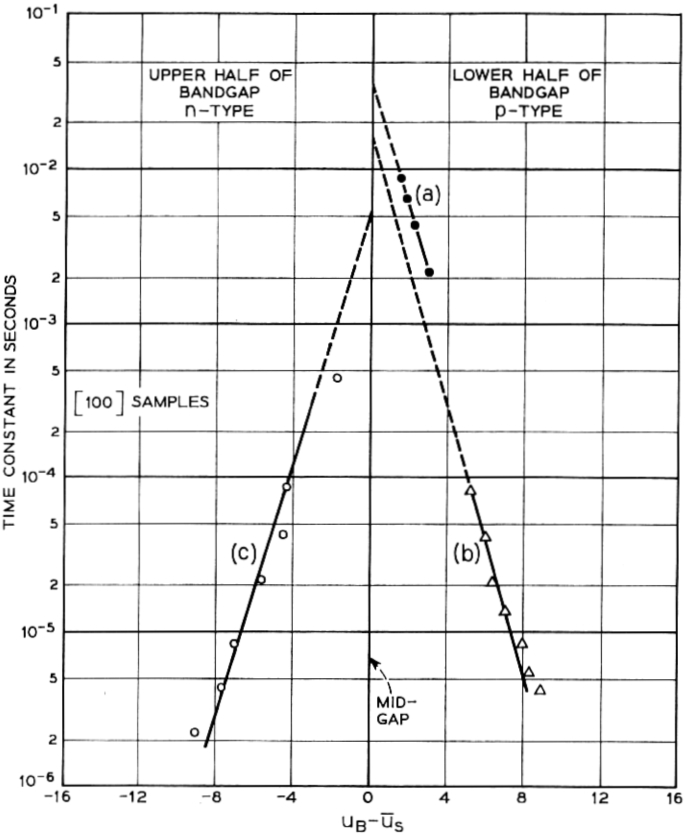

Fig. 17 — Log τ_m vs $u_B - \bar{u}_s$ showing fit of experimentally obtained points to (22) for both holes and electrons using steam-grown oxide and [100] orientation. Points in lower half of gap were measured on a p-type sample. Field plate diameter is 1.37×10^{-1} cm, acceptor density 1.19×10^{16} cm^{-3}, $u_B = 13.5$, and $C_{ox} = 4.94 \times 10^{-8}$farads/cm^2. Points in upper half of gap were measured on an n-type sample. Field plate diameter is 1.43×10^{-1} cm, donor density 1.44×10^{16} cm^{-3}, $u_B = 13.7$, and $C_{ox} = 4.24 \times 10^{-8}$farads/cm^2.

7.1.3 *[111] Orientation, Partly-Dried Oxides*

Fig. 19 shows log τ_m vs $u_B - \bar{u}_s$ for oxides initially grown in steam and partially dried by heating in dry N_2 as described in Section III. Only n-type was analyzed in detail, so that we have

$$\sigma_n \text{ (electrons): } \quad 4.6 \times 10^{-16} \text{ cm}^2 \quad \text{from Fig. 19.}$$

Curve (a) in Fig. 20 shows the corresponding values of N_{ss}. Curve (b) in Fig. 20 is a replot from Fig. 15 for comparison.

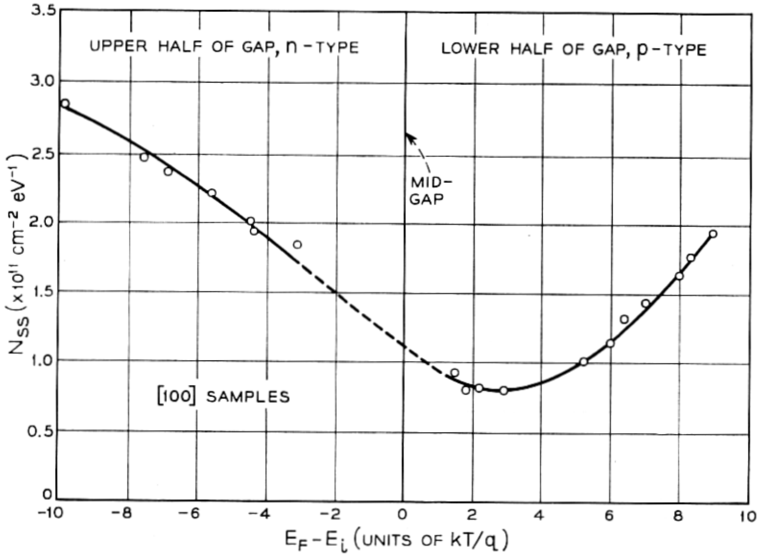

Fig. 18 — N_{ss} from (59) vs $E_F - E_i$ for same two samples in Fig. 17. Flat band is 13.5 in lower half of gap and 13.7 in upper half.

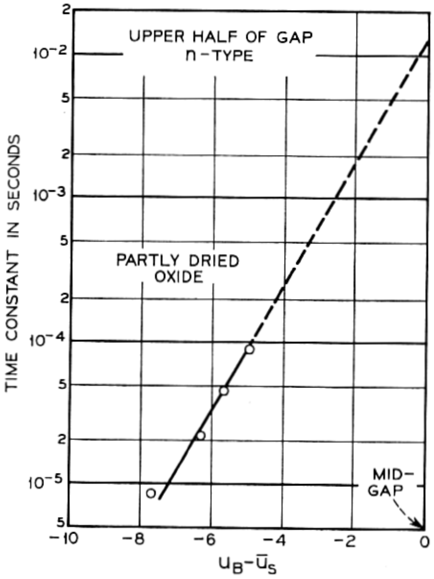

Fig. 19 — Log τ_m vs $u_B - \bar{u}_s$ after partial drying of n-type sample in Fig. 4. Measured points still fit (22). Donor density is 1.4×10^{16} cm^{-3}, $u_B = 13.7$, and $C_{ox} = 5.96 \times 10^{-8}$farads/cm^2.

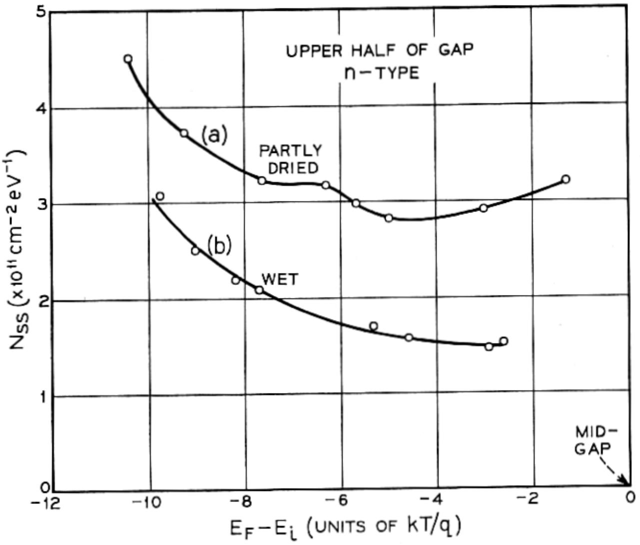

Fig. 20 — N_{ss} from (59) vs $E_F - E_i$. (a) n-type sample in Fig. 19 and (b) replot of n-type sample from Fig. 15 for comparison. For (a), flat band is 13.7 and for (b), flat band is 13.5.

7.1.4 *Range of N_{ss}*

As seen from Figs. 15 and 18, N_{ss} in steam-grown oxides is in the 10^{11} cm^{-2}-eV^{-1} range. Partly drying increases N_{ss} as seen from Fig. 20. Complete drying or growing the oxide in dry oxygen at 1000°C results in N_{ss} values in the 10^{12} cm^{-2}-eV^{-1} range or about an order of magnitude higher than for steam-grown oxides.

N_{ss} can be reduced below the values for steam-grown oxides by an order of magnitude by a method described by Balk.[29] First, an aluminum film serveral thousand Å thick is evaporated over the entire surface of a steam-grown oxide. The sample is then annealed in H_2 or N_2 at 350°C for 1/2 hour. After annealing, field plates are produced using photoresist techniques and chemically etching off the unwanted aluminum. This annealing process is found to reduce N_{ss} from the 10^{11} cm^{-2}-eV^{-1} range characteristic of steam-grown oxides to the 10^{10} cm^{-2}-eV^{-1} range. Thus, N_{ss} can be varied two orders of magnitude from the 10^{12} cm^{-2}-eV^{-1} range characteristic of dry oxides to the 10^{10} cm^{-2}-eV^{-1} range. Capture cross section appears to remain approximately the same throughout.

7.1.5 *Fit to Statistical Model*

The circles in Fig. 21 represent points calculated from measurements on the [111] p-type sample with a steam-grown oxide in Figs. 14 and 15. These points were measured with frequency as variable and bias fixed at −1.9 volts which corresponds to $u_B - \bar{u}_s = 6.04$. The best fit to these points calculated from (39) is the solid curve in Fig. 21. This fit was obtained by substituting (59) for N_{ss} in (39) and using $(G_p/\omega)_{\max}$ from the data. The value of α required for this fit was found to be 2.5×10^{-10} cm².

Fig. 22 shows $\alpha^{\frac{1}{2}}$ vs space-charge width W. The values of α are obtained by finding the best fit of (39) to G_p/ω vs frequency curves taken at different bias settings on the same sample. Space-charge width is calculated from the depletion layer capacitance obtained as described in Section VI from the measured capacitance and equivalent parallel conductance.

7.2 *Weak Inversion*

The G_p/ω curves obtained after correcting for C_{ox} will peak in the weak inversion range at signal frequencies below about 1 kHz. Curves

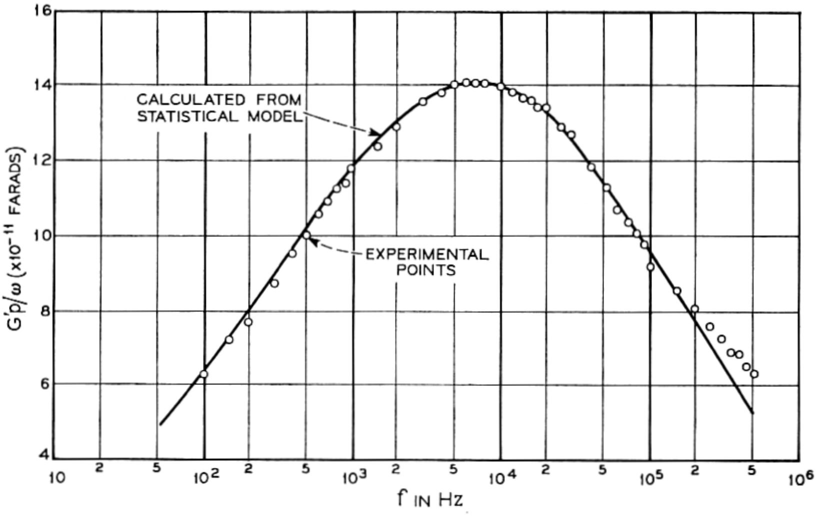

Fig. 21 — G_p'/ω vs frequency showing how experimental points fit (39). An α of 2.5×10^{-10} cm² is used in (39) to get this fit. G_p' is equivalent parallel conductance in mhos derived from the measurements. Experimental points are measured at a bias of −1.9 volts ($\bar{u}_s = 8.1$) in depletion region on the p-type sample of Fig. 14.

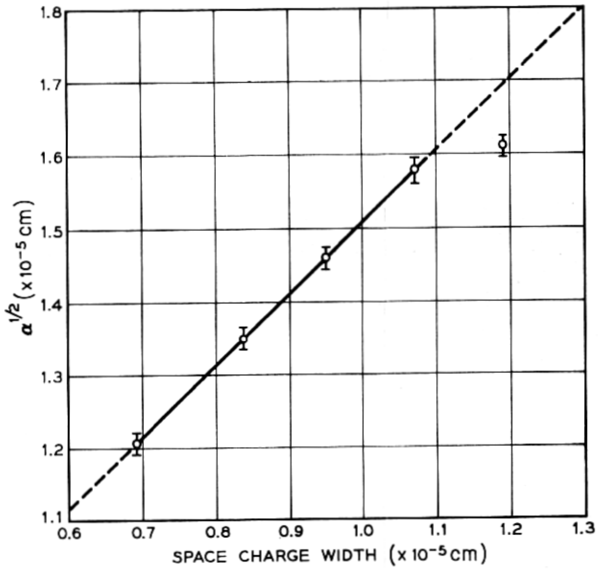

Fig. 22 — $\alpha^{\frac{1}{2}}$ vs space-charge width where α is obtained by finding the best fit of (39) to experimental points taken on p-type sample in Fig. 14. $\alpha^{\frac{1}{2}}$ can have any value within vertical bars through each point. Bars indicate precision with which (39) can be fitted to measured G_p/ω vs frequency points. No reason why the point at the largest W does not lie on the straight line is known.

of capacitance and equivalent parallel conductance vs field plate bias measured at 50 Hz on a [111], n-type sample with steam grown oxide are shown in Fig. 23. An n-type sample was used to avoid the masking effects of lateral ac current flow.[10] The capacitance peak below mid-gap persists down to 1.6 Hz, the lowest frequency used, and disappears at frequencies above 500 Hz. Above 500 Hz, the capacitance curve becomes the usual high-frequency curve.[5] The low-frequency capacitance dispersion is shown in Fig. 24.

G_p/ω vs frequency curves in this range can be fitted with a single time constant as shown in Fig. 25. The circles represent points calculated from measurements on an n-type sample with steam-grown oxide on the [111] orientation. These points were measured with frequency as variable and bias fixed at -3.3 volts which corresponds to $u_B - \bar{u}_s = -6.3$ in weak inversion. The best fit to these points calculated from (17) is the solid curve in Fig. 25. The values of τ and C_s used in (17) to obtain this fit are 1.7×10^{-3} seconds and 5.14×10^{-8} farads/cm^2, respectively.

Fig. 26 is a plot of log τ_m vs $u_B - \bar{u}_s$. These values are calculated from measurements in the weak inversion region using $\omega\tau_m = 1$ to get τ_m. Fig. 27 is a plot of the corresponding values of C_s as a function of the position of the Fermi level with respect to mid-gap potential at the silicon surface. Values of C_s are obtained from the peak of the G_p/ω curve using $\omega\tau_m = 1$ and (17). Values of N_{ss} using $N_{ss} = C_s/q$ are given on the right-hand ordinate axis of Fig. 27. Although these values are not exact, they give the correct order of magnitude of N_{ss}.

Fig. 28 shows a low-temperature drying experiment similar to those in Figs. 3 and 4 except the curves in Fig. 28 were measured at 50 Hz. Curves (a) and (b) of Fig. 28 are plots of capacitance and equivalent parallel conductance vs bias for a steam-grown oxide on n-type silicon. Curves (c) and (d) are the capacitance and equivalent parallel conductance vs bias after heating in dry nitrogen at 350°C for 17 hours.

Fig. 29 is a plot of capacitance and equivalent parallel conductance vs bias measured at 50 Hz on a sample having N_{ss} in the 10^{10} $\text{cm}^{-2}\text{-eV}^{-1}$ range. The sample is n-type oriented in the [100] direction and the low value of N_{ss} is produced by the annealing process described previously.

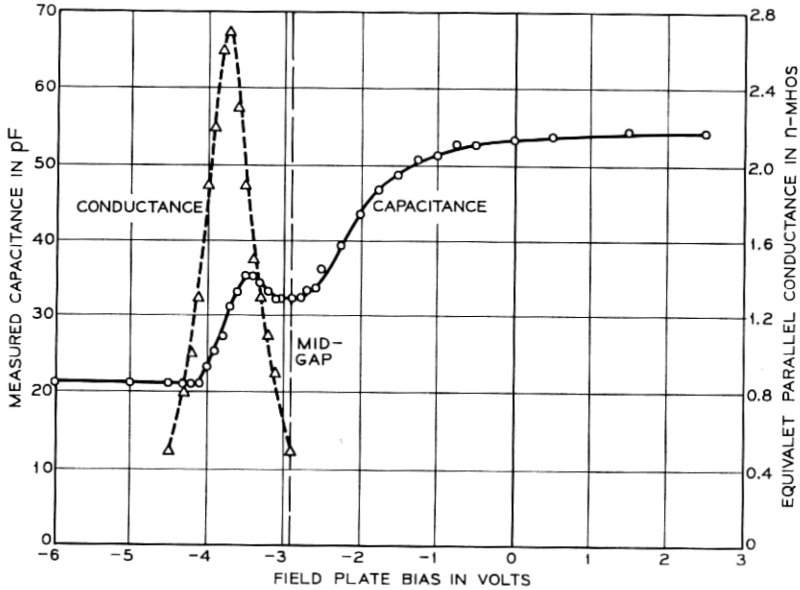

Fig. 23 — Capacitance and equivalent parallel conductance measured at 50 Hz vs field plate bias. Sample is the n-type one in Fig. 14. curves show capacitance and equivalent parallel conductance in the weak inversion region.

VIII. DISCUSSION

The discussion is divided into three parts. The first part deals with the depletion region and the second with the weak inversion region. The third part discusses the limitations of the conductance technique.

8.1 *Depletion Region*

8.1.1 *Steam-Grown Oxides on [111] Orientation*

It is seen from Fig. 14 that points calculated from the measurements fit (22) quite well for both n and p type samples. Time constant decreases by a factor of $1/e$ for every kT/q increase in $u_B - \bar{u}_s$. The exponential dependence of time constant on surface potential shown in Fig. 14 means that time constant is inversely proportional to majority carrier density at the silicon surface as seen from (21) and (22). This

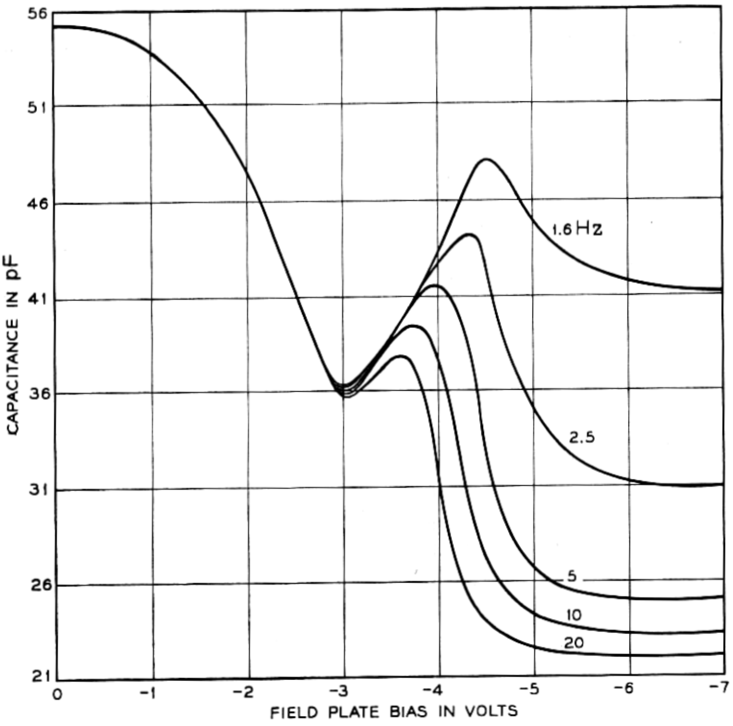

Fig. 24 — Capacitance vs field plate bias with frequency as parameter measured with apparatus described in Ref. 28, on n-type sample of Fig. 14. Curves show capacitance dispersion in weak inversion.

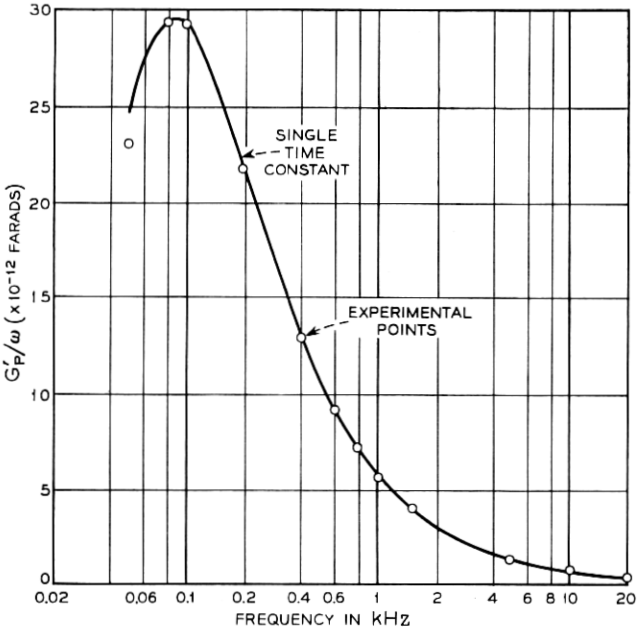

Fig. 25 — G_p'/ω vs frequency showing how experimental points fit (17). $C_s = 5.14 \times 10^{-8}$ farads/cm² and $\tau = 1.7 \times 10^{-3}$ seconds are used in (17) to get this fit. G_p' is equivalent parallel conductance in mhos derived from measurements. Experimental points are measured at a bias of -3.3 volts ($\bar{u}_s = 19.8$) in the weak inversion region on the n-type sample of Fig. 14.

is the experimental evidence that majority carrier transitions make the dominant contribution to measured equivalent parallel conductance in the depletion region.

The exponential dependence of time constant on surface potential is important evidence in addition to the drying experiments of Figs. 3 and 4, that the dominant loss is due to transitions to and from interface states. This follows from the fact that no other relaxation processes in the system have time constants which depend exponentially on surface potential. To see this, we shall consider generation-recombination though impurity levels in the silicon near mid-gap, majority carrier transitions to and from impurity levels in the silicon near the Fermi level and the transit time of majority carriers across the silicon space-charge region. Bulk generation-recombination is insignificant because the Fermi level does not cross impurity levels near mid-gap for band bending in depletion as seen in Fig. 5(b). Therefore, generation-recombination is taken to be zero in Figs. 6 and 8. The Fermi level can

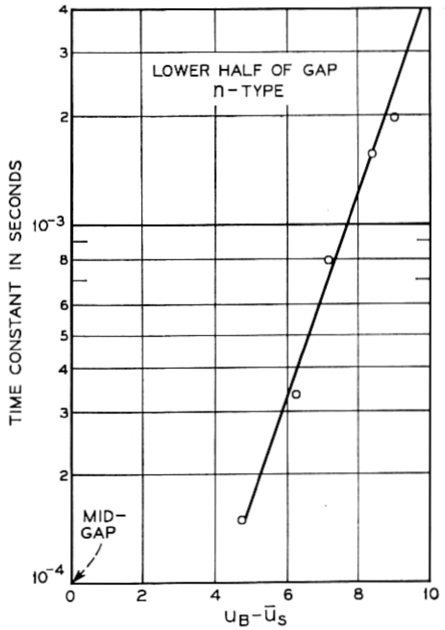

Fig. 26 — Log τ_m vs $u_B - \bar{u}_s$ for n-type sample in Fig. 14 in weak inversion region. Curve shows that equivalent parallel conductance in weak inversion is dependent on majority carrier density at the silicon surface.

still cross any impurity levels located near it [see Fig. 5(b)]. Majority carrier transitions to and from levels at this crossover point will contribute to the loss. The location of the crossover point from the interface will vary with bias. However, the crossover point is always at the same energy with respect to mid-gap so that majority carrier density there will be constant and independent of bias. Therefore, the time constant for this process will be independent of bias. As time constant is not observed to be bias independent in depletion in the samples measured, majority carrier transitions to and from levels in the silicon bulk must make a negligible contribution to the loss compared to majority carrier transitions to and from interface states.

Loss due to the transit of majority carriers across the space-charge layer is negligible.[11,12] It can be shown using the relations given in Refs. 11 and 12 that the value of space-charge resistance for majority carriers when the silicon surface is intrinsic is four orders of magnitude less than majority carrier capture resistance in the samples measured.

The fact that curves (a), (b), and (c) in Fig. 14 can be fitted so well by straight lines implies that capture cross section in these regions is to a good approximation independent of energy in the bandgap. This independence is an experimental fact for which no satisfactory explanation exists at present. Because capture cross sections are independent of energy, they will be unaffected by statistical fluctuations of surface potential. The values of capture cross section obtained from the linear portions of Fig. 14 are therefore truly characteristic of the interface states present. Curves (b) and (c) in Fig. 14 measured on n-type are characterized by two types of interface states having different capture cross sections with an abrupt transition between them. The reason for this is not understood.

Fig. 14 shows that N_{ss} is slowly varying with energy. It is reasonable to conclude from this, and the observation, that the time constant varies monotonically with surface potential over the same portion of the band gap and that there is a continuum of interface states rather than a discrete level. It is also seen from Fig. 15 that N_{ss} is in the 10^{11} $\text{cm}^{-2}\text{-eV}^{-1}$ range. This means that the capturing centers are spaced too far apart in the plane of the interface for the wavefunction of an

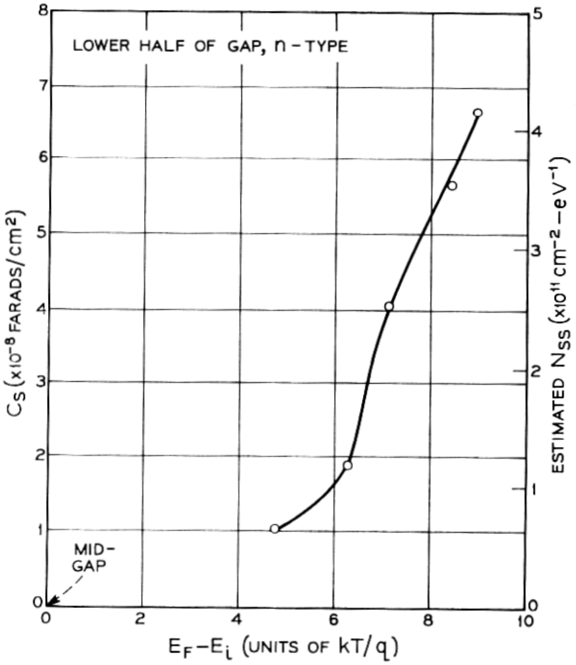

Fig. 27 — C_s and estimated N_{ss} vs $E_F - E_i$ for n-type sample in Fig. 14 in the weak inversion region. Estimated $N_{ss} = C_s/q$. C_s increases monotonically with $E_F - E_i$.

electron in one center to overlap a neighboring center. Transitions from one center to another are therefore, highly improbable so that it may be assumed that we do not have a band but just a continuum of levels closely spaced in energy. Transitions will only occur between the majority carrier band and levels within a few kT/q of the Fermi level. Cascading transitions from center to center which could occur in a band would probably have a different dependence of time constant on surface potential than shown in Fig. 14. Therefore, cascading transitions were not considered in the theory developed in Section IV.

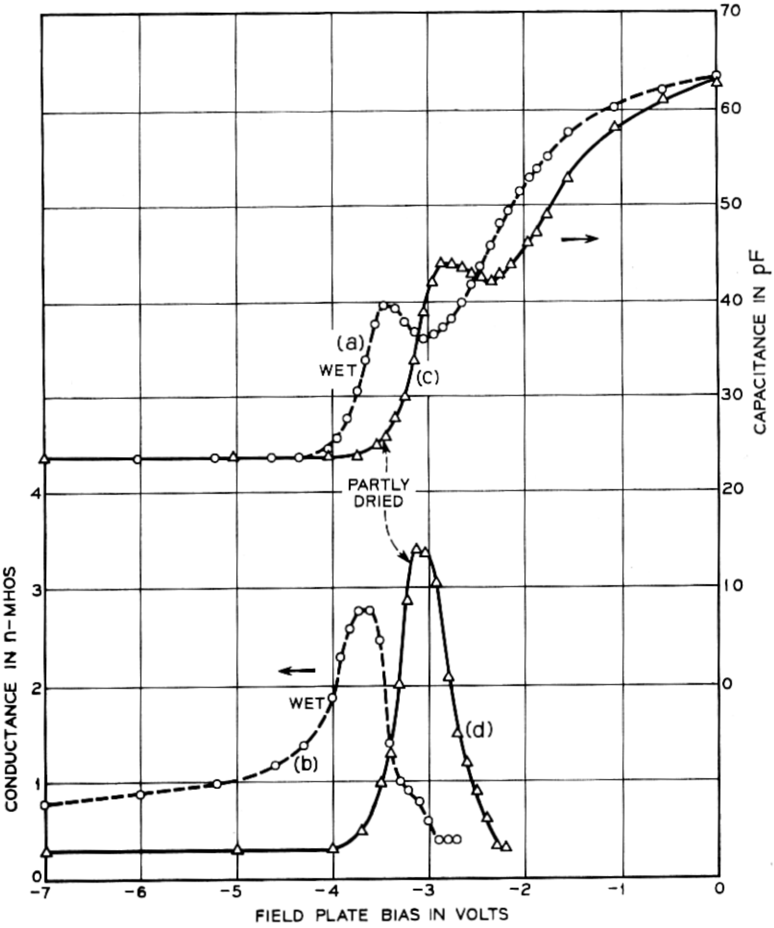

Fig. 28 — (a) capacitance and (b) equivalent parallel conductance measured vs field plate bias before drying; (c) capacitance and (d) equivalent parallel conductance vs field plate bias of same capacitor after heating in dry nitrogen for 17 hours at 350°C. This is the same sample as Fig. 4 only measured at 50 Hz.

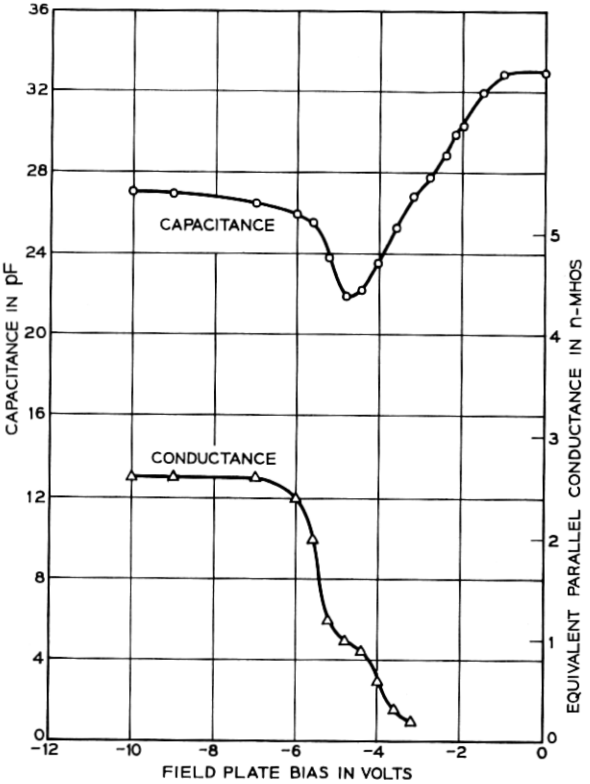

Fig. 29 — Capacitance and equivalent parallel conductance measured at 50 Hz vs field plate bias. Field plate diameter is 3.8×10^{-2}cm, silicon resistivity 1 ohm-cm, n-type, [100] oriented, and $C_{ox} = 2.8 \times 10^{-8}$farads/cm^2. Oxide is grown in steam then H_2 annealed as described in text to produce low N_{ss}.

The integration of (4) and (26) was performed for the case where N_{ss} and capture cross section do not vary appreciably over an energy interval of several kT/q. Figs. 14 and 15 show the experimental justification for this.

In Fig. 16, log τ_m is plotted as a function of $\bar{\psi}_s$ rather than $u_B - \bar{u}_s$ to illustrate more clearly the temperature dependence. An alternate and equivalent form of (22) which follows from (21) for p-type is

$$\tau_m = \frac{1}{\bar{v}\sigma_p N_A} \exp\,(q\psi_s/kT). \tag{60}$$

Curve (a) in Fig. 14 fits (60) quite well with a slope q/kT where $T =$ 204°K. Curve (b) fits with the slope for 300°K. Thus, the temperature dependence expected from (60) is observed. This agrees with Shockley-

Read theory[18] which is used in Section IV. Capture cross section can be obtained from (60) by extrapolating the curves in Fig. 16 to $\bar{\psi}_s = 0$. Doing this, we get $\sigma_p = 8.8 \times 10^{-17}$ cm^2 from curve (a) and $\sigma_p = 2.1 \times 10^{-16}$ cm^2 from curve (b). The capture cross section from curve (b) is of course in agreement with σ_p from curve (a) in Fig. 14. However, it is 2.4 times larger than σ_p from curve (a) of Fig. 16. This is probably due to the error in finding the value of Δ from (52). Normally, Δ can be found only within 20 to 30 mV of its correct value by the technique of Section V. Part of this error is due to neglecting the difference between hole and electron effective mass. It is considered likely that in reality capture cross section is nearly independent of temperature in this range. The apparent variation is ascribed instead to the errors in finding the value of Δ.

8.1.2 *Steam-Grown Oxides on [100] Orientation*

The same general conclusions can be reached for the [100] samples from Figs. 17 and 18 as for the [111] samples just discussed but there are two differences. The first difference is that the time constant curves in Fig. 14 for the upper half of the gap in [111] p-type are similar to the curves for the lower half of the gap in [100] p-type in Fig. 17. Also, the curves in the lower half of the gap in Fig. 14 for [111] p-type are similar to the curves in the upper half of the gap in [100] n-type in Fig. 17. The second difference is that interface state density in Fig. 15 for the lower half of the gap in [111] p-type is about twice as great as in Fig. 18 for the lower half of the gap in [100] p-type. However, interface state density in Fig. 15 for the upper half of the gap in [111] n-type is about the same as in Fig. 18 for the upper half of the gap in [100] n-type. The reason for these differences is not understood.

Built-in charge density in the [111] samples varies from sample to sample from 5×10^{11} cm^{-2} to 9×10^{11} cm^{-2} in both n and p-type. This is about twice as great as in the [100] samples confirming the observations of Balk.[29] Also, the increase of N_{ss} towards the band edges as shown in Figs. 15 and 18 is consistent with the results of Gray and Brown.[30]

8.1.3 *[111] Orientation, Partly-Dried Oxides*

Fig. 19 shows that interface state time constants fit (22) after partial drying. Capture cross section for n-type has not changed appreciably although two types of states were not found. This may be due to the fact that too small a range of surface potential was covered in the measurements. Because capture cross section did not change, partial drying does not appear to introduce new types of states—but only increases the number of states already present.

Comparison of curve (a) to curve (b) in Fig. 20 shows that interface state density for the partly-dried oxide is greater than for the wet oxide. This can be seen directly from Fig. 3 and 4 for p and n-type by the increase in the peak of the measured conductance with drying.

Fig. 3 for p-type shows by the shift of the curves to higher negative bias that a net positive charge has been induced in the oxide by partial drying. Fig. 4 for n-type shows by the shift of the curves towards less negative bias that a net negative charge has been induced in the oxide by partial drying. No conclusions about these results will be drawn because it is not certain that other changes independent of interface state density have not occurred in these samples.

The increase in interface state density with drying can explain the results observed in n^+p junctions in a previous paper.[31] In Ref. 31, it was found that drying an oxide heavily contaminated with alkali metal ions drastically reduced the current flowing in an n channel along the surface of the p-type material. The increased interface state density caused by drying results in more channel electrons being captured by interface states. This markedly reduces field-effect mobility which in turn results in an increase in the sheet resistance of the inversion layer. Although the inversion layer may be terminated by a defect at the edge of the wafer which shorts it to the p substrate, the current flow will be reduced to a much lower level than before.

8.1.4 *Range of N_{ss}*

The range of N_{ss} from 10^{12}cm^{-2}-eV^{-1} for very dry oxides through 10^{11} cm^{-2}-eV^{-1} for wet oxides to 10^{10} cm^{-2}-eV^{-1} for aluminum clad oxides annealed in H_2 or N_2 are empirical observations. It is not definitely known what these centers are. Therefore, an explanation can only be speculative. However, these results are qualitatively similar to previous work on "bare" surfaces which had at most a few monolayers of oxide.[32] Exposure to water vapor always reduced N_{ss} by a few orders of magnitude.[32] For the thick oxides studied in this work, water was most likely present in the form of OH groups which have a permanent dipole moment. An interesting model to explain the annihilation of fast states on "bare" surfaces has been proposed by Rzhanov.[33] Assuming that each interface state also has a permanent dipole moment, the lowest energy configuration will occur at the interface when the dipole moment of an H_2O or OH group is aligned antiparallel with the dipole moment of the nearest interface state. This could reduce the capture probability of the interface state[34] to such an extent that there is no measurable loss due to transitions involving it. The only loss measured would be due to transitions to and from the remaining unpaired states. For thick oxides, this fits the observed decrease in N_{ss} with increasing wetness and the observation that measured capture cross section remains unchanged. This model is, however, only one of several capable of explaining the results.

8.1.5 *Statistical Model*

The time constant dispersion actually observed is much broader than expected for a continuum. This can be seen by comparing Fig. 21 to the curves in Fig. 13. Three mechanisms which could cause this broadening are considered. They are:

(*i*) Tunnelling of majority carriers to states randomly distributed into the oxide.[2]

(*ii*) Fluctuations of capture cross section over the plane of the interface.

(*iii*) Random fluctuations of surface potential over the plane of the interface.

The tunnelling mechanism is unlikely to be the dominant cause of the time constant dispersion for the following reasons. The range of frequencies used in the measurements is from 50 Hz to 500 kHz. Thus, only those interface states having time constants shorter than several milliseconds will be measured. The time constant distribution for tunnelling would be[2]

$$\tau_t = (n_{so}\bar{v}\sigma_o)^{-1} \exp(2K_o\xi), \tag{61}$$

where σ_0 is the true electron capture cross section of the states, $1/2K_o$ is the tunnelling decay constant, and ξ is the distance into the oxide measured from the interface. For any reasonable value of $1/2K_o$, such as 1 Å given in Ref. 2, and reasonable values of σ_o, such as 10^{-15} cm^2 to 10^{-16} cm^2, it can be seen from (61) that states located more than a few Å into the oxide would have time constants much too long to be measured in the frequency range between 50 Hz and 500 kHz. Hysteresis in the capacitance vs bias characteristic which can be observed by slowly varying bias from a low to a high value and then back is negligibly small. Hysteresis would be observed when there is an appreciable density of states deep in the oxide having time constants longer than the time it takes to vary the bias because the charging and discharging of these states would lag behind the bias variation.[2] The fact that no hysteresis is observed means that the density of states in the oxide having time constants of the order of minutes is negligible in the samples measured. Finally, it is shown in Appendix D that for any reasonable value of K_o, the shape of the G_p/ω vs log ω curve calculated using (20) for the continuum of states and (61) for the tunnelling time constant distribution differs markedly from the observed G_p/ω vs log ω curve.

Although capture cross section is observed to be independent of energy over large portions of the band gap from Fig. 14, capture cross section could depend upon the location of the state in the plane of the interface. The assumption that capture cross section fluctuates over this plane does not give a broad enough time constant dispersion to fit the observed G_p/ω vs frequency curves.

Random fluctuations of surface potential over the plane of the interface is the model which fits best. One can see from (22) that small fluctuations of u_s will cause large fluctuations in τ_m. This broadens the time constant dispersion sufficiently to fit the experimental curves.

Random fluctuations of surface potential can be caused by:

(*i*) A random distribution of built-in charges and charged interface states over the plane of the interface.

(*ii*) A random distribution of ionized impurities in the silicon.

(*iii*) Random fluctuations of oxide thickness over the plane of the interface.*

Of these, the first two are the basis for deriving (39) in Section IV. The last was excluded because no dependence of time constant dispersion on oxide thickness was observed between 500 and 2000 Å in the samples measured. Fig. 21 shows the fit of (39) to experimental values of G_p/ω measured in the frequency range between 50 Hz and 500 kHz at a mean surface potential of 8.1. Fig. 21 shows that (39) accurately fits the experimental points. The fit of (39) to the experimental G_p/ω vs frequency points is equally accurate for all the other values of mean surface potential $\bar{u}_s$ measured in depletion. The only independent parameter in (39) varied to obtain these fits is the characteristic area α. By fitting (39) to the experimental G_p/ω vs frequency points at each value of $\bar{u}_s$ in this way, the values of α obtained are found to be related to space charge width W. This relation is shown in Fig. 22 where the length of the side of each square characteristic area $\alpha^{\frac{1}{2}}$ is plotted against

---

* This effect was pointed out to us by C. R. Crowell.[35]

W. Fig. 22 shows that $\alpha^{\frac{1}{2}}$ increases linearly with W and has a larger magnitude. To understand this relationship, consider a positive built-in charge or charged interface state located at a point in the plane of the interface. This charge is sandwiched between the plane of the field plate and the plane at the edge of the space-charge region a distance W away. An image charge is induced in each of these planes. The image charge in each plane induces another in the opposite plane and so on. Thus, the charge at the interface induces an infinite number of image charges in the field plate and space-charge edge planes. Let us now add a number of charges in the neighborhood of this charge in the plane of the interface and calculate the charge distribution induced by all the charges and their images in the plane at the edge of the space-charge region. The largest spacing between charges which results in a constant induced charge distribution in the plane at the edge of the space-charge region is $\alpha^{\frac{1}{2}}$. Thus, charges within a characteristic area are indistinguishable as separate charges which makes surface potential over a characteristic area uniform. The characteristic area is a measure of the granularity of the surface-charge distribution seen at the edge of the space-charge region. A calculation based on the image charge model shows that $\alpha^{\frac{1}{2}}$ has the values in Fig. 22 and a linear dependence on W. The value of α can be obtained within a factor of 2 from the image charge model so that α is not completely arbitrary.

We shall show next that the random distribution of built-in charges and charged interface states is the dominant cause of surface potential fluctuations in the samples measured. Rewriting (36) from Section IV for the standard deviation of surface potential caused by the random distribution of built-in charges and charged interface states gives

$$\sigma_s = \frac{W(\bar{u}_s)\beta(q\bar{Q}/\alpha)^{\frac{1}{2}}}{[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}. \tag{62}$$

Curve (b) in Fig. 30 is the variance σ_s^2 calculated from (62) as a function of W. $\bar{Q}$ calculated from (31) and (32) with $u_s = \bar{u}_s$ is found to be independent of bias from -1.8 volts to -2.2 volts. This bias range corresponds to mean surface potential from 10.3 to 3.8. Assuming a constant value of interface state density of 3×10^{11} cm^{-2}-eV^{-1} in the lower half of the gap from Fig. 15, the change in total interface state density is $3 \times 10^{11} \times (10.3 - 3.9) \times 0.026 = 5.1 \times 10^{10}$ cm^{-2}. This is negligibly small compared to $\bar{Q}/q = 9.8 \times 10^{11}$ cm^{-2}. Curve (b) in Fig. 30 is calculated using $\bar{Q} = 1.57 \times 10^{-7}$ coul/cm^2, $C_{ox} = 5.7 \times 10^{-8}$ farads/cm^2 and the values of α from Fig. 22.

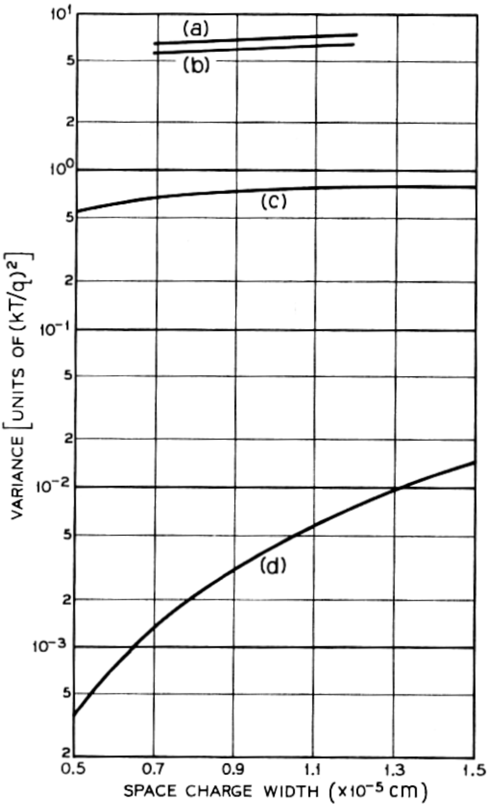

Fig. 30 — (a) total variance of surface potential ($\sigma_s^2 + \sigma_B^2$) caused by random distributions of built-in charges and charged interface states and ionized acceptors in the silicon space-charge region. σ_s and σ_B are calculated from (62) and (63), respectively. (b) variance of surface potential σ_s^2 caused by random distribution of built-in charges and charged interface states alone. σ_s is calculated from (62). (c) variance of surface potential σ_B^2 caused by random distribution of ionized acceptors in silicon space-charge region. σ_B is calculated from (63). (d) variance of surface potential σ_x^2 caused by fluctuations of oxide thickness. σ_x is calculated from (64).

For the standard deviation of surface potential caused by a random distribution of ionized acceptors in the silicon space-charge region, we rewrite (37) derived in paragraph B.1 of Appendix B

$$\sigma_B = \frac{q\beta[\bar{N}_A W(\bar{u}_s)]^{\frac{1}{2}}[1 - \exp(-\bar{u}_s)]}{2[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}. \qquad (63)$$

Curve (c) in Fig. 30 is the variance σ_B^2 calculated from (64) as a function of W. The values used are: $C_{ox} = 5.7 \times 10^{-8}$ farads/cm^2 and $\bar{N}_A = 2.1 \times 10^{16}$ cm^{-3}.

Finally, for the standard deviation of surface potential caused by oxide thickness fluctuation, we rewrite (101), derived in paragraph B.2 of Appendix B,

$$\sigma_x = \frac{qN_A\beta W^2(\bar{u}_s)[1 - \exp(-\bar{u}_s)]\, d\bar{x}}{[W(\bar{u}_s)\epsilon_{ox} + \bar{x}\epsilon_{si}]}. \tag{64}$$

Curve (d) in Fig. 30 is the variance σ_x^2 calculated from (64) as a function of W. The fluctuations of oxide thickness arise from the random nature of the oxidation process. It is reasonable to take $d\bar{x}$ as about one silicon lattice constant or 5Å. The values used in (64) to get curve (d) in Fig. 30 are: $\bar{x} = 595$ Å, $N_A = 2.1 \times 10^{16}$ cm^{-3}, and $d\bar{x} = 5$ Å.

Curve (a) in Fig. 30 is $\sigma_s^2 + \sigma_B^2$ vs W obtained by fitting (39) to the experimental G_p/ω vs frequency points. Fig. 30 clearly shows that the random distribution of built-in charges and charged interface states is the dominant cause of surface potential fluctuations in the samples measured. It can be easily seen that $\alpha^{\frac{1}{2}}$ should vary linearly with W as follows. Over the range of surface potential measured, the experimental G_p/ω vs frequency curves broaden very slightly with increasing surface potential. This is shown quantitatively in curve (a) of Fig. 30. Over the same range of surface potential, curve (c) in Fig. 30 is virtually independent of W. The dominant term σ_s^2 in curve (a) is therefore, nearly independent of W as shown in curve (b) of Fig. 30. For σ_s constant, an inspection of (62) will show that for $\bar{Q}$ constant and $WC_{ox} < \epsilon_{si}$ as is the case in the samples measured, $\alpha^{\frac{1}{2}}$ must vary linearly with W.

The density of built-in charges and charged interface states is in the high 10^{11} cm^{-2} range in the samples measured. Therefore, these charges are about 100 Å apart on the average which is too far for coulombic interaction between them to be important. This makes the assumption that they are randomly distributed a reasonable one. The mean number of built-in charges and charged interface states $\bar{N}$ in a characteristic area α can be calculated from the relation $\bar{N} = \alpha\bar{Q}/q$ (see Section IV) using values of α obtained from Fig. 22. $\bar{N}$ is found to vary from 140 charges at $W = 6.9 \times 10^{-5}$ cm to 246 charges at $W = 1.07 \times 10^{-5}$ cm. For these values of $\bar{N}$, there is negligible error in using the Gaussian approximation to the Poisson distribution (28) in deriving (39) in Section IV. The mean number of ionized acceptors calculated from the relation $\bar{M} = W^3\bar{N}_A$ varies from 26 charges at $W = 1.07 \times 10^{-5}$ cm to 7 charges at $W = 6.9 \times 10^{-5}$ cm. For all values of $\bar{M}$ except perhaps the last, there is negligible error in using the Gaussian approximation to the Poisson distribution (81) in paragraph B.1 of Appendix B in deriving (39). Any error at $W = 6.9 \times 10^{-5}$ cm is not important to the final results because the contribution of σ_B^2 to the total variance is so small. The assumption of small fluctuations made in deriving (39) also seems to be reasonable because of the rather good fit obtained to the experimental points.

One can decide from (62), (63), and (64) which influence will be dominant for samples other than the ones used here. In spite of the fact that the samples investigated here were dominated by random fluctuations of built-in charges and charged interface states, it should be borne in mind that the other influences may become important in other cases. When built-in charge and interface state density is low and the oxide relatively thick, then bulk doping fluctuations will dominate. When the oxide is thin and the impurity density is high, random fluctuations of oxide thickness become important. The general conclusion is that there will always be random fluctuation of surface potential in simiconductor-insulator interfaces.

No data has been taken at flat bands or in the accumulation region. However, it is expected that the equivalent circuit in Fig. 8(a) will also be valid in this region with all the distributed time constants shifted to shorter values.

8.2 *Weak Inversion Region*

In this part, it will be shown that the loss in weak inversion can arise from generation-recombination either through bulk states or interface states and how the conductance technique can be used to distinguish between them. It will be shown also that the observed effects in this region can be explained with a continuum of states rather than a single level near a band edge.

When the silicon surface becomes inverted, the Fermi level will cross impurity levels near mid-gap as seen from Fig. 5(c). Generation recombination through bulk states as well as through interface states can contribute significantly to the measured loss. These two cases can be distinguished by the bias dependence of the measured equivalent parallel conductance. When generation-recombination through interface states dominates the loss, the equivalent parallel conductance goes through a peak as a function of bias in weak inversion because time constant varies inversely with majority carrier density at the surface (see paragraph 4.6). This case is illustrated in Fig. 23 where equivalent parallel conductance goes through a peak in weak inversion and drops to a very low value in strong inversion.

When generation-recombination through bulk states dominates the loss, equivalent parallel conductance does not go through a peak as a function of bias in weak inversion. The reason for this is that the loss due to this process is bias independent. This follows from the fact that hole and electron densities at the crossover point [see Fig. 5(c)], where the Fermi level crosses impurity levels located near mid-gap, are independent of bias. This case is illustrated in Fig. 29. The slight indication of a peak in the conductance curve means that generation recombination through interface states makes a slight contribution to the loss. However, the primary effect is that of a bias independent loss. The variation of measured conductance with bias is mainly due to variation of capacitance with bias.

Fig. 23 is measured on a sample having N_{ss} in the 10^{11} $\text{cm}^{-2}\text{-eV}^{-1}$ range while Fig. 29 is measured on a sample having N_{ss} in the 10^{10} $\text{cm}^{-2}\text{-eV}^{-1}$ range. Thus, samples can be made in which one or the other process dominates the loss. It should be noted that the conductance technique can be used to investigate generation-recombination through interface states or bulk states in the weak inversion region by choosing an appropriate sample. The investigation of generation-recombination through bulk states is beyond the scope of this work so that most of the samples investigated have interface state loss dominant in weak inversion. Therefore, bulk generation-recombination is made zero in the equivalent circuits of Fig. 9.

In general, samples having N_{ss} in the 10^{11} $\text{cm}^{-2}\text{-eV}^{-1}$ range or higher will have interface state loss dominant in the weak inversion range. Additional evidence for this in such samples is shown in Fig. 28. The only difference between Fig. 28 and Fig. 4 is the low signal frequency which is used in Fig. 28 to reveal the behavior of equivalent parallel conductance in the weak inversion region. The conclusion is again drawn from Fig. 28 that interface state loss must be dominant as from Figs. 3 and 4 discussed previously in Section III.

The striking feature of the experimental G_p/ω vs frequency curves at each bias in weak inversion is that they can be accurately fitted by a single time constant using (17). An example of this is shown in Fig. 25. Three possible explanations of this will be considered.

(*i*) A single level state located at an energy in the silicon band gap corresponding to the bias at the capacitance peak in Fig. 23.

(*ii*) A single level near the valence band edge.

(*iii*) A continuum of interface states over the silicon band gap.

The first explanation can be ruled out by an examination of Figs. 24, 26, and 27. Fig. 24 shows that the capacitance peak shifts to greater negative bias with decreasing signal frequency. The peak must always occur at the same bias for a single level. In Figs. 26 and 27, the range of $u_B - \bar{u}_s$ and $E_F - E_i$ corresponds to a range of bias which includes the capacitance peak in Fig. 23. Both τ_m in Fig. 26 and C_s in Fig. 27 increase monotonically through the corresponding capacitance peak in Fig. 23. For a single level state, C_s would go through a peak and τ_m first would be constant for E_F on the valence band side of the single level and then it would decrease exponentially with u_s for E_F on the conduction band side.

The range of $E_F - E_i$ and $u_B - u_s$ in Figs. 26 and 27 is limited by the capacitance dispersion shown in Fig. 24. This means that (52) cannot be used without significant error too far beyond the capacitance peak (at reasonable frequencies).

To rule out the possibility of a single level state near the valence band edge, G_p/ω vs frequency curves were obtained on a p-type sample This sample had steam-grown oxide on a 1 Ω-cm crystal oriented in the [100] direction. The [100] direction has a lower built-in charge density so that masking effects due to lateral ac current flow into the inverted layer beyond the field plate are minimized. A single time constant was again found to be characteristic of the G_p/ω vs frequency curves in the weak inversion region. The experiment was repeated on an n-type sample oriented in the [100] direction having a steam-grown oxide with the same result. The fact that the same result was obtained in weak inversion on both n and p-type samples and the lack of evidence for a single level state in the same energy range in accumulation in these samples makes such a state as an explanation of the effect very unlikely. The observations can be best explained in terms of a continuum of states as done in paragraph 4.6.

Evidence that this effect is dominated by the majority carrier density at the silicon interface is given by Fig. 26. Fig. 27 shows that C_s increases so rapidly with E_F that the approximation of a constant density of states over an interval of kT/q cannot be made. Thus, (18) cannot be simply integrated. The fact that C_s is so rapidly varying is the reason why the curve in Fig. 26 does not have the same magnitude of slope as the curves in Fig. 14. This rapid variation of C_s also explains the capacitance peak in Fig. 23. It can be seen that (16) will go through a peak if C_s has a strong dependence on bias.

Comparison of Fig. 27 with Fig. 15 shows that interface state density in the lower half of the silicon bandgap in n-type has a different magnitude and dependence on position of Fermi level than in the lower half of the gap in p-type. The reason for this is not understood.

8.3 *Limitations of the Conductance Technique*

Interface state densities in the 10^9 cm^{-2}-eV^{-1} range have been measured by this technique using the apparatus described in Section V. For a field plate diameter of 5×10^{-2} cm, this corresponds to about 2×10^6 states/eV. For a signal of 20 mV peak-to-peak, transitions involving only about 4×10^4 states can be detected.

It is important to use as thin an oxide layer as practicable to get large values of capacitance and equivalent parallel conductance for a given N_{ss} and to minimize the error in extracting C_{ox}.

The range of bandgap energy over which measurements were made in this work was limited by the frequency range 50 Hz to 500 kHz of the two bridges used.

IX. SUMMARY AND CONCLUSIONS

(*i*) Majority carrier transitions to and from interface states make the dominant contribution to the measured equivalent parallel conductance in the accumulation, depletion, and inversion regions. This fact makes it possible to use simplified equivalent circuits and extract interface state properties for each type of carrier independntly.

(*ii*) A continuum of Shockley-Read type interface state levels closely spaced in energy across the silicon band gap seems to be characteristic of the Si–SiO$_2$ interface.

(*iii*) Interface state densities in the 10^{11} cm^{-2}-eV^{-1} range are usual for "as-grown" steam oxides.

(*iv*) Interface state density can be increased to the 10^{12} cm^{-2}-eV^{-1} range by drying the oxide.

(*v*) Interface state density can be reduced to the 10^{10} cm^{-2}-eV^{-1} range by H$_2$ or N$_2$ annealing at low temperatures with an aluminum film deposited on a steam-grown oxide.[29]

(*vi*) Capture cross sections for electrons and holes are independent of energy over large portions of the silicon bandgap. They are also independent of temperature, the processes of drying or H$_2$ or N$_2$ annealing, and, except for subtle differences, crystal orientation.

(*vii*) In depletion, a time constant dispersion is observed which is much broader than expected for a continuum. The broadened time constant dispersion is determined mainly by statistical fluctuations of surface potential. In the samples studied, a random distribution of built-in charges and charged interface states was found to be the primary cause of the surface potential fluctuations. This model provides an accurate fit to the experimental points and yields a characteristic area whose linear dimension has about twice the magnitude and has the same bias dependence as the silicon space-charge width. Charges within a characteristic area cannot be distinguished as separate charges. Random fluctuations of surface potential can also arise from a random distribution of ionized impurities in the bulk silicon and by random fluctuations of oxide thickness. In samples where the built-in charge and charged interface state density is low or the oxide layer thin, these can replace the random distribution of built-in charges and charged interface states as the dominant causes of surface potential fluctuations.

(*viii*) In weak inversion, the fast response time of minority carriers and the conductance of the inversion layer results in the observation of a single time constant at each bias.

(*ix*) The interface state branch of the equivalent circuit in the depletion-accumulation region where $\bar{u}_s < u_B$ is given in Fig. 31(a). The equivalent circuit in the mid-gap region where $\bar{u}_s = u_B \pm a$ few kT/q is shown in Fig. 31(b). Finally, the equivalent circuit in the weak inversion region where $u_B < \bar{u}_s < 2u_B$ is shown in Fig. 31(c).

(*x*) It is impossible to make an MIS structure in which surface potential is absolutely uniform because even if built-in charge and interface state density is made small, random distribution of ionized bulk impurities or random fluctuations of oxide thickness will still be present.

(*xi*) The accuracy required for extracting this detailed a picture of the Si–SiO$_2$ interface from MIS capacitor measurements could only be provided by the conductance technique (see Fig. 1).

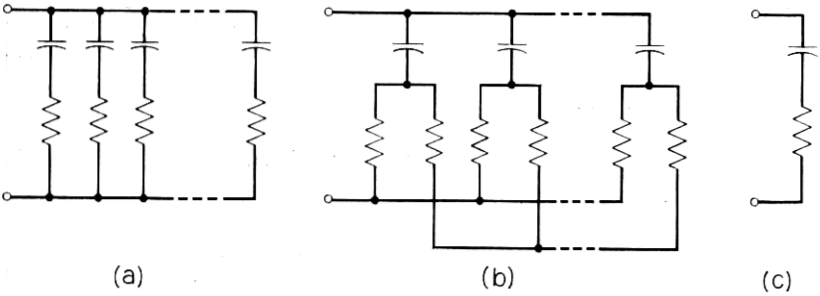

Fig. 31 — (a) Schematic of interface state branch of equivalent circuit in depletion-accumulation region where $\bar{u}_s < u_B$. (b) Schematic of interface state branch of equivalent circuit in mid-gap region where $\bar{u}_s = u_B \pm$ a few kT/q. (c) Schematic of interface state branch of equivalent circuit in weak inversion region where $u_B < \bar{u}_s < 2u_B$.

X. ACKNOWLEDGMENTS

The authors are indebted to Messrs. J. T. Nelson, C. N. Berglund, and C. Goldberg for their criticisms and many valuable suggestions during the course of this work. We also thank Mr. V. Heine for stimulating discussions, Mr. R. M. Ryder for a critical reading of the manuscript, Mr. A. D. Lopez for making the measurements, Mr. J. McGlasson for preparing the samples, Mr. A. Hartman for proofreading the manuscript, and Mrs. J. M. Schilling for help in the computer programming.

APPENDIX A

In this Appendix, interface state admittance is derived from (3):

$$i_s(t) = qN_sc_n[1 - f(t)]n_s(t) - qN_se_nf(t). \tag{65}$$

Expressing $f(t)$ as the sum of a dc and an ac part:

$$f(t) = f_o + \delta f, \tag{66}$$

where f_o is the Fermi function established by the bias and δf the change caused by the ac signal. Similarly,

$$n_s(t) = n_{so} + \delta n_s\ , \tag{67}$$

where n_{so} is the electron density at the silicon surface when the Fermi level is at the trap level and δn_s the change caused by the ac signal. Substituting (66) and (67) into (65) and making the small signal approximation by neglecting second order terms, we get

$$i_s(t) = qN_sc_n[(1 - f_o)n_{so} + (1 - f_o)\ \delta n_s - n_{so}\ \delta f] - qe_nN_s(f_o + \delta f). \tag{68}$$

From detailed balance

$$qN_sc_n(1 - f_o)n_{so} = qe_nN_sf_o\ . \tag{69}$$

Solving (69) for e_n and substituting into (68) we get

$$i_s(t) = qN_sc_n\left[(1 - f_o)\ \delta n_s - n_{so}\frac{\delta f}{f_o}\right]. \tag{70}$$

The net current density can also be expressed as

$$i_s(t) = qN_s\frac{df}{dt}. \tag{71}$$

Equating (71) to (70)

$$\frac{df}{dt} = c_n(1 - f_o)\ \delta n_s - c_n n_{so} \frac{\delta f}{f_o}. \tag{72}$$

The small signal variation of the Fermi function is $\delta f = f_m \exp(j\omega t)$. Combining this with (66) results in

$$\frac{df}{dt} = j\omega\ \delta f. \tag{73}$$

Substituting (73) into (72) and solving for δf

$$\delta f = \frac{f_o(1 - f_o)\ \delta n_s}{n_{so}(1 + j\omega f_o/c_n n_{so})}. \tag{74}$$

Substituting (74) into (70) results in

$$i_s(t) = \frac{j\omega q N_s f_o(1 - f_o)}{(1 + j\omega f_o/c_n n_{so})} \frac{\delta n_s}{n_{so}}. \tag{75}$$

The ratio $\delta n_s/n_{so}$ is equal to the ac surface potential. This can be seen as follows:

$$\frac{\delta n_s}{n_{so}} = \frac{n_i \exp(u_s - u_B)\ \delta u_s}{n_i \exp(u_s - u_B)} = \delta u_s\ , \tag{76}$$

where n_i is the intrinsic carrier density and u_s is the silicon band bending or surface potential in units of kT/q and u_B the potential difference between mid-gap and Fermi level in the quasi neutral region of the silicon in units of kT/q. Expressing δu_s in volts is

$$\delta u_s = \frac{q}{kT}\ \delta\psi_s\ . \tag{77}$$

Equation (75) becomes

$$i_s(t) = j\omega \frac{q^2}{kT} \frac{N_s f_o(1 - f_o)\ \delta\psi_s}{(1 + j\omega f_o/c_n n_{so})}. \tag{78}$$

This can be written as

$$i_s(t) = Y_s\ \delta\psi_s\ , \tag{79}$$

where

$$Y_s = j\omega \frac{q^2}{kT} \frac{N_s f_o(1 - f_o)}{(1 + j\omega f_o/c_n n_{so})}. \tag{80}$$

APPENDIX B

B.1 *Random Fluctuations of Impurity Charges in the Silicon*

We shall use the model proposed by Shockley[23] for this calculation. Otherwise the procedure for finding the standard deviation of the surface potential is similar to that for the case of the surface charges.

Let $\bar{M}$ be the number of ionized acceptors in a characteristic volume in the silicon space-charge region. When $\bar{M}$ is large, the probability that there are M ionized acceptors in a characteristic volume is given by the Gaussian approximation of a Poisson distribution

$$P(M) = (2\pi\bar{M})^{-\frac{1}{2}} \exp [-(M - \bar{M})^2/2\bar{M}]. \tag{81}$$

The characteristic volume from Ref. 23 is a cube in which the length of a side is equal to the space-charge width W. The first transformation of the probability density function is from number of charges to number density

$$P(N_A) = P(M)\, dM/dN_A\ , \tag{82}$$

where N_A is the ionized acceptor density. The relation between M and N_A is

$$M = N_A W^3. \tag{83}$$

Combining (81), (82), and (83), the result of the first transformation is

$$P(N_A) = (2\pi\bar{N}_A W^3)^{-\frac{1}{2}} W^3 \exp [-W^3(N_A - \bar{N}_A)^2/2\bar{N}_A], \tag{84}$$

where $\bar{N}_A$ is the mean acceptor density.

The second transformation from number density to surface potential is

$$P(u_s) = P(N_A)\, dN_A/du_s\ . \tag{85}$$

Rewriting (31) and (32) results in

$$Q = C_{ox}(v_o + u_s/\beta) + (2q\epsilon_{si}N_A/\beta)^{\frac{1}{2}}[\exp(-u_s) + u_s - 1]^{\frac{1}{2}}. \tag{86}$$

Solving (86) for N_A is

$$N_A = \left(\frac{\beta}{2q\epsilon_{si}}\right) \frac{[Q - C_{ox}(v_o + u_s/\beta)]^2}{[\exp(-u_s) + u_s - 1]}. \tag{87}$$

Equation (87) is not a single-valued relation between N_A and u_s. We therefore restrict the problem as with the surface charges to the case where the fluctuations $N_A - \bar{N}_A$ are very small. For very small fluctuations, (87) can be differentiated assuming Q and oxide thickness to be uniform to get

$$dN_A = -\left\{\left(\frac{2N_A}{q\epsilon_{si}\beta}\right)^{\frac{1}{2}} \frac{C_{ox}}{[\exp(-u_s) + u_s - 1]^{\frac{1}{2}}} + \frac{N_A[1 - \exp(-u_s)]}{[\exp(-u_s) + u_s - 1]}\right\} du_s . \tag{88}$$

Equation (88) is now expressed as a function of u_s using (32)

$$C_D = \frac{\epsilon_{si}}{W} = \beta \frac{dQ_{sc}}{du_s} = \left[\frac{q\epsilon_{si}N_A\beta}{2[\exp(-u_s) + u_s - 1]}\right]^{\frac{1}{2}} [1 - \exp(-u_s)]. \tag{89}$$

Solving (89) for $\exp(-u_s) + u_s - 1$, we have

$$\exp(-u_s) + u_s - 1 = \frac{qN_A\beta}{2\epsilon_{si}} W^2[1 - \exp(-u_s)]^2. \tag{90}$$

Substituting (90) into (88) we get

$$dN_A = -\frac{2(WC_{ox} + \epsilon_{si})}{q\beta W^2[1 - \exp(-u_s)]} du_s . \tag{91}$$

Because $\bar{N}_A$ is given by (87) when $u_s = \bar{u}_s$, dN_A can be evaluated about $\bar{N}_A$ at each bias from

$$dN_A = -\frac{2[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}{q\beta W^2(\bar{u}_s)[1 - \exp(-\bar{u}_s)]} du_s . \tag{92}$$

Equation (92) is the transformation equation sought. Replacing dN_A and du_s in (92) by the small fluctuations $N_A - \bar{N}_A$ and $u_s - \bar{u}_s$ respectively, we get

$$N_A - \bar{N}_A = -\frac{2[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}{q\beta W^2(\bar{u}_s)[1 - \exp(-\bar{u}_s)]} (u_s - \bar{u}_s). \tag{93}$$

Combining (84), (85), (92), and (93), we get

$$P(u_s) = (2\pi\sigma_B^2)^{-\frac{1}{2}} \exp[-(u_s - \bar{u}_s)^2/2\sigma_B^2], \tag{94}$$

where the standard deviation σ_B is

$$\sigma_B = \frac{q\beta[\bar{N}_A W(\bar{u}_s)]^{\frac{1}{2}}[1 - \exp(-\bar{u}_s)]}{2[W(\bar{u}_s)C_{ox} + \epsilon_{si}]}. \tag{95}$$

B.2 *Random Fluctuations of Oxide Thickness*

Solving (86) for x, the oxide thickness,

$$\frac{\epsilon_{ox}}{C_{ox}} = x = \frac{\epsilon_{ox}(v_o + u_s/\beta)}{Q - \{(2q\epsilon_{si}N_A/\beta)[\exp(-u_s) + u_s - 1]\}^{\frac{1}{2}}}, \tag{96}$$

where ϵ_{ox} is the dielectric permittivity of the oxide layer. If Q is not zero, it will influence the fluctuations of surface potential caused by fluctuations in oxide thickness even if it is uniform. To get the influence of oxide thickness fluctuations independent of surface charge, we let $Q = 0$. Thus, (96) becomes

$$x = -\frac{\epsilon_{ox}(v_o + u_s/\beta)}{\{(2q\epsilon_{si}N_A/\beta)[\exp(-u_s) + u_s - 1]\}^{\frac{1}{2}}}. \tag{97}$$

Differentiating (97), assuming N_A to be uniform, results in

$$dx = -\frac{\epsilon_{ox}}{(2q\epsilon_{si}N_A/\beta)^{\frac{1}{2}}}\left\{\frac{1}{\beta[\exp(-u_s) + u_s - 1]^{\frac{1}{2}}} - \frac{(v_o + u_s/\beta)[1 - \exp(-u_s)]}{2[\exp(-u_s) + u_s - 1]^{\frac{3}{2}}}\right\} du_s\,. \tag{98}$$

Eliminating $\exp(-u_s) + u_s - 1$ from (98) by using (90) we get

$$dx = -\frac{W\epsilon_{ox} + x\epsilon_{si}}{q\beta N_A W^2[1 - \exp(-u_s)]}\,du_s\,. \tag{99}$$

Equation (97) gives $\bar{x}$ when $u_s = \bar{u}_s$ so that dx evaluated about $\bar{x}$ is

$$d\bar{x} = -\frac{W(\bar{u}_s)\epsilon_{ox} + \bar{x}\epsilon_{si}}{q\beta N_A W^2(\bar{u}_s)[1 - \exp(-\bar{u}_s)]}\,d\bar{u}_s\,. \tag{100}$$

$d\bar{x}$ is the standard deviation of x and $d\bar{u}_s$ the standard deviation of surface potential thus,

$$d\bar{u}_s = \sigma_x = \frac{qN_A\beta W^2(\bar{u}_s)[1 - \exp(-\bar{u}_s)]}{[W(\bar{u}_s)\epsilon_{ox} + \bar{x}\epsilon_{si}]}\,d\bar{x}. \tag{101}$$

APPENDIX C

There is a simple way to relate $C_s(\omega, \bar{u}_s)$ to measured $(G_p/\omega)_{\max}$. The distributed interface state network in Fig. 8(a) can be represented at one frequency and bias by an equivalent series R_sC_s network in which R_s and C_s are frequency and bias dependent. The equivalent circuit will then look like Fig. 6. If we let R_s and C_s be independent of frequency, G_p/ω vs frequency from (17) can be made to coincide with the measured G_p/ω vs frequency curve at each bias at $(G_p/\omega)_{\max}$ as shown in Fig. 32. Therefore, from (17), $C_s(\omega, \bar{u}_s) = 2(G_p/\omega)_{\max}$ at each bias and we obtain (54) by substituting this into (53) and solving for C_D.

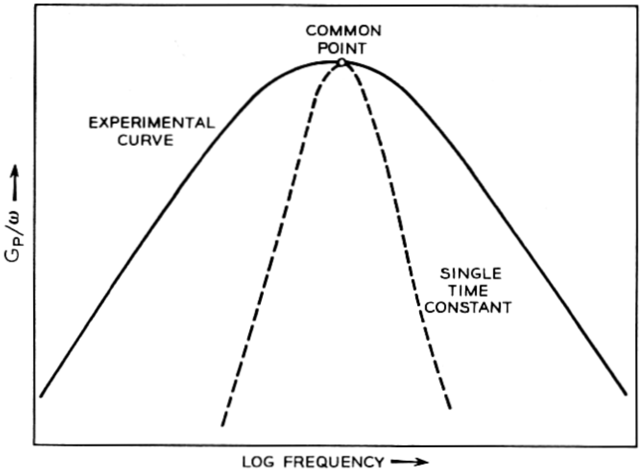

Fig. 32 — Schematic of G_p/ω vs log frequency showing how C_s is obtained from measured G_p/ω peak.

APPENDIX D

We shall calculate G_p/ω vs ω which would result from transitions between the conduction band near the surface and states distributed with uniform density into the oxide. The time constant distribution for tunnelling from (61) is

$$\tau_t = \tau_{ot} \exp (2K_o\xi), \tag{102}$$

where $\tau_{ot} = (n_{so}\bar{v}\sigma_o)^{-1}$.

The contribution to G_p/ω from states located between ξ and $\xi + d\xi$ in the oxide is

$$d(G_p/\omega) = (q/2\omega)(N_{ss}/x_L)\tau_{ot} \exp (-2K_o\xi) \cdot \ln [1 + \omega^2\tau_{ot}^2 \exp (4K_o\xi)]\, d\xi, \tag{103}$$

where (102) has been substituted for τ in (20) for the continuum of states, N_{ss}/x_L is the uniform density of states in the oxide, cm^{-3}-eV^{-1}, and x_L is the oxide thickness, cm.

Total G_p/ω from all the states in the oxide therefore, is

$$G_p/\omega = (q/2\omega)(N_{ss}/x_L)\tau_{ot} \int_0^{x_L} \exp (-2K_o\xi) \cdot \ln [1 + \omega^2\tau_{ot}^2 \exp (4K_o\xi)]\, d\xi. \tag{104}$$

To get G_p/ω vs ω, (104) is numerically integrated on an IBM 7094 computer using the trapezoidal rule. The result is shown in Fig. 33. Because we are interested only in the shape of the curve, the frequency scale is arbitrary.

G_p/ω is normalized to eliminate N_{ss}/x_L. A plot of normalized G_p/ω obtained from experiment is also shown in Fig. 33 for comparison.

The shape of the curve calculated from the tunnelling model is determined essentially by (102). The shortest time constant in this distribution is when $\xi = 0$ which is right at the interface. G_p/ω drops off rapidly as the period of the applied signal becomes comparable to and shorter than this time constant. When the period of the applied signal is long compared to the shortest time constant in (102), G_p/ω becomes a constant as we have assumed a constant density of states into the oxide. The main purpose of assuming a uniform density of states into the oxide is to illustrate the asymmetry of the G_p/ω vs log ω curve introduced by the tunnelling time constant distribution (102). The G_p/ω vs log ω curve will go through a maximum if the density of states decreases with distance into the oxide. As long as the density of states decreases more slowly with distance into the oxide than the time constants given by (102) increase, G_p/ω will decrease faster on the high frequency side of the peak than on the low-frequency side.

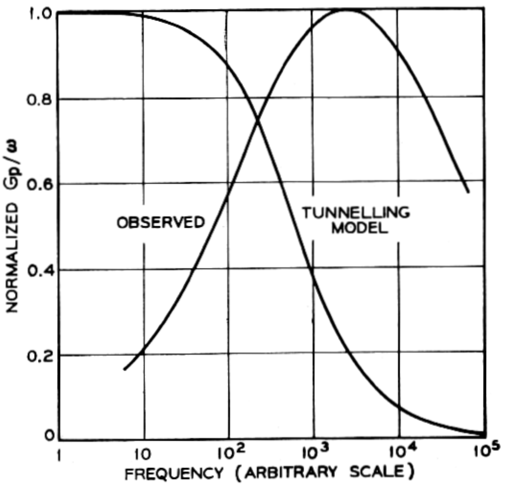

Fig. 33 — Normalized G_p/ω vs log frequency for tunnelling calculated from (104) and normalized G_p/ω vs log frequency obtained for measurement. Frequency scale is arbitrary to compare the shapes of two curves.

The G_p/ω vs log ω curve obtained from measurement is symmetric about its peak. A distribution of states into the oxide which would would give a G_p/ω vs log ω curve symmetric about its peak using the tunnelling model would be a peaked density distribution having its maximum some distance into the oxide. In addition, the width of this distribution would have to be considerably smaller than $1/2K_o$ to get G_p/ω vs log ω curves in the frequency range between 50 Hz and 500 kHz. Such a distribution would be rather arbitrary. Thus, it can be seen from the symmetry of the observed G_p/ω vs log ω curves that tunnelling to states distributed into the oxide is not likely to be the dominant cause of time constant dispersion in the samples measured.

LIST OF SYMBOLS

a	Maximum amplitude of the ac surface potential in volts.
b	Maximum amplitude of applied ac signal in volts.
C	Equivalent parallel capacitance of interface state branch of equivalent circuit having a single time constant in farads/cm^2.
CB	Representation of conduction band.
C_D	Depletion layer capacitance in farads/cm^2.
C_I	Inversion layer capacitance in farads/cm^2.
C_m	Measured equivalent parallel capacitance of the MIS capacitor in farads/cm^2.
c_n	Electron capture probability in cm^3/sec.
C_{om}	Low-frequency capacitance of MIS capacitor in farads/cm^2.
C_{ox}	Oxide layer capacitanec in farads/cm^2.
c_p	Hole capture probability in cm^3/sec.
C_p	Equivalent parallel capacitance of depletion layer and interface state branch of equivalent circuit in farads/cm^2.
C_s	Interface state capacitance associated with a single time constant in farad/cm^2.
$d\nu$	Number of characteristic areas in which the number of surface charges is between N and $N + dN$ or the surface potential is between u_s and $u_s + du_s$.
E_F	Fermi level in units of kT/q.
E_i	Mid-gap energy in units of kT/q.
e_n , e_p	Emission constants for electrons and holes in sec^{-1}.
f_m	Maximum value of the perturbation of the Fermi function by the ac signal.
f_o	Fermi function at the dc bias.

$f(t)$	Fermi function as a function of time.
G_m	Measured equivalent parallel conductance of MIS capacitor in mhos/cm^2.
$G_n(t)$	Emission rate of electrons from a single level interface state in cm^{-2}-sec^{-1}.
G_p	Equivalent parallel conductance of the interface state branch of the equivalent circuit in mhos/cm^2.
G_p'	Equivalent parallel conductance corrected for C_{ox} in mhos.
$G_p(t)$	Emission rate of holes from a single level interface state in cm^{-2}-sec^{-1}.
G_s	Majority carrier capture conductance of a single level interface state in mhos/cm^2.
i_s	Current density charging the interface states in amp/cm^2.
i_{sc}	Current density charging the space-charge layer in amp/cm^2.
i_T	Current density through MIS capacitor in amp/cm^2.
k	Boltzman's constant in eV-coul/°K.
M	Number of ionized acceptors in a characteristic volume.
$\bar{M}$	The mean of M.
N	Number of built-in charges and charged interface states in a characteristic area.
$\bar{N}$	The mean of N.
N_A	Acceptor impurity density in cm^{-3}.
$\bar{N}_A$	Mean ionized acceptor density in cm^{-3}.
N_D	Donor impurity density in cm^{-3}.
n_i	Intrinsic carrier concentration in cm^{-3}.
N_s	Density of a single level interface state in cm^{-2}.
n_{so}	Electron density established by the dc bias in cm^{-3}.
N_{ss}	Density of interface states, $dN_s/d\psi$ in cm^{-2}-eV^{-1}.
$n_s(t)$	Time dependent electron density at the silicon surface in cm^{-3}.
$P(M)$	Probability that there are M ionized acceptors in a characteristic volume.
$P(N)$	Probability of finding N charges in a characteristic volume.
$P(Q)$	Probability that the surface-charge density is Q in cm^2/coul.
p_{so}	Hole density established by the dc bias in cm^{-3}.
$p_s(t)$	Time dependent hole density at the silicon surface in cm^{-3}.
$P(u_s)$	Probability that surface potential in a characteristic area is u_s .
q	Electronic charge in coulomb.
Q	Sum of built-in and interface state charge densities in coul/cm^2.

$\bar{Q}$	Mean surface-charge density in coul/cm^2.
Q_f	Fixed charge density in the oxide in coul/cm^2.
Q_s	Total interface state charge density in coul/cm^2.
Q_{sc}	Silicon space-charge density in coul/cm^2.
Q_T	Total charge density in coul/cm^2.
$R_{n,s}$; $R_{p,s}$	Electron and hole capture resistance of a single level in ohms-cm^2.
$R_n(t)$, $R_p(t)$	Capture rate of electrons and holes by a single level interface state in cm^{-2}-sec^{-1}.
R_s	Majority carrier capture resistance of a single level interface state in ohm-cm^2.
t	time in seconds.
T	Absolute temperature in °K.
u	Potential difference between mid-gap and an interface state level in units of kT/q.
u_B	$= \ln (n_i/N_D)$, for n-type silicon $= \ln (N_A/n_i)$, for p-type silicon.
u_F	Potential difference between mid-gap and Fermi level at silicon surface in units of kT/q
u_o	$\ln \omega/c_n n_i$.
u_s	Surface potential in units of kT/q.
$\bar{u}_s$	Mean surface potential in units of kT/q.
VB	Representation of valence band.
$\bar{v}$	Average thermal velocity of holes and electrons in cm/sec.
$v_a(t)$	Time dependent voltage applied to MIS capacitor in volts.
v_o	dc bias across MIS capacitor in volts.
v_{o1}	Arbitrary dc bias in volts.
W	Depletion layer thickness in cm.
$\bar{x}$	Mean oxide thickness measured from the interface in cm.
y	$\ln \omega\tau_m = u_B + u_o - u_s$.
y_m	$\bar{u}_s - u_s + \ln 2.5$.
Y_s	Admittance of a single level interface state in mhos/cm^2.
Y_{ss}	Admittance of an interface state continuum in mhos/cm^2.
Z_s	MIS capacitor impedance less reactance of the oxide layer in ohms-cm^2.
α	Characteristic area, cm^2
β	$= 38.5\ q/kT$ in volt^{-1}.
Δ	Additive constant in volts.
δf	Time dependent part of Fermi function.
δn_s	Incremental electron surface-charge density induced by δv_a in cm^{-3}.

δv_a	ac signal applied to MIS capacitor in volts.
$\delta \psi_s$	Silicon band bending by ac perturbation in volts.
ϵ_{si} , ϵ_{ox}	Dielectric permittivity of silicon and SiO_2 layer respectively in farad/cm.
ξ	Distance into oxide measured from interface in cm.
σ_B	Standard deviation of surface potential caused by fluctuations of ionized acceptors.
σ_n , σ_p	Electron and hole capture cross sections in cm^2.
σ_s	Standard deviation of surface potential caused by fluctuations of surface charges.
σ_x	Standard deviation of surface potential caused by oxide thickness fluctuations.
τ	Majority carrier time constant of a single level interface state, $R_s C_s$ in seconds.
τ_m	Measured interface state time constant in seconds.
τ_{mo}	Time constant at $\bar{u}_s = u_B$.
$\tau_{n,s}$; $\tau_{p,s}$	Electron and hole time constant for a single level in seconds.
ψ	Energy in electron volts.
ψ_B	$(kT/q)\, u_B$ in volts.
ψ_s	Band bending or surface potential in volts.
$\bar{\psi}_s$	Mean surface potential in volts.
ψ_{se}	Surface potential plus an additive constant in volts.
ψ_{so}	Surface potential at the dc bias in volts.
ω	Angular frequency $= 2\pi f$ in sec^{-1}.

REFERENCES

1. Williams, R., Photoemission of Electrons from Silicon into Silicon Dioxide, Phys. Rev., *140,* 1965, A569.
2. Heiman, F. P. and Warfield, G., The Effects of Oxide Traps on the MOS Capacitance, IEEE Trans. Electron Devices, *ED-12,* 1965, p. 167.
3. Snow, E. H., Grove, A. S., Deal, B. E., and Sah, C. T., Ion Transport Phenomena in Insulating Films, J. Appl. Phys., *36,* 1965, p. 1664.
4. Goetzberger, A., Improved Properties of Silicon Dioxide Layers Grown Under Bias, J. Electrochem. Soc., *113,* 1966, p. 138.
5. Grove, A. S., Deal, B. E., Snow, E. H., and Sah, C. T., Investigation of Thermally Oxidized Silicon Surfaces Using Metal-Oxide-Semiconductor Structures, Solid State Electronics, *8,* 1964, p. 145.
6. Deal, B. E., Grove, A. S., and Snow, E. H., Characteristics of the Surface State Charge Q_{ss} of Thermally Oxidized Silicon, J. Electrochem. Soc., *114,* 1967, p. 266.
7. Terman, I. M., An Investigation of Surface States at a Silicon/Silicon Oxide Interface Employing Metal-Oxide-Silicon Diodes, Solid State Electronics, *5,* 1962, p. 285.
8. Zaininger, K. H. and Warfield, G., Limitations of the MOS Capacitance Method for the Determination of Semiconductor Surface Properties, IEEE Trans. Electron Devices, *ED-12,* 1965, p. 179.
9. Nicollian, E. H. and Goetzberger, A., MOS Conductance Technique for Measuring Surface State Parameters, Appl. Phys. Letters, *7,* 1965, p. 216.

10. Nicollian, E. H. and Goetzberger, A., Lateral AC Current Flow Model for Metal-Insulator-Semiconductor Capacitors, IEEE Trans. Electron Devices, *ED-12,* 1965, p. 108.
11. Hofstein, S. R. and Warfield, G., Physical Limitations on the Frequency Response of a Semiconductor Surface Inversion Layer, Solid-State Electronics, *8,* 1965, p. 321.
12. Lehovec, K. and Slobodskoy, A., Impedance of Semiconductor-Insulator-Metal Capacitors, Solid-State Electronics, *7,* 1964, p. 59.
13. Berz, F., Variation with Frequency of the Transverse Impedance of Semiconductor Surface Layers, J. Phys. Chem. Solids, *23,* 1962, p. 1795.
14. Goetzberger, A. and Nicollian, E. H., Temperature Dependence of Inversion Layer Frequency Response in Silicon, B.S.T.J., *46,* March, 1967, pp. 513-522.
15. Sah, C. T., Solid State Electronics Laboratory Technical Report No. 1, Section 16.1.4, Electrical Engineering Laboratories, University of Illinois, Urbana, Ill., 1964.
16. Lindner, R., Semiconductor Surface Varactor, B.S.T.J., *41,* May, 1962, pp. 803-831.
17. Moll, J. L., Variable Capacitance with Large Capacity Change, 1959 IRE Wescon Conv. Rec.—Part 3, Electron Devices, pp. 32-36.
18. Shockley, W. and Read, W. T., Statistics of the Recombination of Holes and Electrons, Phys. Rev., *87,* 1952, p. 835.
19. Frolich, H., *Theory of Dielectrics,* Oxford at the Clarendon Press, 1958, 2nd Ed., Chapter III.
20. Yager, W. A., The Distribution of Relaxation Times in Typical Dielectrics, Physics, *7,* 1936, p. 434.
21. Lehovec, K., Slobodskoy, A., and Sprague, J., Field Effect Capacitance Analysis of Surface States on Silicon, Phys. Status Solidi, *3,* 1964, p. 447.
22. Lehovec, K., Frequency Dependence of the Impedance of Distributed Surface States in MOS Structures, Appl. Phys. Letters, *8,* 1966, p. 48.
23. Shockley, W., Problems Related to p-n Junctions in Silicon, Solid-State Electronics, *2,* 1961, p. 35.
24. Davenport, W. B., Jr., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise,* McGraw-Hill Book Co., Inc., New York, 1958, Chapter 3.
25. Goetzberger, A., McDonald, B., Haitz, R. H., and Scarlett, R. M., Avalanche Effects in Silicon p-n Junctions. II. Structurally Perfect Junctions, J. Appl. Phys., *34,* 1963, p. 1591.
26. Cramer, H., *The Elements of Probability Theory,* John Wiley & Sons, New York, 1955, Chapter 6.
27. von Hippel, A. R., *Dielectrics and Waves,* John Wiley & Sons, New York, 1954, Appendix I.
28. Berglund, C. N., Surface States at Si-SiO$_2$ Interfaces, IEEE Trans. Electron Devices, *ED-13,* 1966, p. 701.
29. Balk, P., Effects of Hydrogen Annealing on Silicon Surfaces, Extended Abstracts, Electronics Div. Electrochem. Soc., *14,* No. 1, 1965, p. 237.
30. Gray, P. V. and Brown, D. M., Density of SiO$_2$-Si Interface States, Appl. Phys. Letters, *8,* 1966, p. 31.
31. Kuper, A. B. and Nicollian, E. H., Effect of Oxide Hydration on Surface Potential of Oxidized P-Type Silicon, J. Electrochem. Soc., *112,* 1965, p. 528.
32. Many, A., Goldstein, Y., and Grover, N. B., *Semiconductor Surfaces,* John Wiley and Sons, New York, 1965, Ch. 9.
33. Rzhanov, A. V. and Neizvestnyi, I. G., Soviet Physics—Solid State, *3,* 1962, p. 2408.
34. Kurskii, Yu. A., Soviet Physics—Solid State, *4,* 1963, p. 1922.
35. Crowell, C. R., Interpretation of Tunnel and Capacitance Measurements in the Presence of Dielectric Film-Thickness Fluctuations, Appl. Phys. Letters, *8,* 1966, p. 328.