Electron-hole recombination at the $Si\text{-}SiO_2$ interface

E. Yablonovitch
Bell Communications Research, Murray Hill, New Jersey 07974

R. M. Swanson and W. D. Eades
Stanford Electronics Laboratory, Stanford University, Stanford, California 94305

B. R. Weinberger
United Technologies Research Laboratory, East Hartford, Connecticut 06108

(Received 23 August 1985; accepted for publication 15 November 1985)

We have measured the surface recombination current $J(n_s,p_s)$ at high quality thermally grown $Si\text{-}SiO_2$ interfaces as a function of the surface density of electrons and holes, n_s and p_s. We find that the recombination is dominated by centers whose electron capture cross section is about 100 times greater than their hole capture cross section. Therefore, the maximum recombination occurs when $p_s \approx 100n_s$. Recombination is minimized under extreme electron or hole accumulation and is coincidentally the same in both cases: $\exp(qV/kT) \times 10^{-14}$ A/cm^2.

The thermally grown $Si\text{-}SiO_2$ interface is perhaps the dominant material system in modern electronics.[1] Numerous techniques, such as high frequency[2] and quasistatic[3] capacitance-voltage (C-V) methods, ac conductance,[4] and deep level transient spectroscopy[5] (DLTS), have been devoted to the characterization of the electronic properties of this interface. These generally measure majority-carrier properties. Nevertheless, very little is known about minority-carrier properties such as electron-hole recombination.

In this letter we provide a measurement of $J(n_s,p_s)$, the recombination current at the $Si\text{-}SiO_2$ interface as a function of the surface concentration of electrons and holes, n_s and p_s. We will find that J depends strongly on n_s/p_s and less strongly on n_sp_s; therefore, it helps to regard these as the independent variables, $J(n_s/p_s,n_sp_s)$.

We have employed a new experimental method as outlined in Fig. 1. The inductively coupled apparatus is similar to one which is used to make contactless[6] measurements of the minority-carrier lifetime in silicon wafers. The bulk carrier density injected by the strobe lamp is at or near high level injection $n_b \approx p_b$. (The bulk injection levels can be calibrated from the steady state response of the radio frequency bridge to wafers of known resistivity or by use of a calibrated silicon photodiode.) At high level injection, both positive and negative gate bias produce accumulation resulting in an enormous simplification to the band bending problem.[7] It is then an easy matter to accurately calculate both n_s and p_s and especially n_s/p_s, provided that a neutral starting point is known. This was provided by a C-V measurement of the flatband potential in the dark.

The only previous attempts[8,9] to measure J have employed a gate controlled diode or transistor structure in *low level* forward injection. The surface recombination then peaks at a gate bias near the crossover between inversion and depletion. In the vicinity of that crossover, the n_s/p_s ratio varies rapidly and is a *very* sensitive function of gate voltage and is consequently difficult to determine accurately. This has hampered previous attempts to measure $J(n_s/p_s)$ which have generally concluded[8,9] that J peaks at $n_s/p_s \approx 1$ which has led to the usual textbook assumption[7,10] that interface defects have nearly equal capture cross sections for electrons and holes. In striking contrast, we find that J is maximum at $p_s \approx 100n_s$, implying a correspondingly lopsided cross-section ratio. The main reason for our improved accuracy is the enormous simplification engendered in the band bending problem at *high level* injection.

The carrier density decay is monitored by the rf bridge whose output is averaged and displayed by a digital oscilloscope. When the sample thickness L is very thin the decay of excess bulk minority-carrier density n_b is the sum of a bulk and surface term[11]:

$$\frac{dn_b}{dt} = -\frac{n_b}{\tau_b} - \frac{2J}{qL}, \tag{1}$$

where τ_b is the bulk recombination lifetime, and the factor 2 accounts for the front and back surfaces. (The formulas in this article are valid as written for both high and moderate level injection relative to the bulk doping density.) The sur-

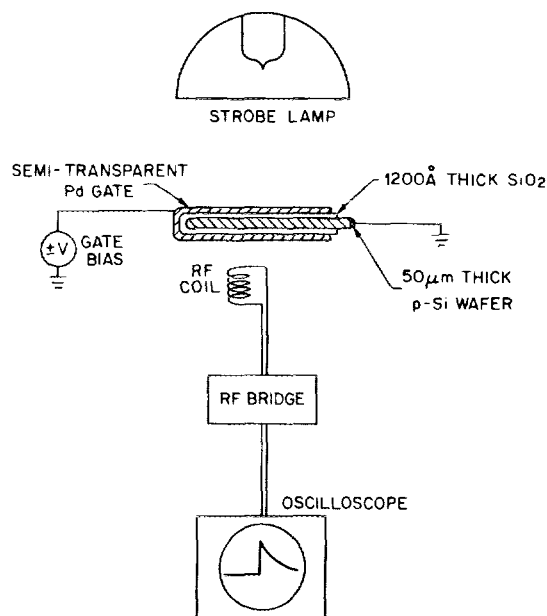

FIG. 1. Contactless minority-carrier lifetime measurement for *in situ* interface measurements.

face recombination velocity may be defined by $J \equiv qSn_b$. Then Eq. (1) can be rewritten

$$\frac{dn_b}{dt} = -\left(\frac{1}{\tau_b} + \frac{2S}{L}\right)n_b . \tag{2}$$

The experimentally measured quantity in brackets has been called the reciprocal filament lifetime.[11] In this work, we selected the highest quality float zone silicon to maximize τ_b and we reduced the silicon wafer thickness as low as $L = 50$ μm, to make the surface recombination velocity S the dominant term in Eq. (2). Notwithstanding the variable absorption depth of the white strobe light source, the injected carrier density n_b will be spatially uniform provided that $L \ll \sqrt{D\tau_b}$ and $L \ll D/S$ where D is the diffusion constant. (Note that S will be time dependent via its dependence on n_b, p_b.)

The thin silicon wafer was oxidized under standard conditions at 1030 °C in dry oxygen to a thickness of 1200 Å. During the initial stages of the oxidation, 0.2% HCl was present. After oxidation the samples were annealed in argon at 1030 °C and then in forming gas at 450 °C. Semitransparent palladium films were subsequently evaporated on both faces of the wafer to act as gate electrodes.

The experiment consisted of measuring the surface recombination velocity S as a function of gate bias which varied from -70 to $+70$ V. Since $S \equiv J(n_s, p_s)/qn_b$ is not purely a surface property but depends upon band bending via the bulk minority-carrier density n_b, it is useful to introduce the generalized surface recombination velocity $S_{gen} \equiv J(n_s, p_s)/q\sqrt{n_s p_s}$ which explicitly depends only on surface properties and is independent of band bending. The generalized surface recombination velocity S_{gen} may be related to the experimentally measured quantity by $S_{gen}\sqrt{n_s p_s} = Sn_b$. Under the assumption of quasi-equilibrium between the surface and the bulk, the quasi-Fermi level separation $V \equiv kT \ln(np/n_i^2) - \Delta$ is independent of position and therefore $n_s p_s = n_b p_b \exp[(\Delta_s - \Delta_b)/kT]$, where Δ_s, Δ_b are the surface and bulk band gap narrowing which occur if the carrier density becomes very high. Neglecting band gap narrowing for the moment, the fundamental surface property S_{gen} can be derived from experimentally measured S provided that the bulk injection ratio n_b/p_b is known: $S_{gen} = S \times \sqrt{n_b/p_b}$.

Based on the above discussion, the generalized surface recombination velocity S_{gen} was measured as a function of the n_s/p_s ratio. One example of such a measurement is plotted in Fig. 2. The curve is shaped like the toe of a right foot, peaking at $\ln(n_s/p_s) \approx -2$ (the big toe) but still having a significant contribution at $\ln(n_s/p_s) > 0$ (the smaller toes). When the thermal oxide preparation conditions were varied the general shape of the curve was unchanged but the absolute value of S_{gen} shifted vertically somewhat.

The standard kinetic model for electron-hole recombination in semiconductors is the Shockley–Read–Hall (SRH) formula.[12] Specializing to the case of surface defects

$$J\left(\frac{n_s}{p_s}, n_s p_s\right) = \sum_t \frac{qN_t v_{th} (\sigma_n \sigma_p n_s p_s)^{1/2}}{(\sigma_n n_s/\sigma_p p_s)^{1/2} + (\sigma_p p_s/\sigma_n n_s)^{1/2} + n_t(\sigma_n/\sigma_p n_s p_s)^{1/2} + p_t(\sigma_p + \sigma_n n_s p_s)^{1/2}}, \tag{3}$$

where N_t is the surface density of defects of type t per cm^2 and the index t is summed over all types of defects, v_{th} is the thermal velocity of carriers, σ_n and σ_p are the electron and hole capture cross sections respectively. Here $n_t = n_i \exp[(E_t - E_i)/kT]$ and $p_t = n_i \exp[(E_i - E_t)/kT]$, where n_i and E_i are the intrinsic density and energy of the semiconductor respectively and E_t is the energy level of the defect. The effect of the last two terms in the denominator of Eq. (3) is to greatly diminish the recombination contribution from any defects with energy near the band edges. Neglecting these defects for now and using Eq. (3), S_{gen} can be written as a sum of simple Lorentzian functions in the variable $\sqrt{n_s/p_s} \equiv x$:

$$S_{gen} = \sum_t \sqrt{\sigma_p \sigma_n} \frac{N_t v_{th} a_t x}{x^2 + a_t^2} . \tag{4}$$

where $a_t \equiv \sqrt{\sigma_p/\sigma_n}$ defines the peak of each Lorentzian which may be different for each type of defect t. Equation (4) is a weighted composite of real part Lorentzian functions $a_t x/(x^2 + a_t^2)$ from different types of defects, each of which is sharply peaked at a value of $x = a_t$ corresponding to its cross-section ratio. The experimental data in Fig. 2 may be fit by a composite of two such Lorentzian functions (the solid curves) with the dominant one peaking at

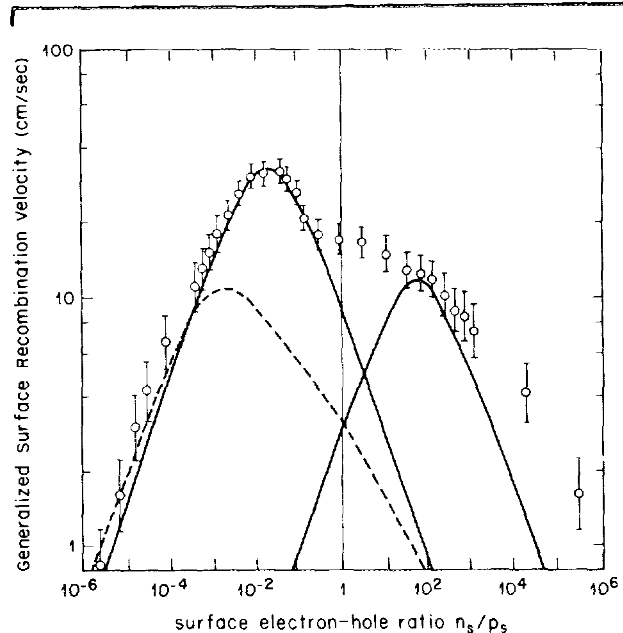

FIG. 2. Generalized surface recombination velocity S_{gen} as a function of n_s/p_s at a thermally grown Si-SiO_2 interface. The injection level was such that $n_b = 3 \times 10^{16}/cm^3$ and $p_b = 5.5 \times 10^{16}/cm^3$. The two solid curves are Lorentzian functions from Eq. (4) respresenting interface states of two different cross-section ratios σ_n/σ_p. The dashed line is a model based on DLTS measurements (Ref. 13).

$n/p = \sigma_p/\sigma_n = 0.01$ and a lesser one at $n/p = \sigma_p/\sigma_n = 80$. In that sense Fig. 2 is an experimental spectrum of interface defects weighted by number density N_t but plotted versus cross-section ratio.

This is not meant to imply that there are only two or a limited number of defect types. In practice there should be a continuum of cross-section ratios. Recent DLTS measurements[13] have accurately measured the electron capture cross sections in the upper half of the band gap and the hole capture cross sections in the lower half of the band gap. A model can be built by making an assumption about the behavior of the cross sections in the complementary half of the band gap where they are not known. The dashed line in Fig. 3 is such a model based on Ref. 13 in which the unknown capture cross sections were simply assumed to be 10^{-16} cm^2. The general preponderance of electron cross section over hole cross section is reproduced by that model, but the details are not necessarily well modeled. Previous work has also shown[14] that the cross-section ratio favors electrons for those defects in the very center of the gap.

The data points in Fig. 2 are measured at low or moderate gate bias. At large positive or negative bias a qualitatively different physical situation emerges. A large density of one carrier type is attracted to within an electrostatic screening length λ (Fermi–Thomas or Debye) of the Si-SiO_2 interface. This length is as short as 10 or 15 Å. The carrier density can become degenerate, as high as $10^{20}/cm^3$. Band gap narrowing, caused by screening which decreases the free energy cost of forming carriers, begins to play a major role. In this limit it makes sense to write $J \equiv J_0 \exp(qV/kT)$, where J_0 may be called the forward leakage current by analogy with the diode equation. This is derived experimentally from $J_0 = (qSn_i^2/p_b) \exp(\Delta_b/kT)$ which can be measured quite accurately. (The band gap narrowing in the bulk Δ_b is only a small correction, while at moderate injection level the majority-carrier density p_b exceeds the bulk doping level only slightly.)

Adapting the SRH model, Eq. (3), to the limit of extreme positive bias:

$$J_0 = \sum_t \frac{qN_t v_{\text{th}} \sigma_p n_i^2 \exp(\Delta_s/kT)}{n_s + n_t + (\sigma_p/\sigma_n) p_t}. \tag{5}$$

This predicts that J_0 should only be a function of n_s and not of p_s. Since that was indeed the case experimentally, we have plotted the measured J_0 vs n_s in Fig. 3(a). For extreme negative bias, we plotted versus p_s in Fig. 3(b).

The highest bias voltage produced the lowest value of J_0, which was coincidentally the same at 10^{-14} A/cm^2 in both limits. The approach toward the limiting value was rather different in the two cases. Under electron accumulation J_0 dropped near its final value as early as $n_s \approx 3 \times 10^{18}/cm^3$. While under hole accumulation J_0 was initially higher, dropping to its final value only after the hole density was quite degenerate ($> 10^{19}/cm^3$).

The theoretical curve[15] in Figs. 3(a) and 3(b) is a simplified form of Eq. (5) in which n_t and p_t are once again neglected: $J_0 \equiv S_n (n_i^2/n_s) \exp(\Delta_s/kT)$, where S_n is a constant representing a sum over all the defect states in the gap, $S_n = \Sigma N_t v_{\text{th}} \sigma_p$. We found $S_n = 330$ cm/s and the corresponding $S_p = 1080$ cm/s.

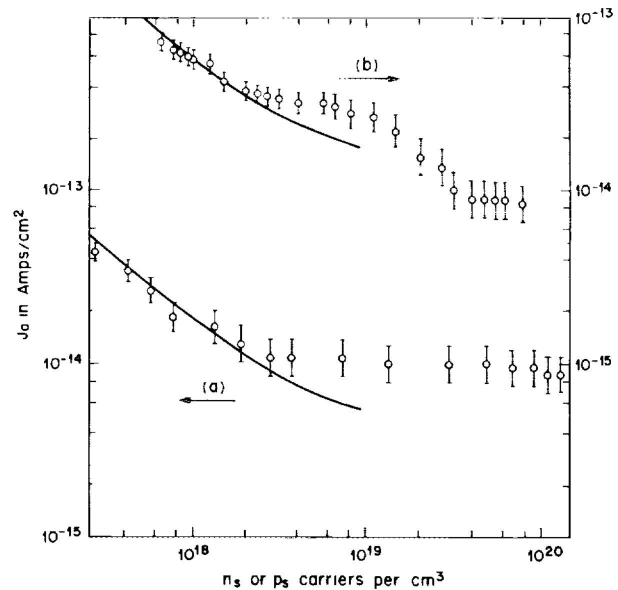

FIG. 3. Forward leakage current J_0 as a function of the carrier density accumulated at the Si-SiO_2 interface: (a) electrons; (b) holes. The solid curves are based upon a band gap narrowing model (Ref. 15).

The theoretical curve is a poor fit, underestimating J_0 at the higher densities. One reason is that the bandtail traps represented by n_t and p_t which were neglected, provide a relatively constant background contribution to J_0. There are many other possible reasons: a change in recombination kinetics in the degenerate limit, roughness at the Si-SiO_2 interface on the same scale as the screening length, inaccuracy of the band gap narrowing correction, etc. Auger recombination, however, seems to be too weak since the accumulation layer is so thin.

[1] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology* (Wiley, New York, 1982).

[2] L. M. Terman, Solid State Electron. **5**, 289 (1962).

[3] A. Goetzberger, E. Klausmann, and M. J. Schulz, CRC Crit. Rev. **6**, 1 (1976).

[4] E. H. Nicollian and A. Goetzberger, Bell System Tech. J. **46**, 1055 (1967).

[5] N. M. Johnson, J. Vac. Sci. Technol. **21**, 303 (1982).

[6] G. L. Miller, D. A. H. Robinson, and S. D. Ferris, Proc. Electrochem. Soc. **78-3**, 1 (1978).

[7] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), p. 386. See also page 368 for a solution of the band bending problem under accumulation.

[8] D. J. Fitzgerald and A. S. Grove, Surf. Sci. **9**, 347 (1968).

[9] M. W. Hillen and J. Holsbrink, Solid State Electron. **26**, 453 (1983).

[10] A. S. Grove, *Physics and Technology of Semiconductor Devices* (Wiley, New York, 1967), p. 301.

[11] W. Shockley, *Electrons and Holes in Semiconductors* (van Nostrand, New York, 1950), p. 318.

[12] A. Many, Y. Goldstein, and N. B. Grover, *Semiconductor Surfaces* (North-Holland, Amsterdam, 1965).

[13] W. D. Eades and R. M. Swanson, J. Appl. Phys. **56**, 1744 (1984).

[14] J. A. Cooper and R. J. Schwartz, Solid State Electron. **17**, 641 (1974).

[15] In this model we have taken $\Delta = q^2/2\lambda$ which is the differential free energy of correlation in a classical plasma. N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill, New York, 1973), p. 65. The value in H. P. D. Lanyon, IEEE Trans. Electron Devices **ED-26**, 1014 (1979) is higher by a factor of 3/2 since it does not include the entropy contribution.