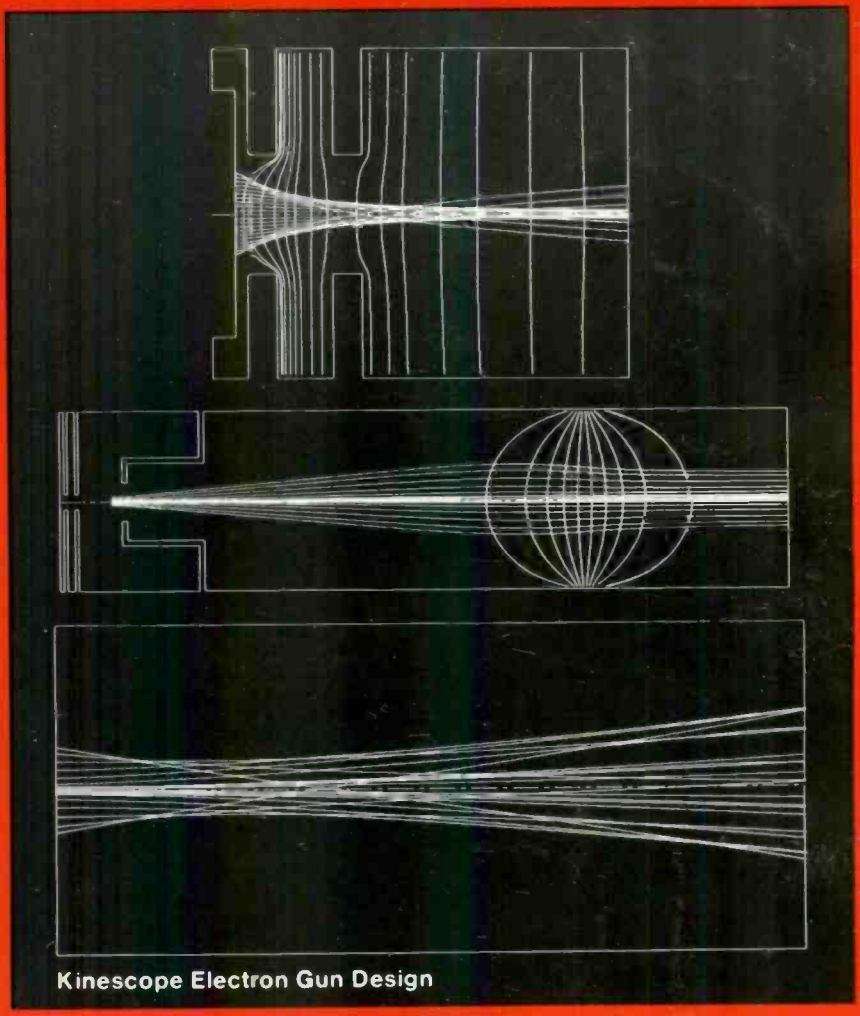
RCA Review



Kinescope Electron Gun Design

December 1980 Volume 41 No. 4

RCARCI 41 (4) 515-622 (1980)

RCA Review, published quarterly in March, June, September and December by RCA Research and Engineering, RCA Corporation, Princeton, New Jersey 08540. Entered as second class matter July 3, 1950 under the Act of March 3, 1879. Second-class postage paid at Princeton, New Jersey, and at additional mailing offices. Effective January 1, 1978, subscription rates as follows: United States and Canada: one year $8.00, two years $14.00, three years $18.00; in other countries, one year $8.60, two years $15.20, three years $19.80. Single copies (except for special issues) up to five years old $3.00.

The cover figure (Figure 4 of the first paper in this issue) shows computer calculations of selected equipotentials and principal electron trajectories in the beam-forming region, main lens, and drift region of a kinescope electron gun.

RCA Review

RCA Review (ISSN 0033-6831) is a technical journal published quarterly by RCA Research and Engineering in cooperation with the subsidiaries and divisions of RCA.

Contents

December 1980 Volume 41 Number 4 ISSN 0033-6831

Kinescope Electron Gun Design

R. Casanova Alig

RCA Laboratories, Princeton, NJ 08540

Abstract—The theoretical description of the undeflected electron beam in an electron gun of a TV kinescope is reviewed. The spot diameter at the screen D_s, is related to the beam diameter in the gun, D_b, by the equation $D_s = (C/D_b) + C_a D_b^3$. A theoretical expression for C is presented, together with the technology limitations that place a lower bound on C. However, this expression for C is incomplete. Complete values of C, denoted C', are calculated numerically for about 125 electron guns with $C_a = 0$ for the first time. The values of C'/C are found empirically to have a lower bound that depends on only the gun dimensions and voltages. Applications of this result to recent examples of electron gun design are described. Also, this result is used to design a gun for use with cathodes emitting electrons at high current densities.

1. The Electron Gun

In vacuum tubes that utilize an electron beam, the source of the beam is called an electron gun. These tubes include display tubes such as oscilloscopes and television picture tubes, as well as travelling wave tubes, electron microscopes, etc. The emphasis here will be on electron guns for television tubes; these tubes are frequently called kinescopes, meaning "to see action."

A cutaway drawing of the inside of a kinescope is shown in Fig. 1. It contains 3 beams, which excite the 3 color phosphors, for which 3 electron guns are required. They are located in the neck at the rear of the tube. Each gun is the source of an electron beam which is focussed on the screen. In a black-and-white kinescope, there is, of course, but one gun.

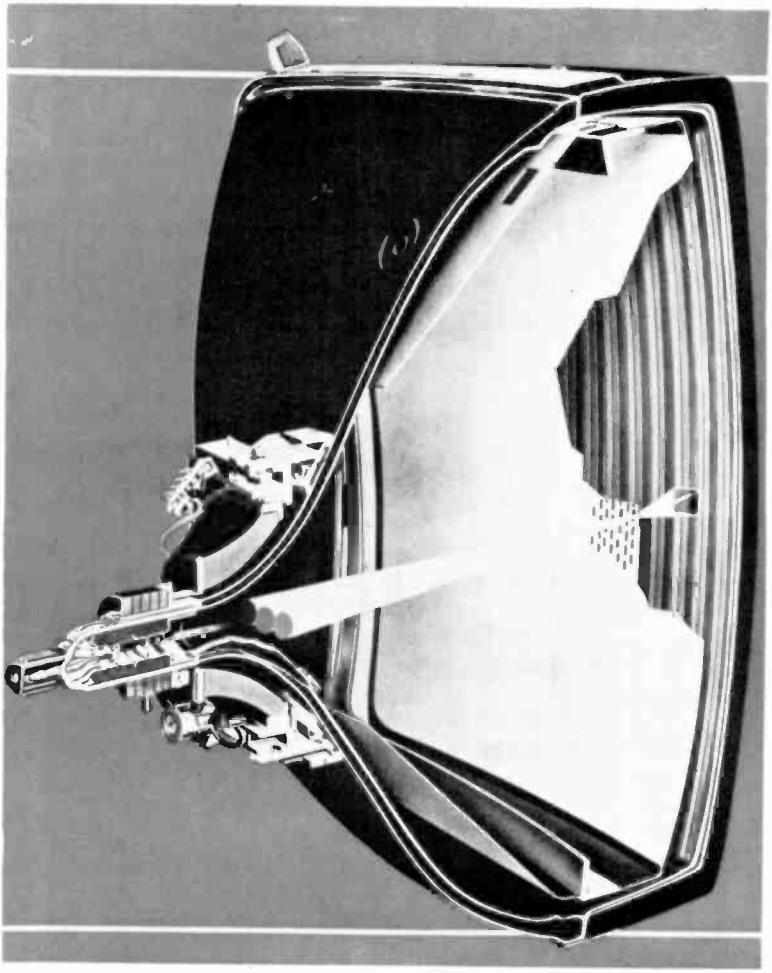


Fig. 1—A cutaway representation of a self-convergent shadow-mask color kinescope. Three electron beams originate from 3 electron guns located in the neck of the tube. The shadow mask is placed between the guns and the screen in such a way that each beam can strike only one color phosphor; for example, a light source defined by the red beam in the figure could only be seen from the red phosphor stripes. Although each beam is shown here to be focussed onto only one aperture in the shadow mask, the focussed spot of each beam generally covers several apertures. The yoke on the outside of the tube neck deflects the 3 beams all over the screen; the in-line arrangement of the guns shown here permits the use of a yoke which maintains the convergence, i.e., the coincidence of the 3 focussed spots, all over the screen. Additional description of the kinescope can be found in Refs. [7] and [9].

A cross-sectional drawing of one of these electron guns is shown in Fig. 2. The electrons are thermally emitted from the cathode, shown at the left of the figure. These electrons are fashioned into a narrow beam and focussed onto the screen by voltages on the electrodes, or grids, of the gun, shown to the right of the cathode. Four grids are shown here; they are labelled G_1, G_2, and G_3, and G_4 in accord with vacuum tube terminology. In addition to these labels, they are sometimes referred to as the control, screen, focus, and ultor* grids, respectively. These electrodes are generally cylindrically symmetric about the axis of each electron gun. Thus the first two grids shown in Fig. 2 are plates with circular apertures and the last two are cylindrical tubes. The electron gun of a kinescope is generally 1 to 3 inches long and 1 inch or less in diameter. Electrically, G_1 is about 100 V negative relative to the cathode. Positive voltages of about 500, 5000, and 25000 V are placed on G_2, G_3, and G_4 respectively.

How did the electron gun come to look this way? A priori, one would expect to need only the focus grid to draw the electrons from the cathode and the ultor grid, which, in combination with the focus grid, focusses the electrons to a spot on the screen. However, the focus grid may then intercept some of the electrons; this current in the focus grid not only wastes power but also alters the focus voltage. Thus, G_1 is added to direct the electrons into the entrance aperture of G_3. This addition was introduced by Wehnelt in 1903, very early in the development of electron guns, and G_1 is frequently called the Wehnelt electrode. In a kinescope, the voltage on G_1 also controls the current in the electron beam; thus G_1 is called the control grid. The cathode, labelled K, and G_1 and G_3 behave

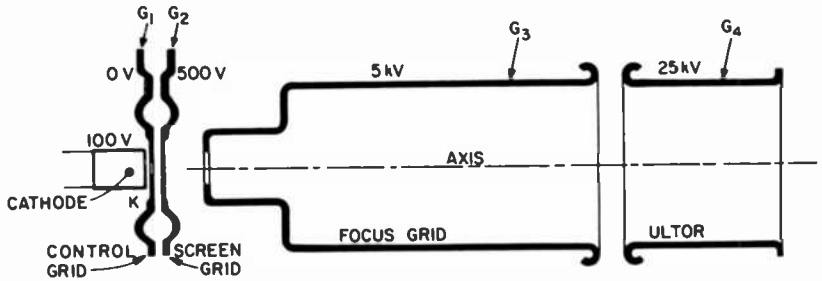


Fig. 2—A cross-sectional drawing of an electron gun similar to one of the 3 shown in Fig. 1. The gun is cylindrically symmetric about the axis. The elements are, from left to right, the cathode and the control, screen, focus, and ultor grids; these are labelled K, G_1, G_2, G_3, and G_4, respectively. The ribs on G_1 and G_2 provide for thermal expansion and the lips on G_3 and G_4 inhibit field emission.

* The ultor voltage is the largest voltage in a vacuum tube; it is also called the EHT, or extra-high tension. The last grid is called the ultor grid only if it is at the screen potential, as it is in a shadow-mask tube.

as the cathode, grid, and anode of a triode vacuum tube.[1] However, G_3 serves both as the anode of this triode and as part of the G_3-G_4 lens which focuses the beam onto the screen. Since a single voltage on G_3 cannot generally fulfill both these functions, the screen grid G_2 is added to screen G_3 from the cathode.

The drawing shown in Fig. 2 is redrawn in Fig. 3 with three axial distance scales. This was done to illustrate the properties of both the detailed structure near the cathode and of the open region between the gun and the screen. Several equipotential surfaces and electron trajectories are also illustrated in Fig. 3. The direction of the force on an electron is indicated at selected points by small arrows.

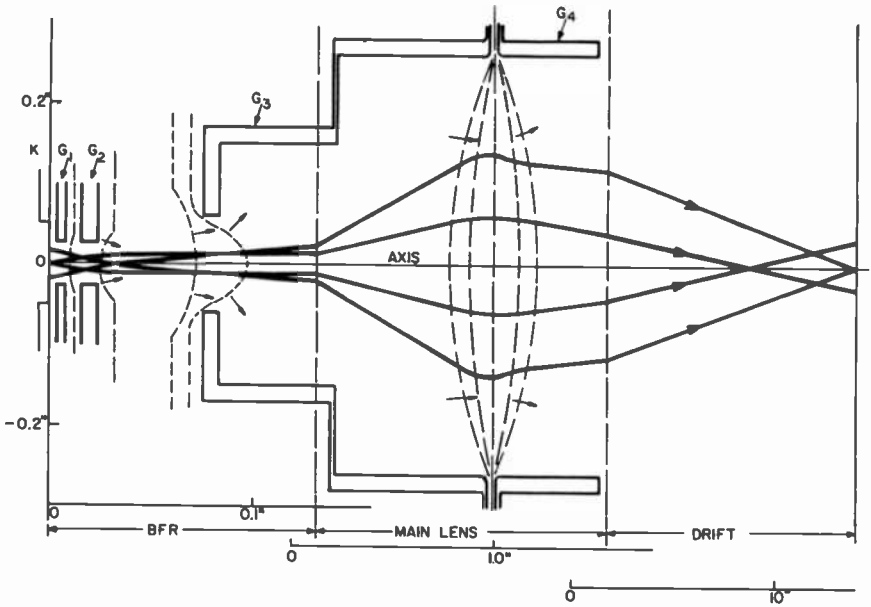


Fig. 3—A diagram of the drawing shown in Fig. 2 showing different portions on different axial scales. The left portion, containing K, G_1, G_2, and part of G_3, is called the beam-forming-region (BFR). The center portion, containing part of G_3 and G_4 is called the main lens; the axial scale is contracted by a factor 10. The right portion, containing the space between the gun and the screen shown in Fig. 1, is called the drift region; the axial scale is contracted by another factor 10. The radial scale applies to the entire diagram. Selected equipotentials are shown by the short-dashed lines, and the directions of the forces on electrons are indicated by the small arrows crossing these lines. Selected electron trajectories are indicated by solid lines originating on the cathode. The changes in their slope at the boundaries between the BFR, main lens, and drift region is illusory, due to the scale change. The trajectories originating normally from the cathode are called principals, and those originating obliquely are called thermals. The crossover is defined by the intersection of the two principals between G_2 and G_3.

Near the cathode the radial component of the force is directed toward the axis and the electrons emitted vertically from the cathode, called principals, are focussed to a small bundle on the axis, called the crossover; the electrons emitted with nonzero radial momenta, called thermals, give the crossover some finite size. The crossover is imaged on the screen by the lens formed between the focus and ultor grids. The similarities of the equipotentials in this region to spherical surfaces illustrates the similarity between electron-optical and light-optical lenses.

In the vicinities of the straight dashed lines shown in Fig. 3, the force on the electrons is very small so the electron trajectories are straight lines and the equipotential surfaces are nearly planes normal to the axis. Thus, these dashed lines define planes which conveniently divide the gun into regions that can be discussed individually. The first region, lying between the cathode and the central portion of the focus grid, is called the beam-forming-region (BFR); it is also called the triode or tetrode.[1] The second region, lying between the central portion of the focus grid and the end of the gun, is called the main lens. The third region, lying between the end of the gun and the screen, is called the drift region. There are no electrodes in this region, but space-charge forces between the electrons of the beam may be quite strong in this region. Computer-generated diagrams showing calculated equipotential planes and electron trajectories in each of these 3 regions are shown in Fig. 4.

2. Review of Theory

The science of electron-beam control is called electron optics, and mathematically there is a one-to-one correspondence between the refractive index of geometrical light optics and a function of the electrostatic potential Φ of electron optics. For motionless electrons at $\Phi = 0$, this function is approximately $(2e\Phi/mc^2)^{1/2}$. The theory of electron optics is described by Klemperer,[1] and by Ramberg in Zworykin, et al.[2] Its application to electron gun design is described by Moss[3,4] and Wilson.[5]

In the application of electron optics to electron-gun design, the goal is to make the display on the screen as sharp and bright as possible. This goal requires that the image size of the crossover at the screen center be minimized at all beam currents, subject to constraints imposed by other parts of the kinescope. There are three major limitations on the image size.

The first limitation is due to the emission of electrons from a cathode of nonzero size and temperature. This limitation follows directly from Abbe's sine law of light optics, i.e., $nr\sin\theta = n'r'\sin\theta'$, where n is the refractive index, and r and θ are the position and angle of the ray relative

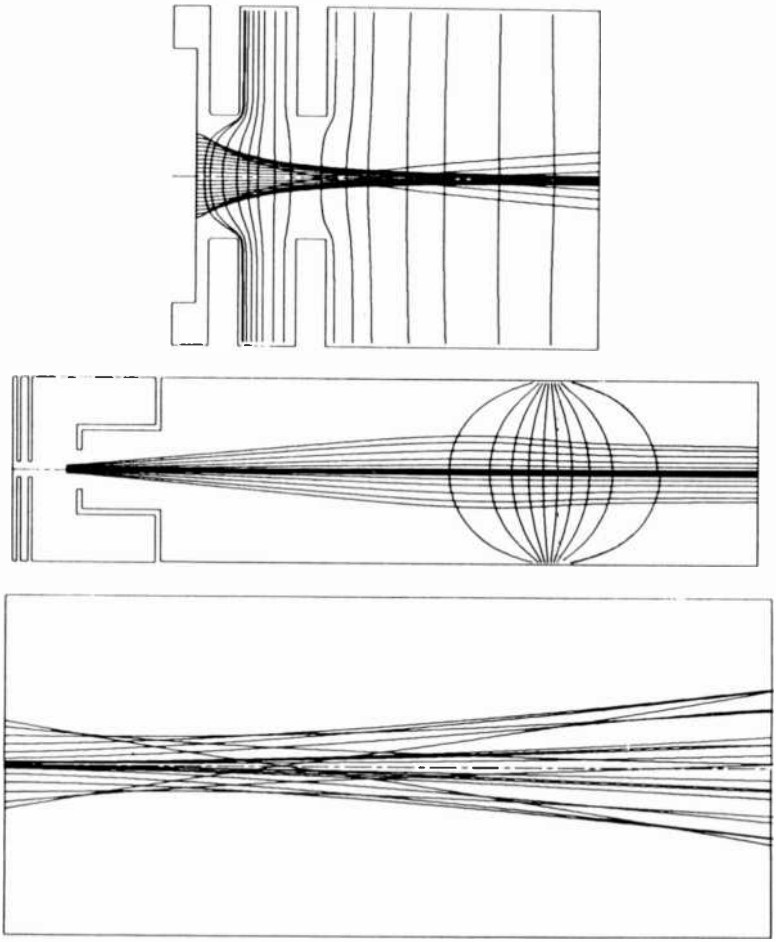


Fig. 4—Computer calculations of selected equipotentials and principal electron trajectories in the BFR, main lens, and drift region are plotted for an electron gun similar to that shown in Fig. 3. The calculations were done for a beam current of 3.5 mA. Due to the space-charge force, the interior trajectories in the drift region never cross the axis. This figure was prepared by F. Campbell.

to the optic axis. The primed and unprimed quantities refer to two planes normal to the axis; here these planes are at the cathode and screen. At the cathode, $n' \sim (kT)^{1/2}$ where T is the cathode temperature. For a ray on the boundary of the beam, $r' = d_c/2$ where d_c is the diameter of the emitting area, and $\sin\theta' = 1$ since electrons are emitted into a hemisphere about the axis. At the screen $n \sim (e\Phi)^{1/2}$, where Φ is the ultor voltage and $e\Phi \gg kT$. For this boundary ray, $r = D_s^t/2$ where D_s^t is the spot diameter on the screen, and $\sin\theta = D_b/2L$, where L is the length of the drift

region and D_b is the beam diameter in the main lens. Then Abbe's sine law becomes

$$D_s^t = \frac{2d_c L(kT/e\Phi)^{1/2}}{D_b} \equiv \frac{C_{th}}{D_b}; \qquad [1]$$

since this equation defines the spot size due to cathode size and temperature, it will be called the thermal magnification. Actually Eq. [1] is a statement of Liouville's theorem,[6] i.e., the conservation of volume in phase space, and it can be derived from thermodynamics.[1]

The second limitation is due to lens aberrations, the ubiquitous condition of lenses that rays from an object point are focused at different image points. In the absence of aberrations, the displacement from the axis in the image plane, r_i, of a ray is a linear function of its displacement r_0 in the object plane, and independent of its displacement r_a in the aperture plane of the lens; this is the Gaussian or paraxial ray approximation. For real lenses, however, r_i will be a function of higher powers of r_0 and r_a. Since r_i must change sign when both r_0 and r_a change sign, r_i is independent of even powers of r_0 and r_a. Thus, if r_i^0 is the displacement in the image plane in the Gaussian approximation, $r_i - r_i^0$ will, in lowest order, depend on a linear combination of r_0^3, $r_0^2 r_a$, $r_0 r_a^2$, and r_a^3. The aberrations associated with these terms are called, respectively, distortion, astigmatism and field curvature, coma, and spherical aberration (also called aperture defect). In an electron gun, r_0 is the cross-over diameter and r_a is the beam radius in the lens, $D_b/2$; since $D_b \gg r_0$, the dominant aberration in kinescope electron guns is spherical aberration. Thus the spot diameter on the screen D_s^a is

$$D_s^a = C_a D_b^3, \qquad [2]$$

where C_a is the aberration coefficient of the lens.[7]

The third limitation on the spot size is space charge, the mutual repulsion of the electrons due to their charge. Space charge, of course, has no optical analog. Space-charge effects are greatest at high beam currents and low electron velocities, and although they are omnipresent, they are most pronounced in the drift region because of its length and in the BFR because of the low electron velocities. Schwartz[8] derived a limitation on the diameter of the spot due to space charge in the drift region, D_s^s. In that derivation $D_s^s D_b$ is related to the parameter $[4\beta_0 IL^2/(D_b^2\Phi^{3/2})]^{1/2}$ as shown in Fig. 5; here I is the beam current and $\beta_0 = (m/2e)^{1/2}/2\pi\epsilon_0$ contains the electron charge and mass, e and m, together with the dielectric permittivity ϵ_0. Not only is this relationship complicated, but also the theory is greatly simplified if D_s^s and D_s^t have the same dependence on D_b. Thus, the Schwartz limitation was found to be closely approximated by

$$D_s^s = 0.8\beta_0 \left(\frac{IL^2}{D_b\Phi^{3/2}}\right) \equiv \frac{C_{sc}}{D_b} \qquad [3]$$

for beam currents from a few tenths to a few mA, typical of those used in kinescopes. This is illustrated in Fig. 5.

These limitations on the spot size must be combined to obtain a limitation D_s on the observed spot size. Moss[4] has suggested D_s^t and D_s^s combine quadratically, and Wilson[5] has suggested D_s^t and D_s^a combine linearly. For this discussion we shall utilize these suggestions to define*

$$D_s \equiv [(D_s^t)^2 + (D_s^s)^2]^{1/2} + D_s^a. \qquad [4]$$

Upon substitution of Eqs. [1], [2], and [3] into Eq. [4], one finds

$$D_s = \frac{C}{D_b} + C_a D_b^3 \qquad [5]$$

where

$$C = (C_{th}^2 + C_{sc}^2)^{1/2}. \qquad [6]$$

A plot of D_s versus D_b is shown in Fig. 6. The minimum of D_s falls at $D_b = (C/3C_a)^{1/4}$ where $D_s = (4/3)(3C_aC^3)^{1/4}$; at this value of D_b, 3/4 of D_s is due to thermal magnification and space charge and 1/4 is due to spherical aberrations.

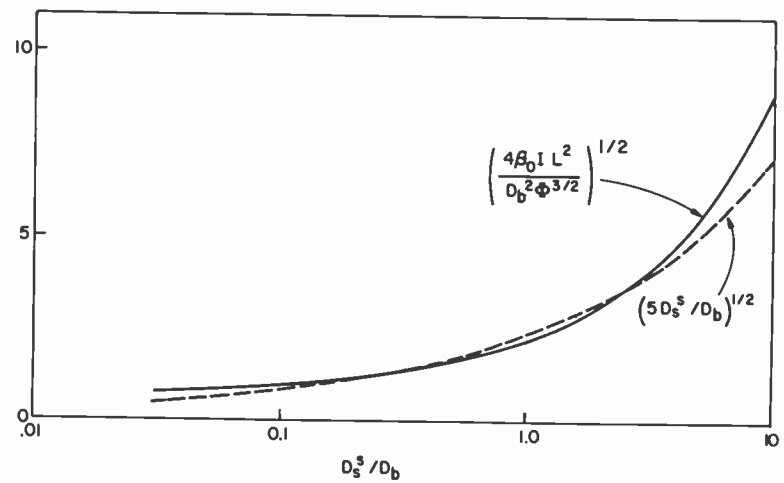


Fig. 5—The functions $f \equiv [4\beta_0 I L^2/(D_b^2\Phi^{3/2})]^{1/2}$, shown by the solid line, and $(5D_s^s/D_b)^{1/2}$, shown by the dashed line, are plotted versus D_s^s/D_b. Except for a scale change, the function f is identical to that shown in Fig. 2 of Ref. [8]. With $\beta_0 = 30345$ $V^{3/2}/A$ and values of L, Φ, and D_b frequently encountered in kinescopes for 0.5 mA $\leq I \leq$ 5 mA, f takes on values between 0.4 and 4.0. In this region f is seen to be closely approximated by $(5D_s^s/D_b)^{1/2}$ so these two functions were equated to obtain Eq. [3].

* Detailed analysis indicates this definition of D_s may not be exact; however, this analysis also indicates it is much simpler and it is a reasonable approximation to the exact result.

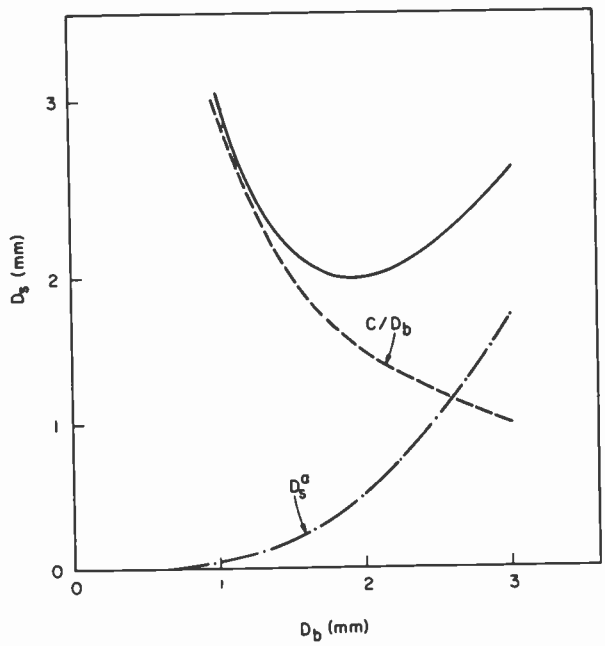


Fig. 6—The spot diameter D_s is plotted versus the beam diameter D_b following Eq. [5] for $C = 3$ mm^2 and $C_a = (16 \text{ mm}^2)^{-1}$. The dashed lines show the separate terms of Eq. [5]. The minimum value of D_s, 2 mm, falls at $D_b = 2$ mm; at this value of D_b, $D_s^a = 0.5$ mm and $C/D_b = 1.5$ mm.

3. Problem

To decrease D_s at a fixed I, the goal† of electron gun design for display tubes, the constants C and C_a in Eq. [5] should be made as small as possible. Reductions in C and C_a affect the display tube performance in different ways, however. In particular, a reduction in C_a decreases D_s, but it increases the value of D_b at which D_s is a minimum. An increase in D_b places additional burdens on other components of the tube. For example, the circle containing the cross-sections of all the beams in the tube is the object of the magnetic deflection lens generated by the yoke, just as the crossover is the object of the electrostatic focussing lens. In the same way as for the electrostatic lens, the aberrations of the magnetic lens increase with the object size; increasing D_b makes this object larger. Also in a shadow mask tube, two phosphor colors must be completely shadowed from each beam; that is, from a given phosphor stripe in Fig. 1, the beam from only one electron gun may be seen. As the beam cross-sections are enlarged by increasing D_b, the size of the apertures

† "The ultimate problem of design is to obtain the smallest and most brilliant electron spot with the minimum of required power for beam generation, modulation, and deflection," H. Moss, Ref. [4].

in the shadow mask must be reduced to ensure that two phosphor colors are fully shadowed from each beam.[7,9] This lowers the transmission of the mask which decreases the brilliance of the spot. Furthermore, in a multi-gun tube a larger beam cross-section forces a greater separation of the guns, with a resulting increase in the convergence force needed to superpose the beams at the screen center. In summary, a reduction in C_a is a useful improvement in a display tube only if the other components can accommodate a larger beam diameter. On the other hand, a reduction in C decreases both the minimum value of D_s and the value of D_b at which it occurs.

Inasmuch as d_c in Eq. [1] grows with I, both C_{th} and C_{sc}, which define C, vary directly with I and L and inversely with Φ. Since I is fixed, only L and Φ can be varied. To decrease L, the deflection angle must be increased; since the deflection angle defines the aperture size of the magnetic lens, this increases the yoke aberration. Nevertheless, the deflection angle of TV display tubes has grown with time.[9] An increase in Φ above 30 kV is accompanied by X-ray emission and browning of the faceplate.[9,10] Thus technological limitations appear to place a lower bound on C.

Can C be made smaller? No, but not all the contributions to the spot size are contained in Eqs. [1]–[4]. For example, Eq. [3] describes space-charge contributions in the drift region; there are additional space-charge contributions from the BFR where the size and velocity of the electron beam are small. Also, Eq. [2] describes lens aberrations only in the main lens; there are additional lens aberrations in the BFR due to the strong electric field gradients. These and other contributions to the spot size will be added to those contained in C to define a new constant C', which will be considered in this paper.

The determination of these contributions to C' is not trivial. In fact, Moss[3] in 1945 wrote:

"The basic theory of the electron gun is incredibly complex if account is taken of all the factors involved. Such an analysis, from first principles, would appear to be almost beyond the mathematics of our age, and would certainly yield no results of real practical value. This need cause no surprise when we reflect that such a theory would have to take into account space charge, the velocity distribution of the emitted electrons, and lens aberrations. Even when considered separately these factors involve weighty analysis."

4. Approach

The calculation of C' described here was done with a digital computer. The computer problem is defined by specifying the electrode geometry

and the potentials on the electrodes. Generally, potentials are specified on a surface enclosing the gun so the electrostatic potential in the interior of the gun can be calculated. This calculation is done by superposing a mesh on the interior and solving Laplace's equation, using the finite difference method of successive overrelaxation.[11] Advantage may frequently be taken of the cylindrical symmetry of the electron gun to solve the problem in only two dimensions. The electron trajectories are then calculated by numerical solution of the differential equation of motion. A large body of computer programs exists to solve Laplace's equation and evaluate electron trajectories.[12]

A current may be assigned to each trajectory using the results of planar diode theory,[13] e.g., Child's law,[1] and the cathode temperature. Once a set of these trajectories has been calculated, they define a space-charge density within the gun. With this charge density, new potentials in the gun may be found by solving Poisson's equation in the manner described in the preceding paragraph. New trajectories may then be calculated. When this procedure converges to a potential distribution and a set of electron trajectories, the calculation of the electron beam with space charge is complete.[14]

Many calculations[15] of C_a have been done in this manner, but none of C'. The calculation of C' was done with a digital computer program[16,17] of the type described above. The main lens was replaced by the mathematical transformation,

$$\begin{bmatrix} y_2 \\ y_2' \end{bmatrix} = \begin{bmatrix} A & B \\ D & E \end{bmatrix} \begin{bmatrix} y_1 \\ y_1' \end{bmatrix}, \tag{7}$$

obtained by following the rules for image transformation.[6] Since this transformation describes an ideal Gaussian lens, C_a is 0. Here y_1 and y_2 are the displacements, and y_1' and y_2' are the angles, of an electron trajectory at the entrance and exit planes of the lens measured relative to the axis. This transformation is illustrated in Fig. 7.

5. Calculations

Calculations were done on the IBM/370 computer for about 125 electron guns. The parameters of the calculations are shown in Fig. 8. The electrode thicknesses, aperture diameters, and separations are labeled t, d, and s, respectively, and subscripted with the grid numbers. Also shown are the cathode, grid, and screen voltages labelled V_K, V_1, V_2, V_3, and Φ, respectively; the G_1 electrode was grounded, i.e., $V_1 = 0$. In all the calculations I was fixed at 3.5 mA, L at 13.5 inches, and t_3 at 0.010 inch. The parameters varied in the calculations were t_1, t_2, d_1, d_2, d_3, s_{K1}, s_{12},

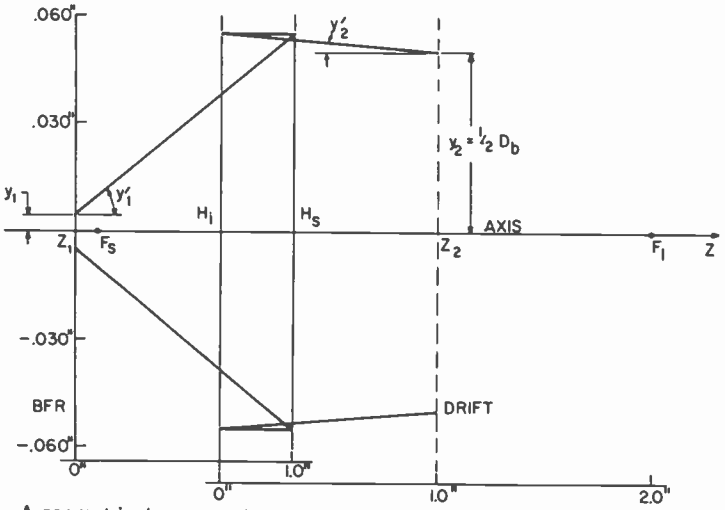


Fig. 7—A geometrical construction of the ideal lens which was used in place of the main lens. The lens is represented by the principal planes H_s and H_i and the focal points F_s and F_i on the lens axis z, as in geometrical optics. The entrace and exit planes are located at z_1 and z_2. The notation is that of Fig. 3.4 of Ref. [6]. The principal planes are reversed; this is shown in Ref. [2] to be a general characteristic of electron lenses. The radial and axial dimensions are typical of those used in the calculations. There are two axial scales because the separation of the principal planes is undefined by A, B, D, and E of Eq. [7]. This construction represents the transformation of y and y' given by $A = 0.5$, $B = 0.950$ inch, $D = -(2 \text{ inches})^{-1}$ and $E = -0.05$ in Eq. [7]; the potential at the exit plane is 5 times that at the entrance plane. The trajectories shown here lie near the edge of the beam. Although they appear to cross the axis 10 inches beyond the entrance to the drift region, space-charge forces will extend this distance.

s_{23}, V_K, V_3, and Φ; the extents of variation of these parameters were

$0.001'' \leq t_1 \leq 0.010''$ $\quad$ $0.003'' \leq t_2 \leq 0.026''$

$0.010'' \leq d_1 \leq 0.042''$ $\quad$ $0.010' \leq d_2 \leq 0.050''$

$0.024'' \leq d_3 \leq 0.060''$ $\quad$ $0.0015'' \leq s_{K1} \leq 0.005''$

$0.004'' \leq s_{12} \leq 0.035''$ $\quad$ $0.030'' \leq s_{23} \leq 0.110''$

$150 \text{ V} \leq V_K \leq 200 \text{ V}$ $\quad$ $2 \text{ kV} \leq V_3 \leq 30 \, kV$

and $18 \, kV \leq \Phi \leq 30 \, kV$

The steps of the calculation included the solution of Laplace's equation in the BFR, the calculation of the maximum V_2 needed to cut off current emission, reduction of V_K (usually by about 100 V) to obtain a 3.5-mA beam current, and calculation of the beam trajectories from the cathode

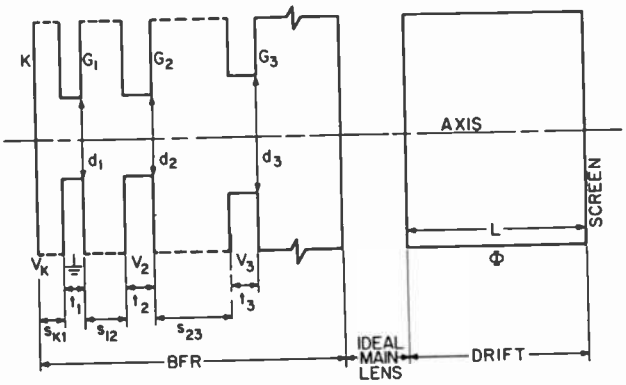


Fig. 8—The parameters of the electron gun used in the calculation of C' are shown in this schematic diagram. This diagram shows the cathode (K) and grids G_1, G_2, and G_3 of the BFR. The beam originates on the cathode, passes through the BFR, is focussed by the ideal lens transformation, and passes through the drift region to the screen. The electrode thicknesses, aperture diameters, and separations are labelled by *t*, *d*, and *s*, respectively. The length of the drift region is *L*. The G_1 electrode is grounded and positive voltages V_K, V_2, V_3, and Φ were applied to K, G_2 G_3, and the screen.

to the screen by iterative solution to Poisson's equation. A combination of the variables A, B, D, and E in Eq. [7] was found to minimize D_s with D_b near 0.080 inch, and each adjustment was accompanied by a calculation of the beam trajectories in the drift region by iterative solution to Poisson's equation. The calculation for each gun took about 15 minutes of central processor time.

The diameters D_s and D_b were defined as the full width of the beam current density function at 5% of its maximum value, and they were measured at the exit and entrance of the drift region. The constant C' is defined as the product D_sD_b. If C' were governed entirely by Eqs. [1] and [3], it would be a function of the gun parameters, independent of D_s and D_b. It was found to be approximately, but not fully, independent of D_s and D_b, indicating the existence of contributions to the spot size not contained in Eqs. [1] and [3].

Each calculation of C' was done for different sets of values for the 11 parameters mentioned above. While there is far too much data for a tabular presentation to be comprehensible, there is far too little data to fit it to a polynomial function of the variables; there is barely enough data to even define C' as a quadratic function of the parameters. Furthermore, the sets of values were chosen with no regard for randomness or for spanning the parameter space. Therefore, an empirical fit of the data to a selected function of the parameters was sought as a means to present the data.

Each value of C' was normalized with the value of C obtained* by using the gun parameters in Eq. [6]. The ratios r_c and r_f are defined empirically as

$$r_c \equiv (s_{K1} + t_1)/d_1$$

and

$$r_f \equiv \left[s_{23}^2 + \left(\frac{1}{2} d_3\right)^2\right]^{1/2} \Big/ d_1. \qquad [8]$$

A plot of C'/C versus $\sqrt{r}$, where

$$r = r_c r_f (\Phi^2/V_K V_3) \times 10^{-3}, \qquad [9]$$

is shown in Fig. 9. The coefficient of correlation of C'/C to $\sqrt{r}$ is 0.85.

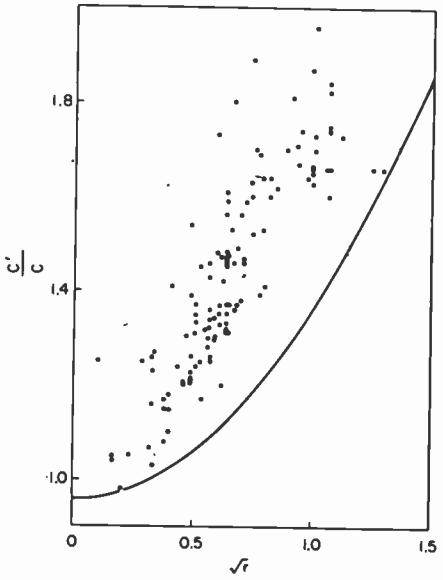


Fig. 9—A plot of the calculated values of C'/C versus $\sqrt{r}$. Each dot, which represents a value of C', was calculated numerically for a set of specific values of the gun parameters, as described in the text; the value of C was calculated from Eq. [6] (see footnote below) for the same set of parameters. This set of parameters also specifies a value of r from Eq. [9]. The curve, from Eq. [10], represents a lower bound on the calculated values of C'/C. Thus, for any set of gun parameters, a value of r is given by Eq. [9], and the smallest value of $D_s D_b$ that can be anticipated is then given by Eq. [10]. The guns with small values of r are, of course, frequently difficult to manufacture.

* Eq. [1] was evaluated with $kT = 0.1$ eV and $d_c = 0.72 d_1$; the latter relation was determined empirically from the calculations. The factor ln20 was included in C_{th} to correct for the beam width being defined at 0.05 instead of e^{-1} of the maximum beam current density. No factor was included in C_{sc} because the beam current density was assumed uniform in Ref. [8]. Thus, for example, with $d_1 = 0.025$ inch and $\Phi = 30$ kV, $C_{th} = 1.71$ mm^2 and $C_{sc} = 1.92$ mm^2.

This correlation is clearly apparent in Fig. 9, but it is even more notable that all the values of C'/C lie above the curve

$$C'/C = 0.96 + 0.40r. \quad [10]$$

For each set of gun parameters a value of V_2 was also calculated. Empirical relations between V_2 and these parameters have been described,[3] a recent one being,[18]

$$(V_2 - V_K)/V_K = 40r_c^2 r_s \quad [11]$$

where

$$r_s \equiv \left[s_{12}^2 + \left(\frac{1}{2}d_2\right)^2\right]^{1/2} \Big/ d_1. \quad [12]$$

The calculated values of $(V_2 - V_K)/V_K$ are plotted versus $r_c^2 r_s$ in Fig. 10; Eq. [11] is shown by the solid line. The coefficient of correlation is 0.88.

It is notable that r_s and t_2 do not appear in Eqs. [1] and [9]. Since the angular divergence of the beam exiting the BFR can be controlled with t_2, several different guns having a single value for each of C'/C, C, and C_a, and different values of D_b can be identified. Computer calculations of D_s and D_b for these guns can then be compared to the theory presented

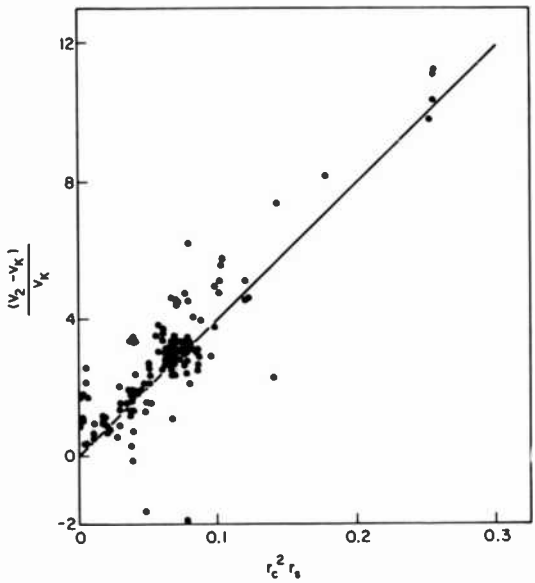


Fig. 10—A plot of $(V_2 - V_K)/V_K$ versus $r_c^2 r_s$. The values of V_2 were calculated from the solution of Laplace's equation for the BFR. The value of V_2 is the maximum voltage on G_2 that will drive electrons back into the cathode everywhere on the cathode surface.

above. The value $C' = 4.0\ \text{mm}^2$ was calculated using Eq. [7] for a BFR having $r = 0.21$ and $C = 3.1\ \text{mm}^2$. Then, with a real main lens having $C_a = (30.3\ \text{mm}^2)^{-1}$ instead of Eq. [7], values of D_s and D_b were calculated for several values of t_2. The results of these calculations are shown in Fig. 11 together with the theoretical curves obtained with Eq. [5], and with C' replacing C in Eq. [5]. The accord of the latter curve and the computed results is quite reasonable.

6. Applications

In some of the examples to be described, tradeoffs between C_A and C'/C will be used to minimize D_s. To relate C_A to the electron lens, a connection between C_A and the lens power S is required.† The coefficients

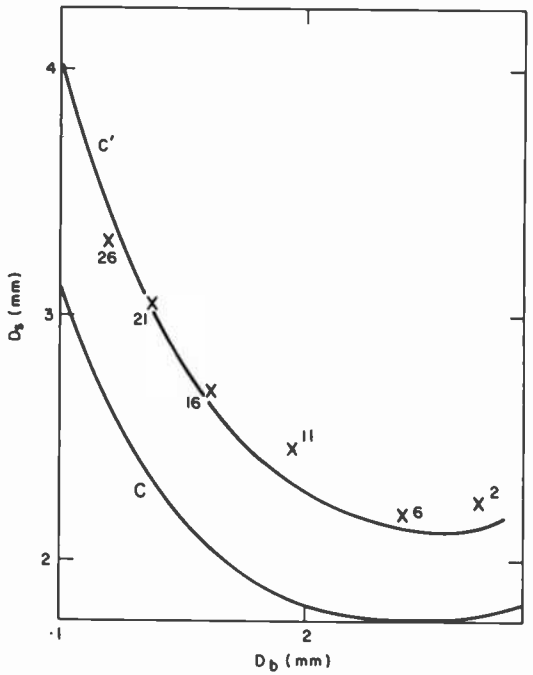


Fig. 11—A plot of D_s versus D_b for several electron guns differing only in their values of t_2. The solid curve labelled C was calculated with Eq. [5] and the solid curve labelled C' was calculated with C' replacing C in Eq. [5]; for both curves $C_a = (30.3\ \text{mm}^2)^{-1}$. The X's indicate calculations of D_s and D_b for guns with a real main lens having this value of C_a. The numbers with each X indicate the t_2 thickness in mils.

† The power of a lens is a measure of the distances of the object and the image from the lens. In light optics the power of a thin lens is the reciprocal focal length.

C_A and S are frequently defined in terms of the axial derivatives of the electrostatic potential on the lens axis.[2,15] These definitions are generally very complicated, as, for example, C_A contains a number of terms depending on the even-order derivatives. In a simple case, for the thin lens,[2] S depends on the first-order derivative. Thus the magnitudes of both S and C_A depend on the magnitudes of these axial derivatives, although they may, of course, depend on them in very different ways.

Since S must be nonzero for a lens to focus the crossover onto the screen, C_A will be nonzero. In a kinescope the values S may assume are further limited by the fixed distance between the crossover and the screen, and by the location of the lens much closer to the crossover than to the screen. Within these limitations, however, gun designs with values of C', C, and C_A that minimize D_s may be found. Yielding a slight increase in D_s from one of these parameters may offer a large decrease from another.

One approach to the minimization of D_s lies in placing the main lens as close as possible to the screen so that S may take on its smallest value for a given type of lens. Since both S and C_A depend on the axial derivatives of the electrostatic potential, C_A will presumably be made smaller. This approach was used in the recent development of the "high-voltage bipotential" electron gun.[17,19] In a bipotential gun the G_3 and G_4 electrodes form the main lens. An increase in the G_3 voltage V_3 weakens the lens and reduces[19] C_A; at the same time, by Eq. [10], C' is reduced due to the increase in V_3. Thus D_s is reduced because of reductions in both C_A and C'. This approach is limited by practical constraints on the length of the gun and the magnitude of V_3.

Another recent approach to the minimization of D_s has been the use of the extended-field lens (EFL).[20,21] In an EFL the abrupt voltage change across the gap between two lens electrodes, e.g., G_3 and G_4, is replaced by smaller voltage changes across gaps between several electrodes bridging the original gap. This, in effect, widens the original gap, something that is otherwise impractical in a kinescope due to the leakage of external fields into the gun, e.g., from charges on the neck of the tube or from adjacent guns in color tubes. The axial potential in an EFL varies quite slowly, so C_A is much reduced. However, S is also reduced somewhat, so if a monotonically increasing axial potential is used, simulating a bipotential, V_3 must be made small. As can be seen in Eq. [10], this increases C'/C. Thus it may be desirable to utilize a design with a large V_3, even at the expense of some increase in C_A, since C_A is so much reduced by the EFL design. This approach is described in Ref. [20] wherein a second screen electrode is placed between the screen and focus electrodes, and the potential on this electrode is intermediate between the focus and ultor voltages.

Another example of a tradeoff arises from the dependence of the spot size on d_c, the cathode diameter. This dependence is direct and inverse in Eqs. [1] and [10], respectively, since d_c is a direct function of d_1. Thus, for a given beam current, there is an optimum value of d_1 which, at the highest commonly used currents, is larger than many of the commonly used values of d_1. This result was confirmed in tests on several kinescope tubes.[22]

In kinescopes with commonly used guns operated at prevailing beam currents, this optimum value of d_1 exceeds the actual value only at the highest beam currents. However, when the emission current density from the cathode is increased, e.g., by making the cathode from other materials or operating it at higher temperatures, the same beam current is obtained from a smaller emitting area. Then, the optimum value of d_1 greatly exceeds the actual value for a substantial range of the higher beam currents, and much of the advantage in Eq. [1] due to the increased emission current density is lost. This advantage cannot be regained by reducing s_{K1} or t_1 in Eq. [8], as the values of these dimensions are already made as small as possible.

To regain this advantage, a platform cathode may be used. In this cathode the emitting surface is raised on a pedestal that extends partially into the G_1 aperture, as shown in Fig. 12. It somewhat resembles the Pierce electron gun,[1] which was developed for applications requiring large electron-beam currents. Since $s_{K1} + t_1$ in Eq. [8] is the distance from the cathode surface to the front surface of G_1, the parameter r_c is much reduced in the platform cathode. Since r_f is unchanged, C' is much reduced and the optimum value of d_1 at high currents can be near the

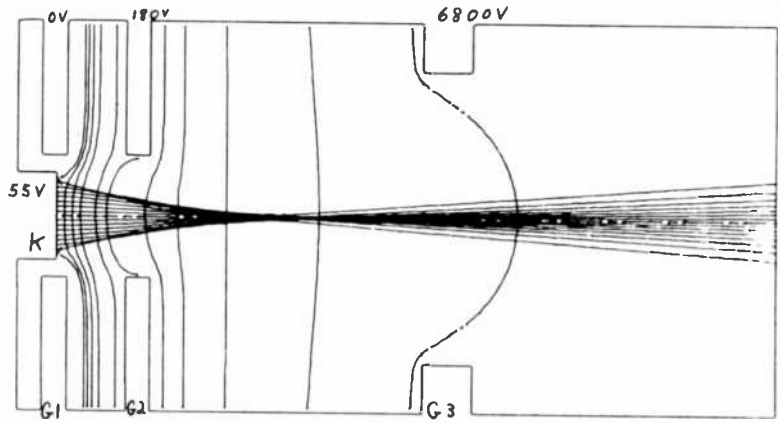


Fig. 12—A sketch of the BFR for the platform cathode showing selected equipotentials and electron trajectories, similar to that shown in Fig. 4 for a planar cathode.

actual value. At low beam currents, however, only the central portion of the cathode emits because the cathode surface is recessed inside the G_1 aperture, unlike the Pierce electron gun. Thus, at low currents, where C_{sc} in Eq. [3] is small, d_c in Eq. [1] is small, so C is small. These results were confirmed in tests on several kinescope tubes.[23]

7. Conclusion

In the preface of his book, Moss[4] observes:

> "The author still retains rather painful memories of his own initiation into these matters. In 1939, fresh from the University, he joined an industrial team engaged in the design of a wide variety of cathode ray tubes. Obviously here was a wonderful opportunity to mount the White Charger and thereby bring the analytic beauty of electron optics to bear on empirical procedures. Alas! after about a year, the White Charger had vanished, Empirical Procedures remained in firm possession of the battlefield, and a much chastened Hilary Moss might have been observed contemplating less ambitious adventures."

He goes on to suggest that the introduction of the digital computer will link gun design to the theory of electron optics.

An example of the introduction of the digital computer to electron gun design has been described in the preceding sections. When the computer calculations of C' were done, however, the results had to be fit to an empirical formula to make them comprehensible. Thus, once again, "White Charger has vanished, and Empirical Procedures remain." The only difference lies in the use of the computer, instead of the laboratory, to generate the data required to obtain empirical formulae, such as Eqs. [10] and [11]. The difficulty, both now and in 1939, lies in the profusion of parameters required to describe the electron gun.

Acknowledgments

The author is deeply indebted to F. J. Campbell, who developed much of the computer program used in these calculations, and to him and D. A. de Wolf for many of the electron-gun calculations. He is indebted to them, and to R. W. Cohen, D. P. Bortfeld, H. E. McCandless, C. H. Meltzer, H. E. P. Schade, and L. J. Vieland for many helpful discussions. He is also indebted to them, and to S. Bloom, C. W. Struck, and P. J. Wojtowicz for reviewing this paper.

References:

[1] O. Klemperer, *Electron Optics*, 3rd Ed., in collaboration with M. E. Barnett, Cambridge, London (1971).

[2] V. K. Zworykin, G. A. Morton, E. G. Ramberg, J. Hillier, and A. W. Vance, *Electron Optics and the Electron Microscope*, John Wiley, N.Y. (1945).

[3] H. Moss, "The Electron Gun of the Cathode Ray Tube—Part I," *J. Brit. Inst. Radio Eng.*, **5**, p. 10 (1945); "Engineering Methods in the Design of the Cathode Ray Tube," **5**, p. 204 (1945); and "The Electron Gun of the Cathode Ray Tube—Part II," **6**, p. 99 (1946).

[4] H. Moss, *Narrow Angle Electron Guns and Cathode Ray Tubes*, Academic Press, N.Y. (1968).

[5] I. M. Wilson, "Theoretical and Practical Aspects of Electron Gun Design for Color Picture Tubes," *IEEE Trans. Consumer Electron.*, **CE-21**, p. 32 (1975).

[6] P. Dahl, *Introduction to Electron and Ion Optics*, Academic Press, N.Y. (1973).

[7] V. K. Zworykin and G. A. Morton, *Television*, 2nd Ed., John Wiley, N.Y. (1954); P. W. Hawkes, "On the Nature of Spherical Aberration," *Amer. J. Phys.*, **46**, p. 433 (1978).

[8] J. W. Schwartz, "Space-Charge Limitations on the Focus of Electron Beams," *RCA Rev.*, **18**, p. 3 (1957).

[9] A. M. Morrell, H. B. Law, E. G. Ramberg, and E. W. Herold, *Color Television Picture Tubes*, Academic Press, N.Y. (1974).

[10] C. B. Braestrup and R. T. Mooney, "X-Ray Emission from Television Sets," *Science*, **130**, p. 1071 (1959).

[11] T. Saito, S. Kubota, and S. Miyaoka, "Digital Computer Studies of Electron Lenses for Cathode Ray Tubes," *J. Appl. Phys.*, **44**, p. 4505 (1973); M. Distasio and W. C. McHarris, "Electrostatic Problems? Relax!" *Am. J. Phys.*, **47**, p. 440 (1979).

[12] P. W. Hawkes, "Compilation of a Catalogue of Computer Programs in Electron Optics," *Computer Phys. Commun.*, **5**, p. 399 (1973); W. B. Herrmannsfeldt, *Electron Trajectory Program*, SLAC Report No. 166 (U.S. Dept. Commerce, 5285 Port Royal Road, Springfield, Va., 1973).

[13] P. T. Kirstein, G. S. Kino, and W. E. Waters, *Space-Charge Flow*, McGraw Hill, N.Y. (1967).

[14] D. L. Say, R. B. Jaeger, and J. D. Ryan, "Electron Optics for CRT's: Computer Modeling," *IEEE Trans. Consumer Electron.*, **CE-21**, p. 57 (1975).

[15] E. Harting and F. H. Read, *Electrostatic Lenses*, Elsevier, N.Y. (1976); T. Saito and O. J. Sovers, "Numerical Calculations of Focal and Aberration Properties of Unipotential Electron Lenses," *J. Appl. Phys.*, **48**, p. 2306 (1977); and, "Spherical Aberration Coefficients of Double-Cylinder Electrostatic Lenses Over Wide Magnification Ranges," *J. Appl. Phys.*, **50**, p. 3050 (1979).

[16] F. J. Campbell, private communication; this program is described in Ref. (17).

[17] R. H. Hughes and H. Y. Chen, "A Novel High-Voltage Bipotential CRT Gun Design," *IEEE Trans. Consumer Electron.*, **CE-25**, p. 185 (1979).

[18] R. Stone, private communication; for grid drive the left-hand side of Eq. [11] is V_2/V_K.

[19] D. L. Say, "The 'High Voltage Bipotential' Approach to Enhanced Color Tube Performance," *IEEE Trans. Consumer Electron.*, **CE-24**, p. 75 (1978).

[20] A. P. Blacker, I. M. Wilson, and J. W. Schwartz, "A New Form of Extended Field Lens for Use in Color Television Picture Tube Guns," *IEEE Trans. Consumer Electron.*, **CE-22**, p. 238 (1976).

[21] J. W. Schwartz and I. M. Wilson, "Extended Field Electron Lenses," *1975 IEDM Digest*, p. 269.

[22] H. Y. Chen, private communication.

[23] H. E. McCandless and C. H. Meltzer, private communication.

Evaluation of CMOS Transistor Related Design Rules

Alfred C. Ipri

RCA Laboratories, Princeton, NJ 08540

Abstract—In order to evaluate integrated circuit design rules that are specifically related to the MOS structure, a transistor array (TRAY) was designed. Fabricating and testing this array has produced optimum rules for the fabrication of LSI circuits using the specific process evaluated. It is seen that defects can cause either functional failures or increased array leakage and that, depending upon the leakage sensitivity of the integrated circuit, either may cause loss of circuit functionality.

1. Introduction

The use of various test patterns to evaluate individual aspects of MOS processing such as polysilicon continuity, metal step coverage, and contact opening integrity has been discussed previously.[1] The results of the previous evaluation show that meaningful yield curves can be generated to aid in the IC design and layout, and to maximize the number of working chips per wafer.

A number of design rules, however, involve mask level-to-mask level interactions and, hence, a test vehicle specifically oriented towards the MOS transistor is needed. The transistor array, or TRAY, described here was designed for this purpose. It includes many of the features of the previous test arrays, such as the use of standard cells that are serially connected to form strings of ever increasing numbers of cells. Simple dc testing is used to interrogate the strings for short or open circuits. The yield of functional arrays is then plotted vs. the number of transistors.

2. Transistor Array (TRAY)

The standard cell used for design rule analysis in this array is the CMOS inverter. The advantage of the CMOS inverter is its ability to monitor both open and short circuits that may occur in the transistor structure. For example, if a break occurs in the polysilicon gate of an MOS device over the channel region, a source-to-drain short can result. On the other hand, if a break occurs in the polysilicon away from the channel region, an open circuit can result. The inverter configuration is the simplest element whose lack of functionality reflects both open and short circuits. Previous test vehicles have incorporated either series connected MOS transistors[1] or transistors connected in parallel.[2] The series technique can easily monitor open devices but not shorted devices, while the parallel approaches test for shorts but not opens. The inverter configuration is, in effect, a combination of both the series and parallel approaches since the inverters are connected in parallel (between V_{DD} and V_{cc}) while the input signal is conveyed serially from the input to the output of the series connected string.

Fig. 1 shows the basic transistor inverter building block. The output of each inverter is connected to the input of the next inverter to form strings of inverters containing 200, 400, 600, 1200, 2400, 4800, and 9600 cells per string. Each string has independent input, output, V_{DD}, and V_{cc} connections. The total number of inverters in each chip is 19,200. Also included on each chip are individual test transistors and resistors for obtaining parametric data. The array measures 6.96 × 6.45 mm and, hence, about 100 chips are contained on a 3 inch diameter wafer.

The testing procedure involves the use of a computer controlled parametric IC tester. Each inverter string is powered on at 5 V (V_{DD}) and ground (V_{cc}). The input to the first inverter in the string is grounded and the output signal of the last inverter is measured, as well as the standby or leakage current of the string. The input is then switched high (5 V) and the output and leakage current again measured. An inverter string is considered functional if the difference between the two output signals is greater than 4.0 V. The leakage current is only cataloged on those strings that are considered operational. A statistical compilation is made for each wafer giving the number of operational strings as a function of the string length. Leakage current histograms are also generated for each string and each wafer. Yield curves are then plotted showing the variation in functionality, with both the number of inverters and the specific design rule being varied. Each set of wafers contains a specific design rule variable. It is possible, therefore, to determine an optimum and minimum set of design rules that are specifically related to two-level mask interactions.

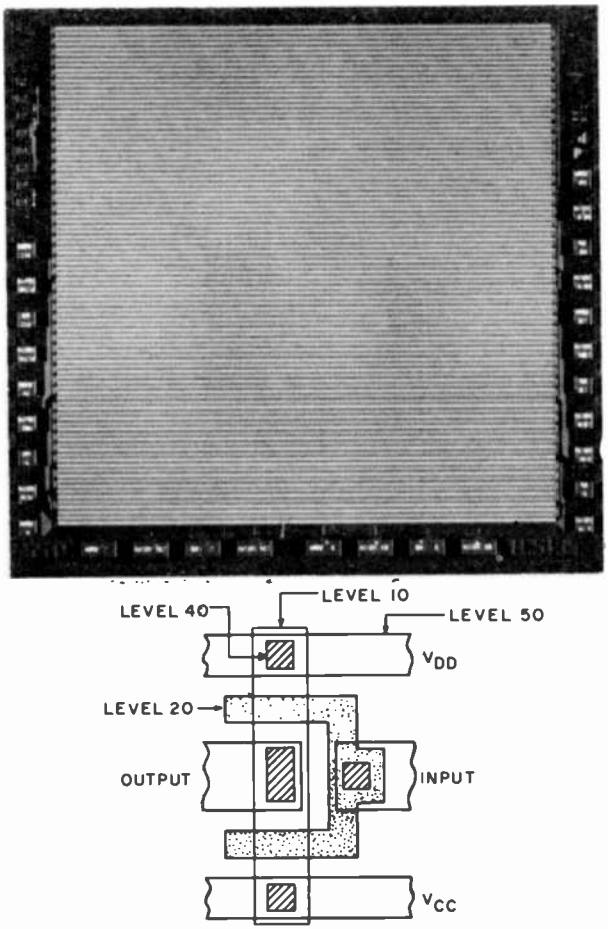


Fig. 1—Photomicrograph of the TRAY and the inverter cell.

The technology used to fabricate the transistor array was the CMOS/SOS process. Undoped silicon on sapphire films, 0.6-μm thick, were grown in the usual manner.[3] The silicon islands were defined using an anisotropic etchant (KOH, n-propanol and water) and doped using ion implantation. Channel oxides were grown in HCl steam at 900°C before deposition of the LPCVD polysilicon layer. Anisotropic wet etching was also used to define the polysilicon. The sources and drains were doped with boron and phosphorus by ion implantation using doses of 4×10^{15} cm^{-1} and 2×10^{15} cm^{-2}, respectively. A low temperature phosphorus glass was deposited and then flowed over the surface at 1050°C. Contact openings were then chemically etched. Using an In-

source® deposition system, 1.5 μm aluminum was evaporated over the substrates and chemically defined. The process, therefore, contained seven mask steps through metal definition.

3. Experimental Results

The TRAY was used to study such CMOS/SOS design rules as:

(a) epi-silicon to polysilicon spacing,
(b) poly-gate overlap of the silicon island,
(c) source-drain spacing,
(d) contact opening to epi-silicon edge,
(e) contact to poly-gate spacing, and
(f) N+ and P+ overlap of epi-silicon island.

The Tray mask set includes 8 epi-silicon definition levels, 11 polysilicon levels, 4 source-drain doping levels, 2 contact levels, and an aluminum definition level, for a total of 26 levels. A particular set of levels is used to analyse a particular design rule. To study the epi-silicon to polysilicon spacing, for instance, several wafers are processed using each of three different epi-silicon mask levels (levels 15, 16, 17). The same poly mask, however, is used on all wafers and, hence, the epi-to-poly spacing can be set to values of 2.5 μm, 1.0 μm, and 0.0 μm. A complete list of the various mask levels and their respective dimensions is given in Appendix 1. It should be noted that the functionality of the inverter structure is relatively insensitive to variations in device threshold, breakdowns, and leakage currents. Its functionality is designed to test for catastrophic failure mechanisms such as opens or shorts, which in turn, are defect related.

3.1 Epi-Silicon to Polysilicon Spacing

Fig. 2 is a plot of operational yield as a function of the number of inverters tested. Data were obtained for spacings of 2.5 μm, 1.2 μm, and 0, or line-to-line. As can be seen from the figure there is essentially no difference in functional yield. Fig. 3, however, shows a series of leakage histograms for inverter strings containing 9600 inverters (19,200 transistors). All three histograms show average leakage currents of about 3–4 μA. The only dependence on poly-to-epi spacing appears to be the population at the 100 μA level. As the number of poly-to-epi bridges increases, therefore, the number of arrays having leakage currents in

® Applied Materials Inc., Santa Clara, California.

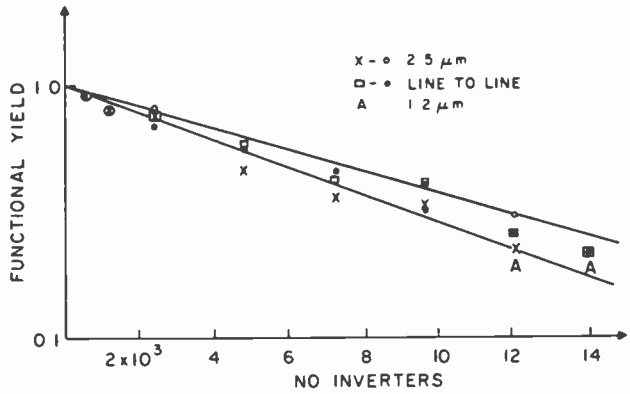


Fig. 2—Functional yield as a function of the number of inverters and the epi-silicon to polysilicon spacing

excess of 30 μA also increases. From an initial value of less then 6% of the total when wide (2.5 μm) spacings are used, the number of high leakage arrays increases to about 25% of the total when heavy bridging is present. The conclusion, therefore, is that poly-to-epi bridging can influence the yield of LSI arrays if they are leakage current sensitive. the total when wide (2.5 μm) spacings are used, the number of high leakage arrays increases to about 25% of the total when heavy bridging is present. The conclusion, therefore, is that poly-to-epi bridging can influence the yield of LSI arrays if they are leakage current sensitive.

3.2 Poly-Gate Overlap of the Epi-Silicon Island

Fig. 4 shows the results of varying the poly-gate overlap of the silicon island. It is seen that there is relatively little decrease in yield when using a value of 3–4 μm. Optimization calculations[1] show that the decrease in yield when using a value of 4.0 μm instead of 5.0 μm is more than compensated for by the increase in packing density and, hence, is the preferable value. Fig. 5 shows the effect of gate overlap on leakage currents. The average leakage value and the width of the overall distribution increases as the overlap is reduced. The data shown are for strings of inverters with 9600 cells. It is felt that the optimum gate overlap value is about 3–4 μm.

3.3 Source-Drain Spacing

Fig. 6 shows the functional yield dependence on channel length. It is seen

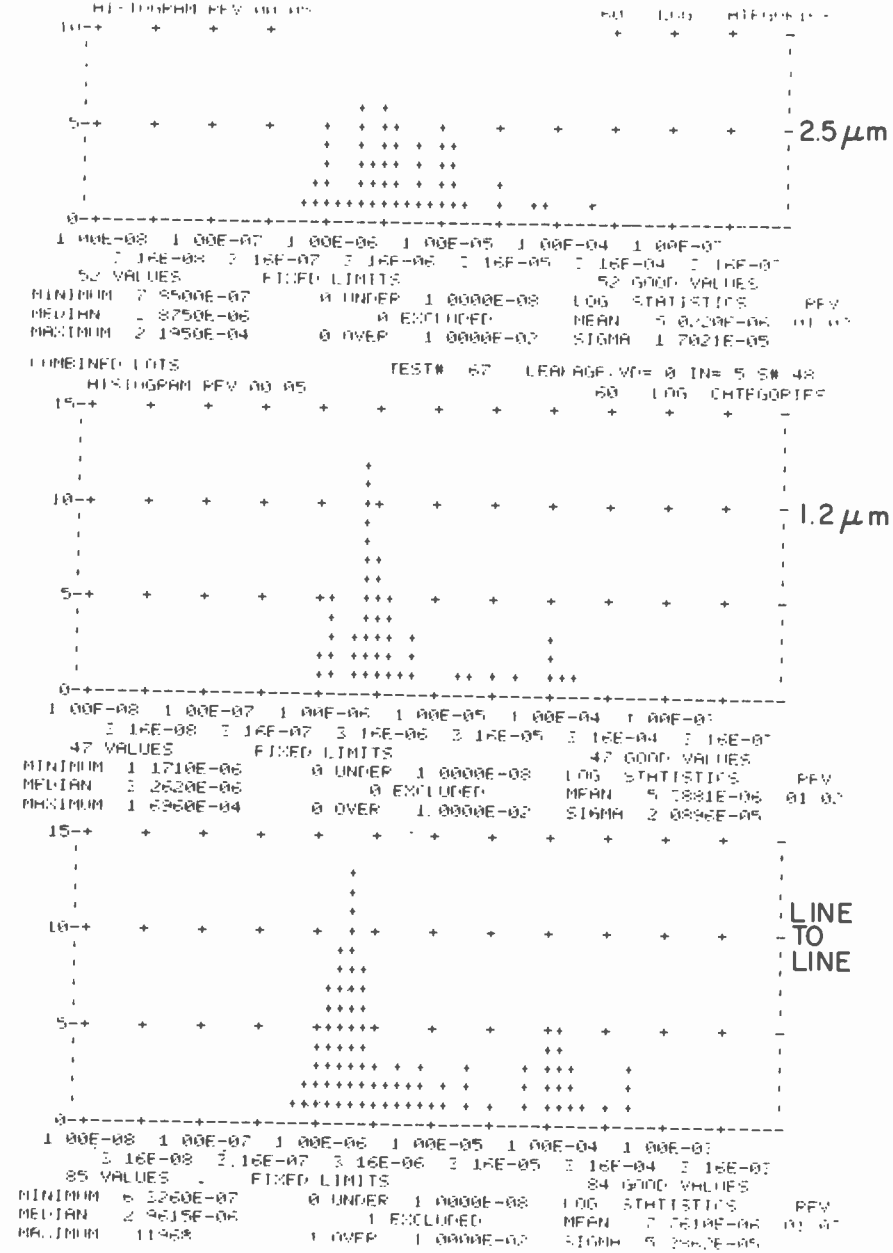


Fig. 3—Leakage current distributions as a function of epi-silicon to polysilicon spacing.

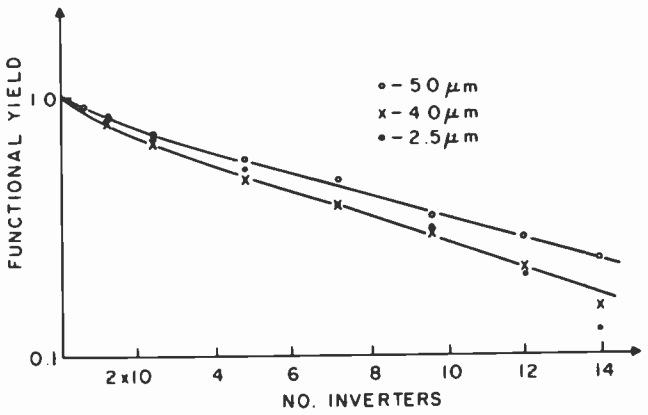


Fig. 4—Function yield as a function of the poly gate overlap dimension.

that the yield drops off substantially for channel lengths of 2.5 μm. Parametric effects such as substrate doping can also play a part in functional yield reduction as the channel length is reduced. Examination of leakage current data (Fig. 7), however, shows little increase in the leakage distribution as the channel length is reduced from 4.0 to 2.5 μm and, hence, defects rather than parametric effects are probably the cause of the yield reduction.

3.4 Contact Positioning

Tests were made to analyse the position and size of contacts with respect to the size of the underlying epitaxial silicon, the size of the metal line, and the proximity of the polysilicon gate. The results, in general, show a substantial reduction in functional yield when the contact opening is larger than either the underlying silicon or the metal interconnect line. The yield reduction can be as much as a factor of 3 or 4 when the metal does not completely encompass the opening. The yield is essentially zero when the metal must cross exposed epitaxal silicon edges. SEM inspection in the past has shown substantial negative tapering of etched silicon edges, while similar effects have not been observed with the oxide profile. Obviously these effects are strongly process dependent and, hence, similar experiments can also be used to compare various etching and flowing process techniques.

Results on the position of the contact with respect to the polysilicon gate, which measures alignment accuracy as well as etching control, show

little yield degradation with contact-to-poly spacings as small as 2.5 μm. Measurements incorporating a 1-μm contact to polygate spacing, however, show almost total shorting of the source to the gate. Leakage data also show no dependence on contact position for spacings above 1.0 μm.

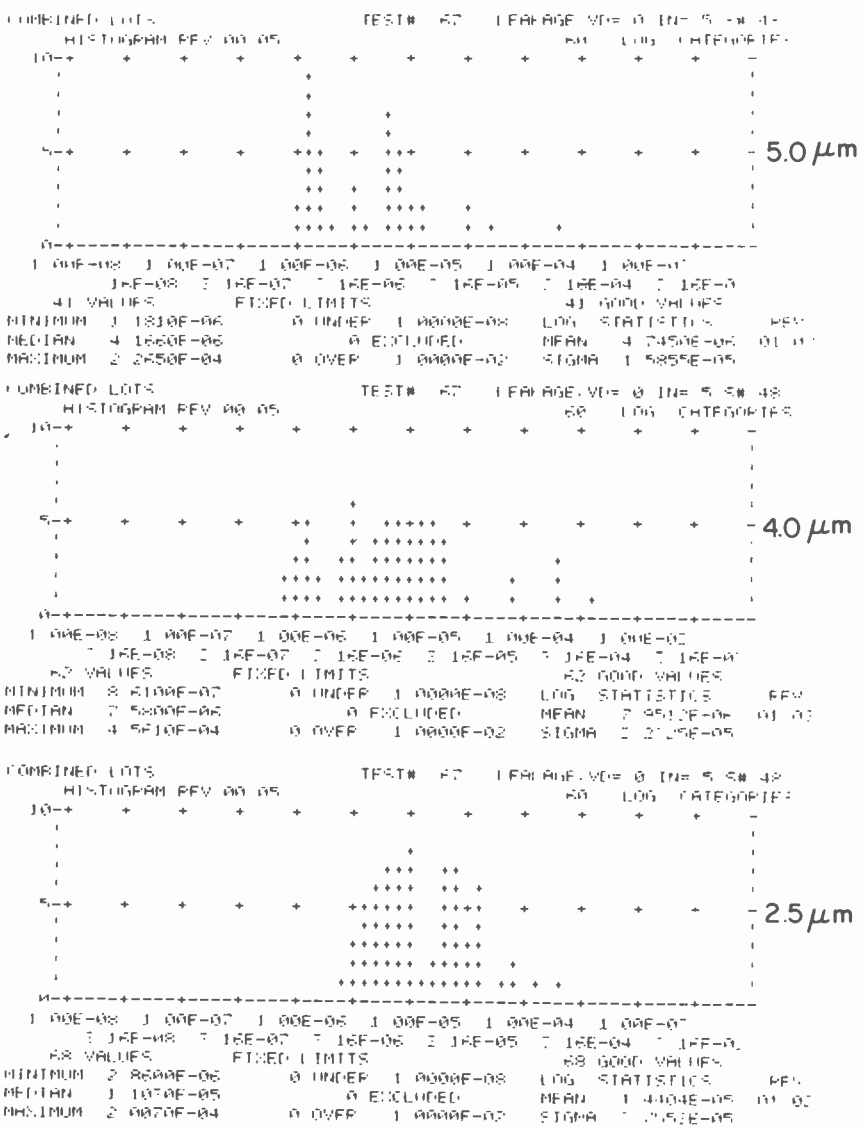


Fig. 5—Leakage current distributions as a function of poly gate overlap dimension.

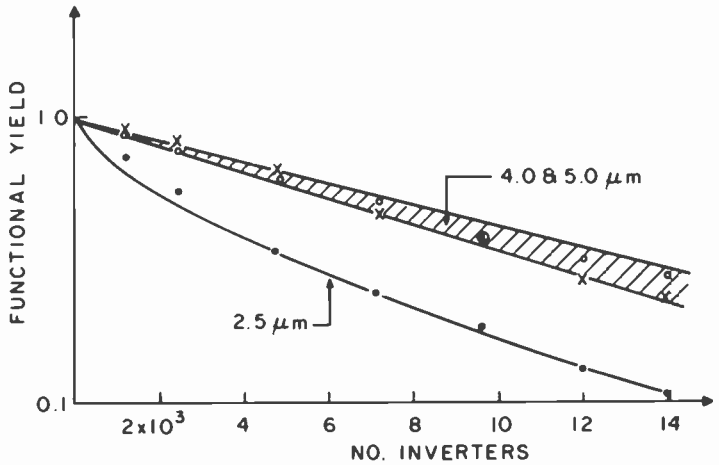


Fig. 6—Functional yield as a function of the source-to-drain spacing.

3.5 N+ and P+ Overlap of the Epi-Silicon Island

In this particular test sequence, two interrelated parameters were varied. One was the spacing between the N+ or P+ mask opening and the exposed epitaxial silicon island and the second was the spacing between the mask opening and the unexposed portion of the silicon island. Since the N+ and P+ levels are reverse tones of each other, both the N+ and P+ levels were evaluated simultaneously. The results indicate that no significant functionality or leakage problems result from either of these parameters even for spacings as low as 1.0 μm.

4. Conclusions

Data obtained from the transistor array indicate that high yields can be achieved on LSI circuits if the proper design rules are used and the limitation imposed by such electrical parameters as leakage currents and threshold voltages are minimal. Fig. 8 shows the variation in leakage current with the number of transistors tested. It is seen that, while individual wafers contain arrays whose leakage current increases linearly with the number of transistors, the leakage from wafer to wafer is ex-

tremely variable. Most arrays with 10,000 inverters, however, have leakage currents below 10 μA when proper design rules are used. The transistor array, therefore, demonstrates that high-yield low-leakage VLSI arrays can be fabricated using CMOS/SOS technology.

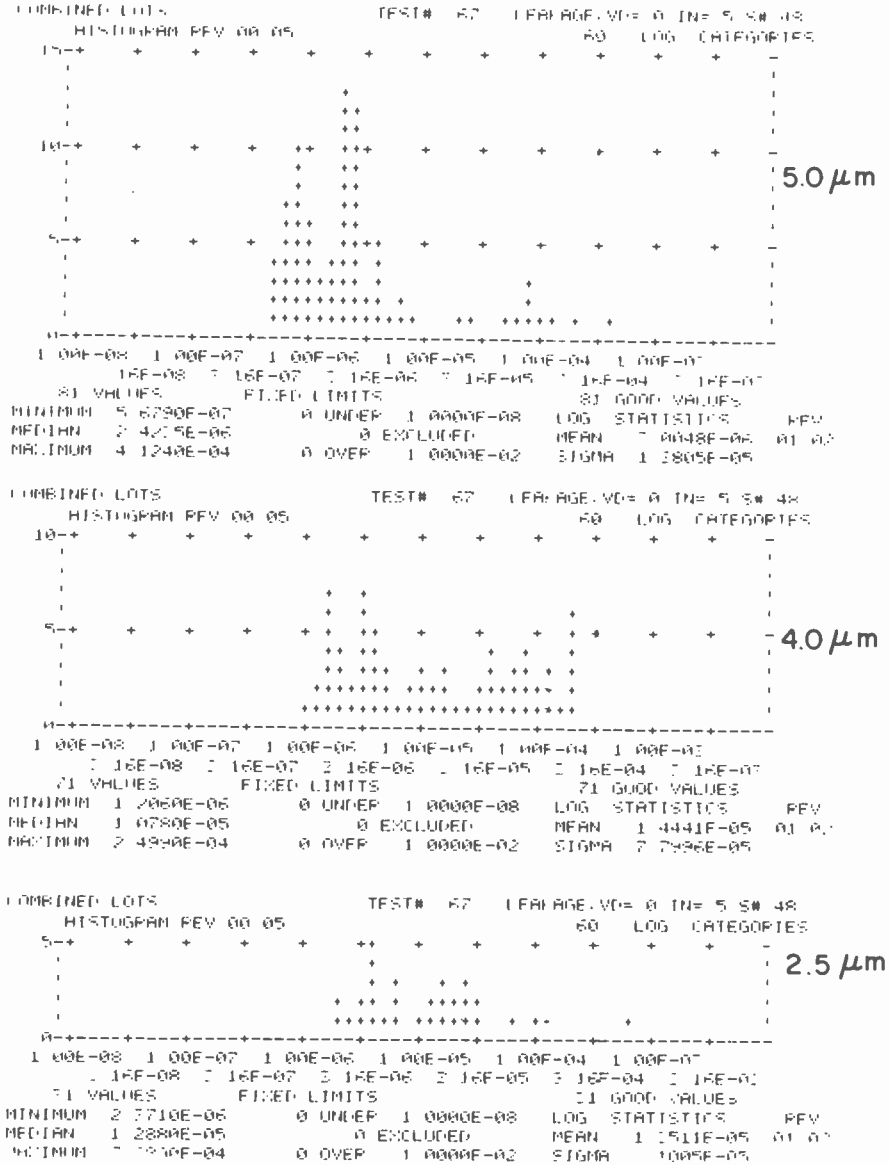


Fig. 7—Leakage current distributions as a function of source-to-drain spacing.

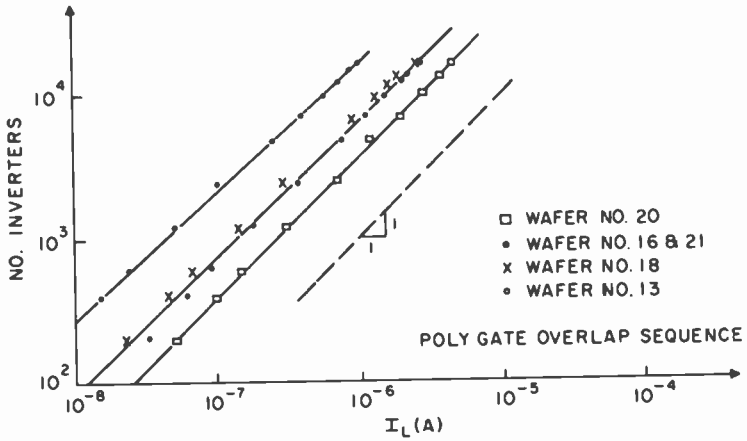


Fig. 8—Array leakage current as a function of the number of inverters in the array.

References:

[1]A. C. Ipri and J. C. Sarace, "Integrated Circuit Process and Design Rule Evaluation Techniques" *RCA Review*, **38**, p. 323 (1977).

[2]W. E. Ham, "Comprehensive Test Pattern and Approach for Characterizing SOS Technology" NBS Special Publication 400-56 (1980).

[3]G. W. Cullen, "The Preparation and Properties of Chemically Vapor Deposited Silicon on Sapphire and Spinal," *J. Crystal Growth*, **9**, 107 (1971).

Appendix 1

Transistor Array

Mask #	*Level*	*Dimension (Mils)*
10	Epi	Width = 0.4
11	Epi	Width = 0.3
12	Epi	Width = 0.2
13	Epi	Width = 0.1
14	Epi	Width = 0.05
15	Epi	Epi-to-Poly Spacing = 0.1
16	Epi	Epi-to-Poly Spacing = 0.05
17	Epi	Epi-to-Poly Spacing = 0.0
20	Poly-Gate	Channel Length = 0.2
21	Poly-Gate	Channel Length = 0.15
22	Poly-Gate	Channel Length = 0.10
23	Poly-Gate	Channel Length = 0.05
24	Poly-Gate	Gate Overlap = 0.15
25	Poly-Gate	Gate Overlap = 0.10

26	Poly-Gate	Special (Metal to Poly Shorts)
27	Poly-Gate	Contact-to-Poly Spacing = 0.15
28	Poly-Gate	Contact-to-Poly Spacing = 0.10
29	Poly-Gate	Contact-to-Poly Spacing = 0.05
230	Poly	Special (Poly-to-Epi Spacing)
30	N+	Epi to N+ Spacing = 0.2
31	N+	Epi to N+ Spacing = 0.15
32	N+	Epi to N+ Spacing = 0.10
33	N+	Epi to N+ Spacing = 0.05
40	Contacts	Poly/Source = 0.2 × 0.2 Drain = 0.2 × 0.4
140	Contacts	Poly/Source = 0.3 × 0.3 Drain = 0.3 × 0.5
50	Metal	V_{CC}/V_{DD} = 0.3 Poly/Drain = 0.5
90–93	Reverse Tone of 30–33	
130–133	Same as 30–33	

Carrier Lifetime in Photovoltaic Solar Concentrator Cells by the Small-Signal Open-Circuit Decay Method

A. R. Moore

RCA Laboratories, Princeton, NJ 08540

Abstract—The conventional open circuit voltage decay method of lifetime determination in junction structures has been modified by the addition of dc bias. This converts the decay from linear to exponential in time. Ambiguity in the junction boundary condition no longer influences the result. Application to concentrator solar cells operating over a wide range of solar intensity is discussed.

Introduction

Solar cells designed for use in solar concentrator systems operate at solar insolation levels of 1 to 100 W/cm^2 (10 to 1000 suns). From a purely semiconductor point of view, the chief difference between the conventional cell (one sun) and the concentrator cell is that the latter operates at an injected hole–electron pair density Δp of 10^{15} to 10^{17}/cm^3 which is often large compared to the thermal equilibrium majority carrier concentration in the base. As a consequence it becomes important to know how the minority-carrier lifetime in the base varies with injection (sun) level. The lifetime in the surface heavily doped layer is ordinarily not affected except at extremely high insolation levels not considered in the present discussion.

A well-known technique for lifetime measurement in ordinary diodes is the open-circuit decay method, originally proposed by Gossick[1] and by Lederhandler and Giacolletto[2] for use with germanium diodes and

more recently discussed by Wilson[3] in connection with p^+-π-n^+ silicon diodes. The diode under test is pulsed into the forward direction through an auxiliary fast switching diode. The pulse current is high enough to bring the test diode to the injection level desired and the pulse width is wide enough to attain an approximately steady-state condition in the base. After the pulse goes off, one observes the open circuit voltage across the diode with a high impedance probe. The auxiliary diode serves to isolate the test diode from the pulse generator.

In the special case of solar cells, the forward current pulse can be replaced by a fast light flash. A GaAs light emitting diode is particularly suitable for silicon solar cells since the emitted infra red light is just the right wavelength to penetrate deeply into the silicon base. No auxiliary diode is needed. However, with most available LED lamps, it is difficult to get sufficient light onto the solar cell to operate at the highest injection levels.

In any event, from the form of the V_{oc} versus time curve after the pulse goes off one calculates the lifetime, using known theory. The purpose of the present note is to show how a simple modification of the conventional procedure, particularly applicable to solar cells, enables several deficiencies in the calculation to be corrected, resulting in a more reliable measurement of the lifetime as a function of injection level.

Theory

The analysis of the open-circuit decay lifetime method requires the knowledge of the time dependence of the minority-carrier concentration in the base at the junction edge and the voltage response of the junction to that concentration. Because all recent concentrator cells incorporate the back surface field (BSF) concept, in which a heavily doped n^+ layer on the n-type base is used to reduce the minority-carrier recombination to almost zero at the back contact, these cells lend themselves to a particularly simple derivation. During the injection phase when the system has reached a steady state, virtually no pairs are lost to the back contact, and since the base width is normally only about one diffusion length thick, the diffusion gradient across the base is small. When the injection pulse (or light flash) is cut off ($t = 0$) and the diode is open circuited, the pair concentration in the base at the junction edge decays almost entirely by recombination; the diffusive flow can be neglected. It is, therefore, unnecessary to solve the time dependent diffusion equation and the decay can be written simply as

$$(p_n)_0 - p_{n0} = \Delta p(t) = \Delta p_0 e^{-t/\tau}. \qquad [1]$$

Here p_{n0} is the thermal equilibrium hole concentration, $(p_n)_0$ is the hole

concentration in the base at the junction edge ($x = 0$), and Δp_0 is the excess pair concentration at $x = 0$, $t = 0$.

For the response of the junction to the excess minority concentration in the base, early authors[2] assumed the Shockley small-signal boundary condition

$$(p_n)_0 = p_{n0}e^{qV_j/kT}. \qquad [2]$$

Eq. [2] leads directly to the expression for the open-circuit voltage in terms of the excess pair concentration which, when combined with Eq. [1], gives the measured quantity, the time dependence of the open circuit voltage across the junction,

$$V_j = \frac{kT}{q} \ln (1 + (\Delta p_0/p_{n0})e^{-t/\tau}). \qquad [3]$$

As long as $\Delta p_0 e^{-t/\tau} \gg p_{n0}$, we may neglect the one in the argument of the logarithm, yielding

$$V_j \cong \frac{kT}{q} \ln \left(\frac{\Delta p_0}{p_{n0}}\right) - \frac{kT}{q}\frac{t}{\tau}. \qquad [4]$$

Since the first term on the right is a constant, equal to the voltage at $t = 0$, Eq. [4] shows a *linear* decay of the open-circuit voltage with time whose slope gives the lifetime. This linear decay is the characteristic feature of the method. On the other hand, if we wait until $(\Delta p_0/p_{n0})e^{-t/\tau} \ll 1$,

$$V_j \cong \frac{kT}{q}\frac{\Delta p_0}{p_{n0}} e^{-t/\tau}. \qquad [5]$$

For an initial level corresponding to 100 suns ($\Delta p_0 \sim 10^{16}$), this *exponential* decay would begin at about 25 lifetimes.

The chief difficulty with applying this commonly accepted theory to actual silicon solar cells over a wide range of injection levels is that Eq. [2] does not really represent the true situation. If it did, silicon diodes would follow the ideal current-voltage law. Instead, due to various forms of nonlinearity at high levels and to recombination in the space-charge region at low levels, practical diodes follow the modified law

$$I = I_0(e^{qV/nkT} - 1). \qquad [6]$$

The diode quality factor n is a function of I (i.e., the injection level) and I_0 is set by some combination of the minority-carrier concentration at thermal equilibrium and thermal generation in the space-charge region (leakage current). The value of n is commonly observed to vary between 1 and about 3.

An obvious way to see the variations possible in the boundary condi-

tion of Eq. [2] is to consider the BSF concentrator solar cell already mentioned. The cell is really much closer to the p^+-π-n^+ diode considered by Wilson[3] than to the simple junction at low level envisioned by Eq. [2]. For the equivalent p^+-n-n^+ solar cell, one must consider the voltage drops across the p^+-n and n-n^+ junctions separately. Wilson gives the boundary condition

$$np = n_i^2 e^{qV/kT}, \tag{7}$$

where V is the sum of the drops across the junctions, the electron and hole concentrations referring to the central base region where space-charge neutrality also requires that $n = p + N_d$, where N_d is the donor density in the base. Then

$$p(p + N_d) = n_i^2 e^{qV/kT}. \tag{8}$$

Under high injection conditions $p \gg n_0 = N_d$ and Eq. [8] becomes

$$p = n_i e^{qV/2kT}, \tag{9}$$

while at low level $p \ll n_0$ and Eq. [8] reverts to Eq. [2]. The result of using Eq. [9] instead of Eq. [2] for the boundary condition in developing the response of the junction to the decaying pair concentration in the base is mainly to put the factor 2 into the determining equation for the linear decay when in the clearly high-level regime, and having the factor vary from 2 to one as the level decreases toward $p < n_0$. On the other hand Van Vliet[4] rejects Eq. [8], arguing that it is only correct if V is interpreted as the applied terminal voltage under conditions of current flow. Since there is no net current flow in the open-circuit decay method, it would appear that under the Van Vliet prescription yet another version of the boundary condition should be used. We shall not repeat that version here. Our only point is to show that the value of the multiplier n is ambiguous. Add to that the questionable influence of the space-charge recombination current present in all practical silicon diodes on the boundary condition at low injection levels. The result is that the defining equation for the open circuit decay lifetime method can be written

$$V_{oc} = \frac{nkT}{q}\frac{t}{\tau}, \tag{10}$$

with n an empirical constant varying with injection level. We take this n to be the same as the n in Eq. 6.

An additional problem arises in the application of Eq. [5]. The current-voltage curve of Eq. [6] and its equivalent short-circuit current-open circuit voltage curve for solar cells represents the static situation. In a dynamic case, the junction capacitance shunting the voltage source discharges through the junction dynamic resistance. At the low junction

voltages envisioned in Eq. [5], this effect can dominate the decay.[5] It is easy to show this analytically. The dynamic $I_{sc} - V_{oc}$ equation including the capacitance term is

$$I_0(e^{qV_{oc}/nkT} - 1) + C\frac{dV_{oc}}{dt} = I_{sc}. \quad [11]$$

Although the junction capacitance, C, is voltage dependent, the variation with voltage is only proportional to $(V + V_d)^{-1/2}$ where V_d is the diffusion potential. For small voltages, C may be taken as a constant. In the same voltage range, the exponential may be expanded to linearize the differential equation. Setting $nkT/(qI_0) = R_0$ and $R_0C = \tau_j$, we get

$$\tau_j\frac{dV_{oc}}{dt} + V_{oc} = R_0I_{sc}. \quad [12]$$

R_0 is the dynamic resistance of the junction near $V = 0$ and τ_j is the junction time constant.

I_{sc} itself is a function of time, the exact mode of decay depending on the boundary conditions, dimensions, and material constants. If we represent this as $I_{sc} = I_de^{-t/\tau}$, where I_d is a constant,

$$\tau_j\frac{dV_{oc}}{dt} + V_{oc} = R_0I_de^{-t/\tau} = f(t), \quad [13]$$

which is a linear differential equation with constant coefficients and a forcing term. With the initial condition that the system is in a steady state at $t = 0$, the complete solution to Eq. [13] is

$$V_{oc} = \frac{R_0I_m}{\tau_j - \tau}[\tau_je^{-t/\tau_j} - \tau e^{-t/\tau}]. \quad [14]$$

The decay of the open-circuit voltage is dominated by τ or τ_j, whichever is the larger. τ is typically 10 μs or so. Using actual values of I_d, $\tau_j = R_0C$ is typically 100 μs or more. Thus, a nonlinear region normally observed in the tail of the open-circuit voltage decay does not measure the base lifetime as predicted by the ideal Eq. [5], but is essentially a lumped circuit artifact.

In summary, the open-circuit decay lifetime method can give useful results at high and intermediate injection levels where the decay is linear, provided the proper value of n is applied at each point over the range, while at low levels where the decay is approximately exponential it does not measure lifetime at all.

Small-Signal Open-Circuit Decay

The rather inconvenient situation described above in the application of theoretical formulas to the open-circuit lifetime measurement can be readily corrected to give a simpler and more accurate modification which we call the small-signal open-circuit decay. Suppose we illuminate the cell with a dc light in addition to the injection pulse or light flash. Then the time dependence of the pair concentration after termination of the injection is

$$\Delta p(t) = P + \Delta p_0 e^{-t/\tau}, \qquad [15]$$

where P is the steady or dc value of the pair concentration generated by the dc light and the second term is the decaying portion as in Eq. [1]. Following the same procedure as before using the Shockley boundary condition, but modifying Eq. [3] to take into account the diode factor n, assuming $(\Delta p_0/p_{n0})e^{-t/\tau} \gg 1$ and making the small-signal approximation $\Delta p_0 \ll P$, we find

$$V_{oc} = \frac{nkT}{q} \ln\left(\frac{P}{p_{n0}}\right) + \frac{nkT}{q}\left(\frac{\Delta p_0}{P}\right) e^{-t/\tau}. \qquad [16]$$

The first term on the right is again a constant, but now equal to the open-circuit voltage at $t = \infty$ determined by the dc injection level. The second exponential term carries the time dependence. Its amplitude depends on n and $\Delta p_0/P$, but the characteristic decay constant depends only on τ. Thus the lifetime τ can be determined at any level set by the dc light free of the ambiguities of the relationship between the open-circuit voltage and the pair concentration characterized by the "constant" n. Furthermore, the interfering effect of the junction time constant, which prevents the simple application of the exponential decay in Eq. [5] at low level, is removed. It requires only a small amount of dc light to bias the junction sufficiently in the forward direction to reduce the dynamic junction resistance to the point where the RC time constant is *smaller* than the lifetime, whence the junction capacitance is no longer a factor (Eq. [14]). The only precaution required is that the dc light level be large enough to satisfy the small-signal approximation and to maintain the open-circuit boundary condition in the presence of the leakage current. This is ordinarily no problem above 0.01 sun. At the other end of the injection level scale above 1000 suns where the open-circuit voltage begins to saturate, both the Shockley and the Wilson or Van Vliet boundary conditions, even with adjustable n factor, are very poor approximations. If one uses a more precise but more complicated equation which exhibits this saturation,[6] it can be shown that if the small-signal condition is retained, exactly the same result is obtained: the decay back

to the dc level is exponential with characteristic time constant τ. Of course, if the open-circuit voltage is totally saturated, the amplitude of the exponential decay will go to zero, but this is a limiting case of no practical import.

In practice, the small-signal method is extremely easy to apply. With the dc light focussed on the sample the resultant dc short-circuit current is measured. The short is removed, the pulse injection applied, and lifetime is determined directly from the "scope" trace as the $1/e$ decay time. The result of measurements at various dc light levels is a curve of lifetime versus short-circuit current, which is just the information desired for concentrator solar cell development. The pair injection level can be estimated from the dc open-circuit voltage and the use of Eq. [8]. within the limitations already discussed. Alternatively, the double pulse method of Collet[7] allows Δp to be determined experimentally from the base resistance during injection.

As the dc light level is reduced, it is not necessary to simultaneously reduce the pulse injection level to maintain the small-signal condition. All that is required is that the decay be observed somewhat later in time after the injection ceases ($t = 0$). Then the pulse level will be seen to sink naturally back to the exponentially decaying region. At that point and beyond the small-signal condition is guaranteed.

Experimental Results

The pair lifetime in the base of a number of concentrator solar cells of different base resistivity has been examined in order to illustrate the small-signal decay method. Some measurements were also made by the conventional open-circuit decay method and also by the double pulse conductivity method of Collet et al[7] for comparison purposes.

Fig. 1 shows the decay of the open-circuit voltage in the tail beyond the linear decay with and without dc line bias. With no light bias, the decay is never quite exponential because the junction RC time constant is controlling. As the voltage decays toward zero, the dynamic resistance of the junction continuously increases producing an increasing time constant. With a light bias equivalent to a dc open-circuit voltage of 0.25 V, short-circuit current of less than 0.1 mA, approximately 0.01 sun, the decay becomes quite exponential with a characteristic time constant of 7.04 μs. This measures the true decay of the pair concentration at the injection level set by the light bias because it is not influenced by the junction capacitance.

Fig. 2 shows how the characteristic time constant changes as the dc light level is varied, in this case, over a range of 30 to one, from about 0.3 to 10 suns. Thus, we find that the lifetime depends on the dc injection

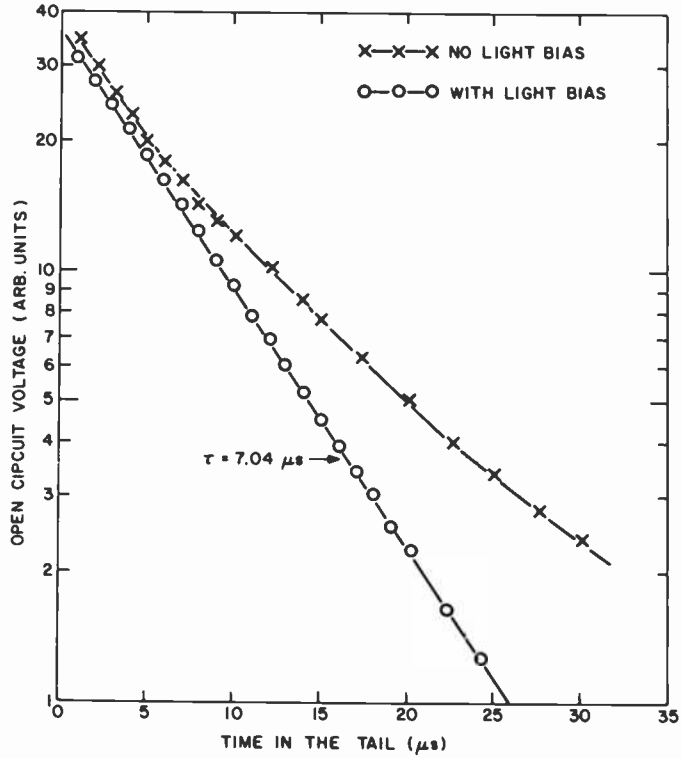


Fig. 1—Open-circuit voltage in the tail of the decay. The upper curve is with no light bias, the lower curve is with a small light bias equivalent to less than 0.1 mA short-circuit current. The normal operating range of these cells is near 1 ampere.

level, the exact details of which are important in operation of solar cells at high light level. Table 1 gives a direct comparison of the lifetime obtained by the conventional open-circuit linear decay and the small-signal exponential decay for this same sample. The value of n, needed in Eq. [10], was obtained separately from the slope of the $\log I_{sc}$ versus V_{oc} curve at each injection level, using the same dc light source. As can be seen, the results by the two methods are in general agreement provided the value of n is interpreted as that needed in Eq. [10]. Since the small-signal method is simpler and faster, as already explained, all subsequent measurements by open circuit were done small-signal.

Fig. 3 shows the small-signal lifetime versus short-circuit current for two samples. One sample had a very high thermal equilibrium base resistivity of about 5000 Ω-cm, the other a relatively low base resistivity of 2 Ω-cm. Both show the same general characteristic that we have found common to all samples tested; namely that the lifetime is small at very low levels, then rises at some intermediate level, but falls again at high

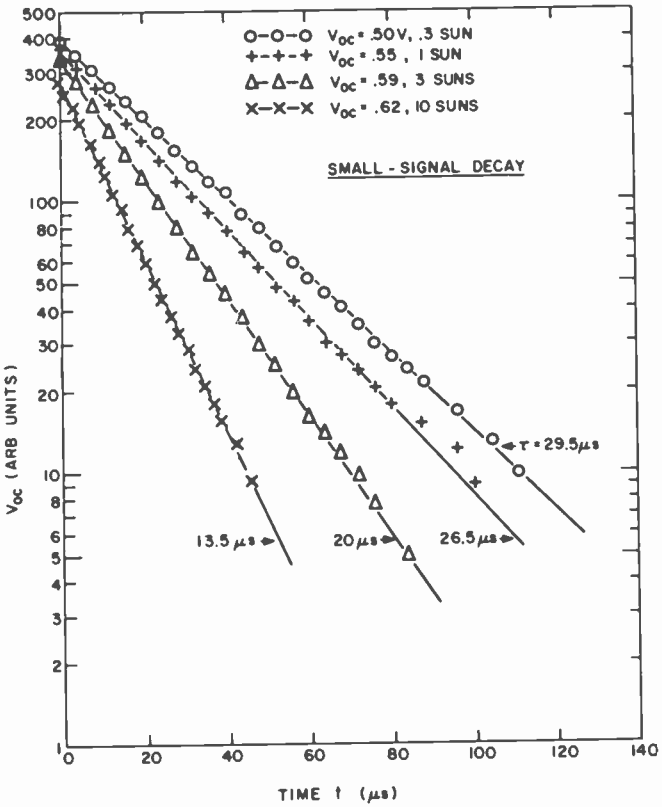


Fig. 2—Decay of the small-signal open-circuit voltage at various dc light levels from 0.3 to 10 suns.

level. Although the high resistivity sample shows much higher lifetime in the range below a few mA (one sun), the low resistivity sample is definitely superior at the higher levels near one ampere (about 100 suns). This illustrates an important point: the suitability of the base material must be determined by lifetime measurements made at the level at which

Table 1—A Comparison of Open-Circuit Linear and Small-Signal Exponential Decay

V_{oc} (Volts)	I_{sc} (mA)	Sun Equiv.	n	Eq. [10] τ from linear decay	Eq. [16] τ from Exp. decay (μs)
0.25	0.027	3.8×10^{-3}	2.83	7.5 μs	7.04
0.35	0.098	1.4×10^{-2}	2.50	9.75	10.0
0.40	0.28	4.0×10^{-2}	2.0	20.8	19.1
0.50	2.8	4.0×10^{-1}	1.67	26.0	29.1
0.59	23.1	3.3	1.50	18.3	20.0
0.66	280	40	1.0	8.2	8.2

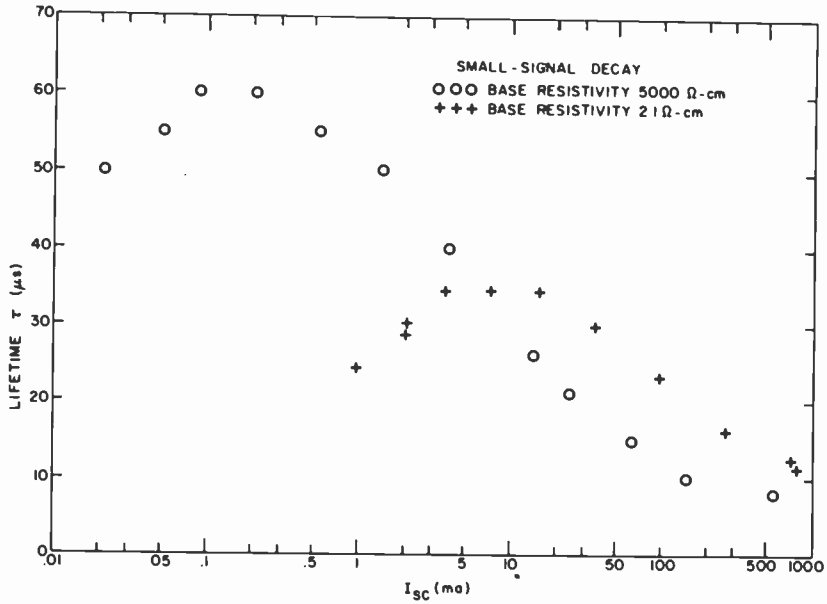


Fig. 3—Lifetime as a function of short circuit current, obtained by the small-signal open-circuit decay method, for two samples of widely different base resistivity.

the cells are to be used, not at lower levels. Other things being equal, the lower resistivity material is to be preferred, allowing a thicker optimum base width. We do not mean to imply that a low resistivity base is necessarily better, only that the shape of the lifetime versus short-circuit current curves must be taken into account. The small-signal open-circuit method provides a rapid and accurate means of doing it.

For completeness it seemed desirable to make a comparison with a wholly independent lifetime method. This was done using the method of Collet et al already mentioned. This double pulse method has the advantage that it also determines the injection level in absolute terms, although it is more troublesome in application. Thus Fig. 4 shows the lifetime determined by both a small-signal open-circuit decay and by double pulse conductivity versus the injection level obtained from the latter. The open-circuit method, as pointed out by Wilson,[3] is particularly well adapted to low injection levels. This is true for the open-circuit method, whether for linear or small-signal exponential decay. In the figure, the low range for the small-signal method is fully a decade below that which is possible by the double pulse method. In the region of overlap the agreement is seen to be quite satisfactory.

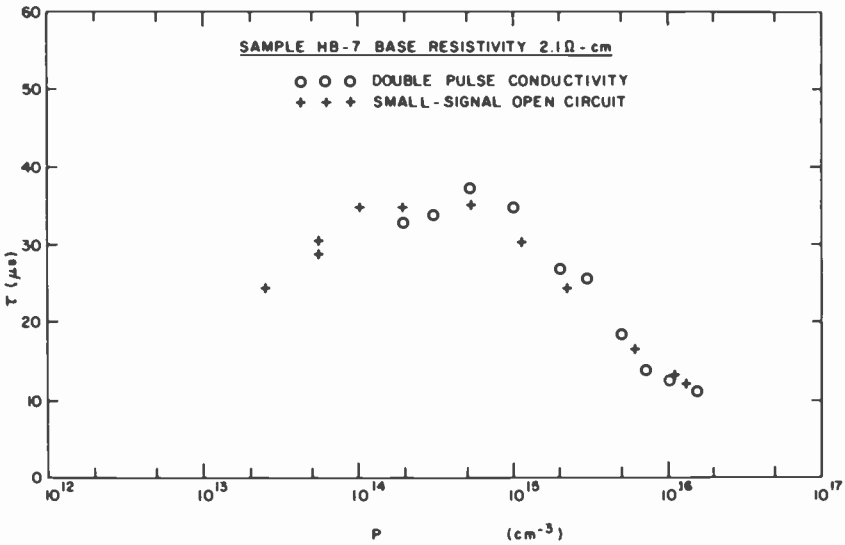


Fig. 4—Comparison of lifetime versus pair density obtained by double pulse conductivity and small-signal open-circuit decay.

Discussion

The question remains of how to understand the observed behavior of the lifetime as the injection level is increased. Within the context of the Shockley–Read–Hall theory,[8] recombination is generally assumed to take place through the intermediary of the recombination center or trap. A single level trap system with two states, negative and neutral, can yield either a τ increasing with injection level or τ decreasing with level depending on the energy of the center in the forbidden gap.

Consider the two cases separately. A center near the middle of the gap is the most efficient (low lifetime) because re-emission of electrons and holes once captured is minimized. In an n-type sample at very low injection levels near thermal equilibrium, the traps are filled with electrons. A nonequilibrium hole is captured by a negatively charged center making it neutral, then the trap is immediately refilled with an electron from the copious supply in the conduction band. The rate-determining step is the slower one, i.e., hole capture. At higher injection level, $\Delta n = \Delta p > n_0$. Now so many recombinations are occurring per second that a substantial fraction of the traps become filled with holes or, alternatively, it is said that the quasi Fermi level for holes moves down toward the trap level, partially emptying it of electrons, which is the same thing. Now the rate depends on both filling of centers by electrons and filling by holes. The lifetime rises. When $\Delta n = \Delta p \gg n_0$, the lifetime limit is

reached and the traps are completely loaded with holes (emptied of electrons). The trap level is now saturated.

Now consider a recombination center located near the conduction band edge, *above* the thermal equilibrium Fermi level in this n-type material. At low injection level this is an inefficient recombination center because an electron falling into it from the conduction band has a high probability of re-emission before recombining with a hole. The traps are only partially filled, the lifetime is long. But the fraction of the traps filled by electrons is a dynamic balance between filling from the conduction band and emptying by re-emission to the conduction band. At high injection levels, as $\Delta n = \Delta p \gg n_0$, the balance is driven toward increased filling by the increased number of electrons in the conduction band. Alternatively, one says that the quasi Fermi level for electrons rises toward the trap level. As the trap level fills up with electrons, it becomes a more efficient medium for hole recombination and the lifetime falls.

It is, therefore, clear that a system of two recombination levels strategically located in the forbidden gap can explain a lifetime that first increases with level until the first trap set is saturated, then decreases as the level goes higher and the second trap set comes into play.[9,10] The theoretical parameters required are the hole and electron lifetimes when the centers are completely full or completely empty, respectively, for both trap levels, and a parameter that contains the trap depth of level 2 in the form $K = e^{(E_t - E_f)/kT}$. Fig. 5 shows how well the Shockley–Read–Hall theory with two trap levels can be made to match the data.

Conclusion

A modification of the open-circuit decay method of lifetime determination employing bias light and exponential decay has been found particularly suitable for rapid and accurate lifetime versus injection level measurements useful in solar cell development. The method avoids the ambiguity inherent in the linear decay approach from which it derives. The lifetime in the base of a silicon solar concentrator cell cannot be judged by measurements at low injection level but must be measured at the level at which it is used. In general, the lifetime has been found to go through a maximum at some injection level that is larger the lower the resistivity of the base material. A two-level recombination theory based on the Shockley–Read–Hall recombination mechanism can explain the observed results.

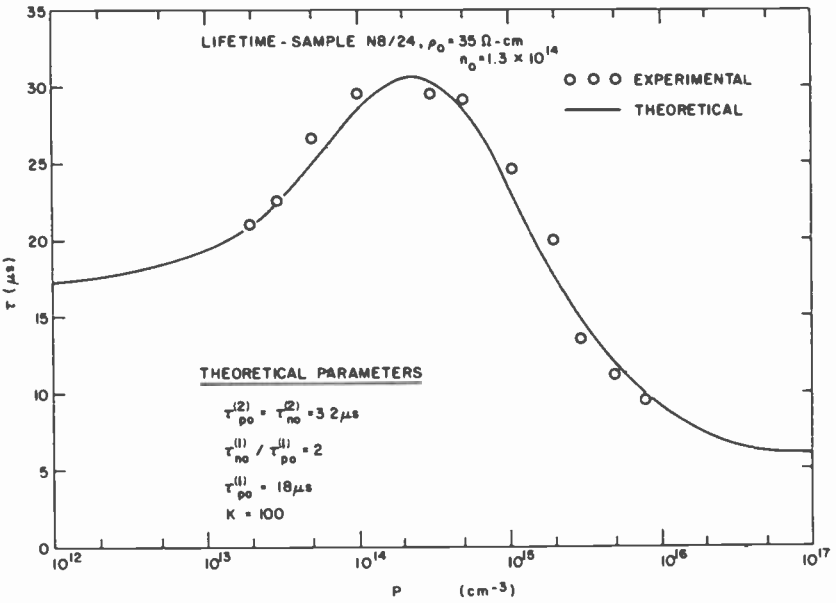


Fig. 5—Match of two-level Shockley–Read–Hall recombination theory to lifetime data. Parameters for the two levels used to calculate the theoretical curve are given in the figure. $K = 100$ is equivalent to a trap depth of 0.2 eV below the conduction band.

Acknowledgment

I would like to acknowledge the cooperation of George Swartz in the application of the double pulse conductivity method to solar cells and for supplying me with a variety of concentrator cells from current production. Without his help this work would not have been possible. I would also like to thank D. Redfield for his careful reading of the manuscript.

References:

[1] B. R. Gossick, "Post-Injection Barrier Electromotive Force of p-n Junctions," *Phys. Rev.*, **91**, p. 1012 (1953).

[2] S. R. Lederhandler and L. J. Giacoletto, "Measurement of Minority Carrier Lifetime and Surface Effects in Junction Devices," *Proc. I.R.E.*, **43**, p. 477 (1955).

[3] P. G. Wilson, "Recombination in Silicon p-π-n Diodes," *Solid State Elec.*, **10**, p. 145 (1967).

[4] K. M. Van Vliet, "High Injection Theories of the p-n Junction in the Charge-Neutrality Approximation," *Solid State Elec.*, **9**, p. 185 (1966).

[5] B. R. Gossick, *Potential Barriers Semiconductors*, pp. 129–130, Academic Press, New York (1964).

[6] Lee A. M. Vasil'ev, V. M. Evdokimov, A. P. Landsman and A. F. Milovanov, *Applied Solar Energy* (Geliotekhma) **11**, p. 72 (1975) or S. R. Dhariwal, L. S. Kothari, and S. C. Jain, "Saturation of Photovoltage and Photocurrent in p-n Junction Solar Cells," *IEEE Trans. Elec. Devices*, **Ed-23**, p. 504 (1976). Both give similar formulas for the boundary condition that are more general than the Shockley equation.

[7] J. Collet, L. Bailon, J. C. Brabant, J. Barrau and M. Brousseau, "New Methods for Carrier Lifetime Measurements in p-π-n Structures," *Solid State Elec.*, **16**, p. 999 (1973).

[8] R. N. Hall, "Electron-Hole Recombination in Germanium," *Phys. Rev.*, **87**, p. 387 (1952) and W. Shockley and W. T. Read, "Statistics of the Recombination of Holes and Electrons," *Phys. Rev.*, **87**, p. 835 (1952).

[9] V. L. Dalal and A. R. Moore, "Design Considerations for High-Intensity Solar Cells," *J.A.P.*, **48**, p. 1244 (1977).

[10] For a general discussion, see J. Moll, *Physics of Semiconductors*, Chapter 6, McGraw-Hill, New York (1964).

MOS Threshold Voltage Monitoring

W. A. Bosenberg

RCA Solid-State Division, Solid State Tech. Center
Somerville, NJ 08876

Abstract—The threshold voltage of an MOS transistor is an important parameter that affects the performance of integrated circuits. Threshold voltage depends on oxide thickness, channel doping, substrate bias, metal work function, and surface state density. The dc monitoring of the threshold voltage can be greatly improved by including the substrate bias effect in the measurement. The procedure used extrapolates (to a zero drain current) threshold voltage from the drain-current/gate-voltage square-law characteristics. A second least-square fit is used for the dependence of the extrapolated threshold voltage with substrate bias. The fitting parameter is the channel concentration. The procedure makes it possible to analyze the threshold voltage as a function of an effective surface-state density. The values calculated from the experiments are invariable over a limited temperature range.

Introduction

During the last ten years, the number of MOS transistors on a single integrated circuit chip has been increasing at a very large rate. This increase was accomplished in the early years by clever circuit design together with the use of polysilicon layers for busing, thus reducing a static memory cell of six transistors to a functionally equivalent dynamic cell consisting of a single transistor and a storage capacitor. Recently, overall scaling of MOS integrated circuit devices has been used to reduce the chip size. Presently, IC's using channel lengths as small as 4 μm are in the production stage and experimental devices with channel lengths as small as 0.5 μm have been made with 1.5 to 2 μm being more typical. As device dimensions are scaled down, other parameters associated with the IC, such as substrate dopant concentration, oxide thickness, and supply voltage, must also be scaled. Smaller supply voltage leads to the requirement of increased control of threshold voltage. While most digital systems seem to stay with 5-volt supplies, there is an appreciable consumer segment of digital wrist watches with 1.5 to 3 volt supply voltages. For these present applications, threshold voltage tolerances are 75 mv compared with 250 mv for the older technologies.

Traditionally, MOS capacitor monitoring has been used to obtain silicon–silicon-dioxide interface charge information. This technique, however, is very insensitive to minor changes in substrate surface concentration that may occur during thermal processing. This paper describes a technique for monitoring the basic device parameters that determine the threshold voltage, i.e., substrate donor or acceptor surface concentration, effective silicon–silicon-dioxide interface charge, and oxide thickness. In addition, this technique will give information on the MOS transistor gain, i.e., the K-factor. For the circuit designer, the full drain-current/drain-voltage characteristics as a function of gate and substrate voltage are of interest. The data measured has been used in this fashion and indicates, as expected, the need for the modeling of carrier mobility saturation effects. Substantially more computer power is required for this and it is better left for off-line analysis with a non-linear regression.

MOS Transfer Basics

Fig. 1 shows a cross-sectional view of a typical MOS transistor. Figs. 2a and 3a show the drain-current/drain-voltage characteristics of such a device. In the saturated region, the drain current I_d is independent of the drain voltage V_d and depends on the difference between applied gate voltage V_g and the threshold voltage V_t:[1]

$$I_d = K(V_g - V_t)^2 \qquad [1]$$

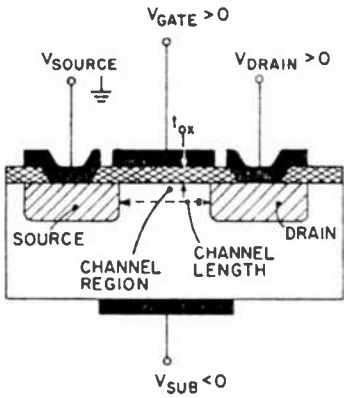


Fig. 1—Cross section of MOS transistor with source, gain drain and substrate contacts. For an NMOS transistor, the drain and gate voltages are positive. The substrate voltage is negative against the grounded source. For PMOS transistors, the polarities are reversed. There is little practical difference between the case in which the drain voltage is fixed and the gate voltage is changed to obtain a certain drain current and the case in which both the drain and gate electrodes are interconnected. The latter case is more convenient to use, as in our case.

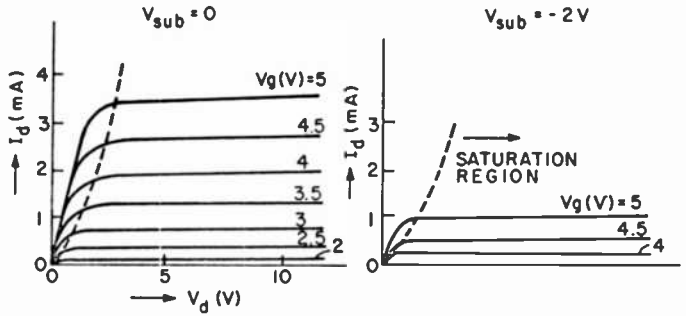


Fig. 2—Drain-current/drain-voltage characteristics as a function of gate voltage for CD4007 NMOS transistor (a) for zero substrate bias and (b) for $V_{sub} = -2$ volts. The substrate effect is very strong.

where $K = C_{ox}\mu W/(2l)$ is the K factor, μ the channel mobility, W the channel width, and l the channel length. This threshold voltage V_t is the same as the voltage required to reach strong inversion in the MOS capacitor having the same substrate dopant concentration.

Figs. 2b and 3b show the drain-current/drain-voltage characteristics when a substrate bias has been applied. The reduction in drain current with substrate bias is larger with higher substrate dopant concentrations, which is usually the case for NMOS transistors with diffused p-wells.

Threshold Voltage Parameters

The threshold voltage V_t of a MOS transistor is a complex function of oxide thickness t_{ox}, channel concentration, surface state densities N_{ss}, and substrate voltage V_{sub}. In its simplest form, it can be written[2]

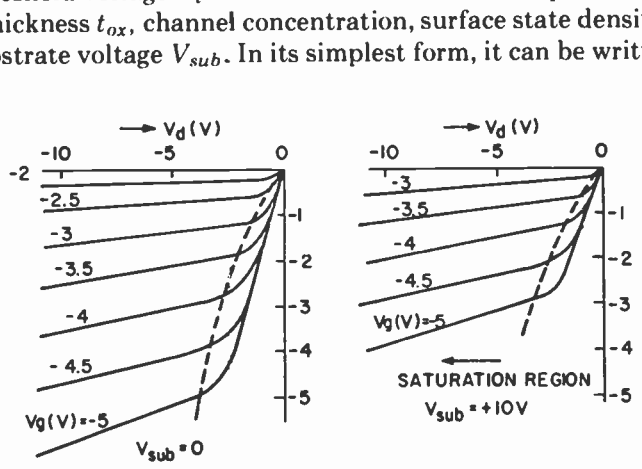


Fig. 3—Drain-current/drain-voltage characteristics as a function of gate voltage for CD4007 PMOS transistor (a) for zero substrate bias and (b) $V_{sub} = 10$ volts. The substrate effect is relatively weak.

$$V_t = -\frac{qN_{ss}}{C_{ox}} \pm V_f - 0.45 + \frac{\sqrt{2q\epsilon_o\epsilon_{Si}N(|V_{sub}| + 2V_f)}}{C_{ox}} \quad [2]$$

$$= V_o + C\sqrt{|V_{sub}| + 2V_f} \quad [2a]$$

for devices fabricated with aluminum or $n+$ poly gates, and

$$V_t = -\frac{qN_{ss}}{C_{ox}} \pm V_f + 0.65 - \frac{\sqrt{2q\epsilon_o\epsilon_{Si}N(|V_{sub}| + 2V_f)}}{C_{ox}} \quad [3]$$

$$= V_o - C\sqrt{|V_{sub}| + 2V_f} \quad [3a]$$

for devices fabricated with $p+$ poly gates.

In these equations the plus sign is used for NMOS transistors, the minus sign is used for PMOS transistors, and

$\epsilon_o\epsilon_{Si}$ is the dielectric constant of silicon,

$C_{ox} = \frac{\epsilon_o\epsilon_{SiO2}}{t_{ox}}$ is the oxide capacitance,

$V_f = \frac{kT}{q} \ln (N + N_i)/N_i$ is the Fermi voltage,

q is the electronic charge, and

N_i is the intrinsic carrier concentration of the silicon.

It should be noted that in Eqs. [2a] and [3a] the Fermi voltage V_f is a weak function of the channel doping concentration and its value must be found using an interactive procedure.

Measurement Procedure

The measurement procedure consists of first measuring the drain current as a function of gate and substrate bias, and then performing a least-square fit using the relationship

$$\sqrt{I_d} = c(V_g - V_t). \quad [4]$$

An error parameter R (correlation factor) can be used to determine how valid the least-square fit is. For a perfect fit, R is 1. In the present case, the fit is excellent, and for convenience, 1-R expressed in parts per million is used. A good fit shows a value between 50 and 400 ppm. The major problem comes from sub-threshold leakage when the semiconductor surface is in weak inversion and the square law Eq. [1] is not obeyed.

The original concept included a matrix of $m = 9$ drain currents and $n = 8$ substrate biases, and measurements were made from 25 to 200°C in 25°C intervals. At the present, the measurement has been simplified

to five drain currents (10, 55, 140, 250 and 400 μA) and to five substrate biases of 0, 1, 3, 6, and 10 volts at room temperature. The values have been picked to give roughly equal weight along the line for the least square fits. The lowest current value of 10 μA was chosen over the original lowest current value of 1 μA so that the MOS system would be in strong inversion and the effect of subthreshold leakage would be minimized. Experimentally it is easier to connect the gate terminal to the drain terminal, force a drain current, and measure the gate voltage, rather than have a constant drain bias and to vary the gate voltage in a computer loop to obtain a certain drain current. The extrapolated threshold voltages show little change, while the K factor is slightly higher due to the fine slope in the saturated drain current characteristic, especially for PMOS (see Fig. 3a).

Extrapolated Threshold Voltage

The summations needed for performing the least square fit[3] to experimental data taken from m discrete points defined by I_{di} and V_{gi}, where i ranges from i to m, are

$$\sum_{i=1}^{i=m} \sqrt{I_{di}}\,,\ \sum_{i=1}^{i=m} I_{di},\ \sum_{i=1}^{i=m} V_{gi},\ \sum_{i=1}^{i=m} V_{gi}^2,\ \text{and}\ \sum_{i=1}^{i=m} \sqrt{I_{di}}\,V_{gi}.$$

Averages are defined by

$$\overline{\sqrt{I_d}} = \frac{\sum_{i=1}^{i=m} \sqrt{I_{di}}}{m} \text{ and } \overline{V_g} = \frac{\sum_{i=1}^{i=m} V_{gi}}{m}\,. \qquad [5]$$

Then the threshold voltage is

$$V_t = \overline{V_g} - \frac{\overline{\sqrt{I_d}}\left[\sum_{i=1}^{i=m} \sqrt{I_{di}}\,V_{gi} - m\,\overline{\sqrt{I_d}}\,\overline{V_g}\right]}{\sum_{i=1}^{i=m} I_{di} - m\,\overline{\sqrt{I_d}}\,\overline{\sqrt{I_d}}}\,, \qquad [6]$$

The correlation coefficient for this least square fit is:

$$1-R=\frac{1\times 10^{6}\left[1-\left|\sum_{i=1}^{i=m}\sqrt{I_{d_i}}\,V_{g_i}-m\,\sqrt{\overline{I_d}}\,\overline{V_g}\right|\right]}{\sqrt{\left[\sum_{i=1}^{i=m}I_{d_i}-m\,\sqrt{\overline{I_d}}\,\sqrt{\overline{I_d}}\right]\left[\sum_{i=1}^{i=m}V_{g_i}^{2}-m\,\overline{V_g}\overline{V_g}\right]}}\ (\text{ppm}). \quad [7]$$

Channel Concentration

A second least square fit is done for V_{tj} and V_{subj}, where j ranges from 1 to n by using the following summations:

$$a=\sum_{j=1}^{j=n}V_{tj}$$

$$b=\sum_{j=1}^{j=n}V_{tj}^{2}$$

$$c=\sum_{j=1}^{j=n}\sqrt{|V_{subj}|+2V_f}$$

$$d=\sum_{j=1}^{j=n}|V_{subj}|+2V_f$$

$$e=\sum_{j=1}^{j=n}V_{tj}\sqrt{|V_{subj}|+2V_f}$$

Averages are defined as

$$\overline{V_t}=\frac{a}{n},\ \sqrt{\overline{|V_{sub}|+2V_f}}=\frac{c}{n}\,. \quad [8]$$

The channel concentration is obtained using the equation

$$N=\frac{C_{ox}^{2}}{2q\epsilon_o\epsilon_{si}}\,\frac{e-n\,\sqrt{\overline{|V_{sub}|+2V_f}}\,\overline{V_t}}{d-n\,\sqrt{\overline{|V_{sub}|+2V_f}}\,\sqrt{\overline{|V_{sub}|+2V_f}}}\,. \quad [9]$$

Solving for V_o in Eq. [2a], one obtains

$$V_o=\overline{V_t}-\frac{C_{ox}}{\sqrt{2q\epsilon_o\epsilon_{si}N}}\times\frac{e-n\,\sqrt{\overline{|V_{sub}|+2V_f}}\,\overline{V_t}}{d-n\,\sqrt{\overline{|V_{sub}|+2V_f}}\,\sqrt{\overline{|V_{sub}|+2V_f}}}\,. \quad [10]$$

N_{ss} can then easily be calculated.

An error factor Q for the second least square fit is obtained from

$$1 - Q = \frac{1 \times 10^6 [1 - |(e - n\sqrt{|V_{subj}| + 2V_f}\,\bar{V}_t)|]}{\sqrt{(b - n\bar{V}_t\bar{V}_t)(d - n\sqrt{|V_{subj}| + 2V_f}\sqrt{|V_{subj}| + 2V_f})}} \text{ (ppm).} \quad [11]$$

Experimental Results

Tables 1 and 2 show the data matrix on a CD4007 NMOS and PMOS transistor. Temperature steps of 25°C from 25°C up to a temperature of 200°C were used. As is well known,[4] the threshold voltage of MOS transistors decreases with temperature by between 0.4 and 0.5 volts in this temperature range (see Tables 3 and 4). The channel width of the p-type transistor is larger than that of the n-type MOS transistor. Eq. [2] is plotted in Fig. 4 and Eq. [3] is plotted in Fig. 5, both for room temperature. Over the 25–200°C temperature range, the channel doping concentration is constant. This indicates that the approach taken is reasonable. The temperature dependence of the threshold voltage is basically the temperature dependence of the Fermi voltage. The slight increase in the error factors 1-R and 1-Q with temperature may have been caused by the slightly larger fluctuations of temperature during the measurements. In any case the correlation of the least square fits is excellent. Up to 150°C, the surface state density of both the NMOS and PMOS is rather constant, especially that of the PMOS transistor. The K-factor of both units decreases with temperature. The observed temperature dependence of $T^{-1.683}$ for the NMOS and $T^{-1.707}$ is slightly higher than the theoretical value for the mobility dependence of $T^{-1.5}$. The higher K-factor for the PMOS is not a reflection of hole mobility being larger than the electron mobility but rather of the factor-of-two bigger channel width. The remaining differences can be explained by slight variations in channel length and the dependence of mobility on doping levels in the channel region. The effective surface mobility is typically about ⅓ that of the bulk mobility for the same concentration.

Threshold Adjustment by Ion Implantation

Very often ion implantation is used to adjust the threshold voltage. Another term, V_{bulk}, must then be added to Eqs. [2] and [3];[5,6]

$$V_{bulk} = \frac{q}{C_{ox}} \int_0^{w_d} N(x)dx. \quad [12]$$

Table 1—Gate voltage V_g at various drain currents I_d as a function of substrate voltage V_{sub} for CD4007 NMOS transistor. The temperature ranges from 25 to 200°C in 25°C increments.

Temp (°C)	V_{sub} (volts)	V_g (volts) at I_d (μA) = 16	49	100	169	256	361	484	625
25	0.0	1.60	1.80	1.98	2.16	2.34	2.51	2.69	2.87
25	−0.4	1.99	2.18	2.36	2.53	2.70	2.87	3.05	3.22
25	−1.0	2.48	2.66	2.83	2.99	3.16	3.33	3.49	3.66
25	−1.8	3.02	3.19	3.35	3.51	3.67	3.83	3.99	4.16
25	−3.0	3.69	3.86	4.01	4.17	4.32	4.48	4.63	4.79
25	−5.0	4.61	4.77	4.92	5.06	5.21	5.36	5.51	5.67
25	−8.0	5.73	5.88	6.02	6.16	6.30	6.45	6.59	6.74
25	−12.0	6.95	7.09	7.23	7.36	7.49	7.63	7.77	7.91
50	0.0	1.54	1.75	1.95	2.14	2.32	2.51	2.70	2.89
50	−0.4	1.94	2.14	2.33	2.51	2.69	2.87	3.05	3.24
50	−1.0	2.44	2.63	2.80	2.98	3.15	3.33	3.50	3.68
50	−1.8	2.98	3.16	3.33	3.50	3.67	3.84	4.01	4.18
50	−3.0	3.66	3.83	4.00	4.16	4.32	4.48	4.65	4.82
50	−5.0	4.58	4.75	4.90	5.06	5.22	5.37	5.53	5.69
50	−8.0	5.71	5.86	6.01	6.16	6.31	6.46	6.61	6.77
50	−12.0	6.93	7.08	7.22	7.36	7.50	7.65	7.79	7.94
75	0.0	1.49	1.71	1.92	2.12	2.32	2.52	2.71	2.91
75	−0.4	1.90	2.11	2.31	2.50	2.69	2.88	3.07	3.27
75	−1.0	2.40	2.60	2.79	2.97	3.16	3.34	3.53	3.71
75	−1.8	2.94	3.14	3.32	3.50	3.67	3.85	4.03	4.21
75	−3.0	3.63	3.81	3.98	4.16	4.33	4.50	4.68	4.85
75	−5.0	4.56	4.73	4.89	5.06	5.22	5.39	5.56	5.73
75	−8.0	5.68	5.85	6.01	6.16	6.32	6.48	6.64	6.80
75	−12.0	6.91	7.06	7.21	7.36	7.51	7.66	7.82	7.98
100	0.0	1.43	1.67	1.89	2.10	2.31	2.52	2.73	2.94
100	−0.4	1.85	2.07	2.28	2.49	2.69	2.89	3.10	3.30
100	−1.0	2.35	2.57	2.77	2.96	3.16	3.35	3.55	3.75
100	−1.8	2.91	3.11	3.30	3.49	3.68	3.87	4.06	4.25
100	−3.0	3.60	3.79	3.97	4.15	4.34	4.52	4.70	4.89
100	−5.0	4.53	4.71	4.89	5.06	5.23	5.41	5.59	5.77
100	−8.0	5.66	5.83	6.00	6.16	6.33	6.50	6.67	6.84
100	−12.0	6.88	7.05	7.20	7.36	7.52	7.68	7.85	8.01

Table 1 (Contd.)

125	0.0	1.36	1.62	1.86	2.09	2.31	2.53	2.76	2.98
125	−0.4	1.80	2.04	2.26	2.48	2.69	2.91	3.12	3.34
125	−1.0	2.32	2.54	2.75	2.96	3.17	3.37	3.58	3.79
125	−1.8	2.87	3.09	3.29	3.49	3.69	3.89	4.09	4.30
125	−3.0	3.56	3.77	3.96	4.16	4.35	4.54	4.74	4.94
125	−5.0	4.50	4.69	4.88	5.06	5.25	5.43	5.62	5.81
125	−8.0	5.63	5.82	5.99	6.17	6.35	6.52	6.70	6.89
125	−12.0	6.86	7.04	7.20	7.37	7.54	7.71	7.88	8.06
150	0.0	1.30	1.57	1.82	2.06	2.31	2.54	2.78	3.02
150	−0.4	1.74	2.00	2.23	2.47	2.69	2.92	3.15	3.38
150	−1.0	2.27	2.51	2.73	2.95	3.17	3.39	3.61	3.83
150	−1.8	2.83	3.06	3.28	3.49	3.70	3.91	4.12	4.34
150	−3.0	3.53	3.75	3.95	4.16	4.36	4.56	4.77	4.98
150	−5.0	4.47	4.68	4.87	5.06	5.26	5.46	5.66	5.86
150	−8.0	5.61	5.80	5.99	6.17	6.36	6.55	6.74	6.93
150	−12.0	6.84	7.02	7.20	7.37	7.55	7.73	7.92	8.10
175	0.0	1.24	1.53	1.80	2.05	2.30	2.55	2.80	3.05
175	−0.4	1.70	1.97	2.22	2.46	2.70	2.94	3.18	3.42
175	−1.0	2.23	2.49	2.72	2.95	3.18	3.41	3.64	3.87
175	−1.8	2.80	3.04	3.27	3.49	3.71	3.93	4.16	4.38
175	−3.0	3.50	3.73	3.95	4.16	4.37	4.59	4.81	5.02
175	−5.0	4.45	4.66	4.87	5.07	5.28	5.48	5.69	5.90
175	−8.0	5.59	5.79	5.98	6.18	6.38	6.57	6.77	6.98
175	−12.0	6.81	7.01	7.20	7.38	7.57	7.76	7.95	8.15
200	0.0	1.17	1.48	1.76	2.03	2.30	2.56	2.82	3.09
200	−0.4	1.65	1.93	2.19	2.45	2.70	2.95	3.20	3.46
200	−1.0	2.19	2.45	2.70	2.94	3.19	3.43	3.67	3.92
200	−1.8	2.76	3.02	3.25	3.49	3.72	3.95	4.19	4.43
200	−3.0	3.47	3.71	3.94	4.16	4.39	4.61	4.84	5.07
200	−5.0	4.41	4.64	4.86	5.07	5.29	5.51	5.73	5.95
200	−8.0	5.56	5.77	5.98	6.18	6.39	6.60	6.81	7.02
200	−12.0	6.79	6.99	7.19	7.39	7.58	7.78	7.98	8.19

Table 2—Gate voltage V_g at various drain current I_d as a function of substrate voltage V_{sub} for CD4007 PMOS transistor. The temperature ranges from 25 to 200°C in 25°C increments.

Temp (°C)	V_{sub} volts	V_g (volts) at I_d (μA) = −16	−49	−100	−169	−256	−361	−484	−625
25	0.0	−1.53	−1.70	−1.87	−2.04	−2.20	−2.37	−2.54	−2.72
25	0.4	−1.66	−1.82	−1.98	−2.15	−2.31	−2.48	−2.65	−2.82
25	1.0	−1.79	−1.96	−2.12	−2.28	−2.44	−2.60	−2.77	−2.94
25	1.8	−1.94	−2.10	−2.26	−2.41	−2.57	−2.74	−2.90	−3.07
25	3.0	−2.11	−2.27	−2.42	−2.58	−2.74	−2.90	−3.06	−3.23
25	5.0	−2.34	−2.49	−2.64	−2.80	−2.95	−3.11	−3.28	−3.45
25	8.0	−2.59	−2.74	−2.90	−3.05	−3.21	−3.37	−3.53	−3.70
25	12.0	−2.86	−3.01	−3.16	−3.32	−3.47	−3.63	−3.79	−3.96
50	0.0	−1.50	−1.69	−1.87	−2.04	−2.22	−2.40	−2.58	−2.77
50	0.4	−1.63	−1.81	−1.98	−2.16	−2.33	−2.51	−2.69	−2.87
50	1.0	−1.77	−1.94	−2.12	−2.29	−2.46	−2.64	−2.82	−3.00
50	1.8	−1.92	−2.09	−2.26	−2.43	−2.60	−2.77	−2.95	−3.13
50	3.0	−2.09	−2.26	−2.43	−2.59	−2.76	−2.94	−3.11	−3.29
50	5.0	−2.32	−2.48	−2.65	−2.81	−2.98	−3.15	−3.33	−3.51
50	8.0	−2.58	−2.74	−2.90	−3.07	−3.23	−3.41	−3.58	−3.76
50	12.0	−2.85	−3.01	−3.17	−3.34	−3.50	−3.67	−3.85	−4.02
75	0.0	−1.47	−1.67	−1.86	−2.05	−2.24	−2.43	−2.63	−2.82
75	0.4	−1.60	−1.79	−1.98	−2.16	−2.35	−2.54	−2.73	−2.93
75	1.0	−1.74	−1.93	−2.12	−2.30	−2.48	−2.67	−2.86	−3.05
75	1.8	−1.89	−2.08	−2.26	−2.44	−2.62	−2.81	−3.00	−3.19
75	3.0	−2.07	−2.25	−2.43	−2.61	−2.79	−2.97	−3.16	−3.35
75	5.0	−2.30	−2.47	−2.65	−2.83	−3.01	−3.19	−3.38	−3.57
75	8.0	−2.56	−2.73	−2.91	−3.08	−3.26	−3.44	−3.63	−3.82
75	12.0	−2.83	−3.00	−3.18	−3.35	−3.53	−3.71	−3.90	−4.09
100	0.0	−1.43	−1.64	−1.85	−2.05	−2.25	−2.46	−2.67	−2.88
100	0.4	−1.56	−1.77	−1.97	−2.17	−2.37	−2.57	−2.78	−2.98
100	1.0	−1.71	−1.91	−2.11	−2.31	−2.50	−2.70	−2.90	−3.11
100	1.8	−1.87	−2.06	−2.26	−2.45	−2.64	−2.84	−3.04	−3.25
100	3.0	−2.05	−2.24	−2.43	−2.62	−2.81	−3.01	−3.21	−3.41
100	5.0	−2.27	−2.46	−2.65	−2.84	−3.03	−3.23	−3.43	−3.63
100	8.0	−2.54	−2.72	−2.91	−3.10	−3.29	−3.48	−3.68	−3.88
100	12.0	−2.81	−2.99	−3.18	−3.37	−3.56	−3.75	−3.95	−4.15

Table 2 (Contd.)

125	0.0	−1.39	−1.62	−1.84	−2.05	−2.27	−2.48	−2.70	−2.92
125	0.4	−1.53	−1.75	−1.96	−2.17	−2.38	−2.60	−2.81	−3.03
125	1.0	−1.69	−1.90	−2.11	−2.31	−2.52	−2.73	−2.94	−3.16
125	1.8	−1.84	−2.05	−2.25	−2.46	−2.66	−2.87	−3.08	−3.29
125	3.0	−2.02	−2.22	−2.43	−2.63	−2.83	−3.04	−3.25	−3.46
125	5.0	−2.25	−2.45	−2.65	−2.85	−3.05	−3.26	−3.46	−3.68
125	8.0	−2.52	−2.71	−2.91	−3.11	−3.31	−3.51	−3.72	−3.93
125	12.0	−2.79	−2.99	−3.18	−3.38	−3.58	−3.78	−3.99	−4.20
150	0.0	−1.35	−1.59	−1.82	−2.05	−2.28	−2.51	−2.74	−2.97
150	0.4	−1.50	−1.73	−1.95	−2.18	−2.40	−2.62	−2.85	−3.08
150	1.0	−1.66	−1.88	−2.10	−2.32	−2.54	−2.76	−2.98	−3.21
150	1.8	−1.81	−2.03	−2.25	−2.46	−2.68	−2.90	−3.12	−3.34
150	3.0	−2.00	−2.21	−2.42	−2.63	−2.85	−3.07	−3.29	−3.51
150	5.0	−2.23	−2.44	−2.65	−2.86	−3.07	−3.29	−3.51	−3.73
150	8.0	−2.50	−2.70	−2.91	−3.12	−3.33	−3.54	−3.76	−3.98
150	12.0	−2.77	−2.98	−3.18	−3.39	−3.60	−3.81	−4.03	−4.25
175	0.0	−1.30	−1.56	−1.81	−2.05	−2.29	−2.53	−2.77	−3.01
175	0.4	−1.47	−1.71	−1.94	−2.18	−2.41	−2.65	−2.88	−3.12
175	1.0	−1.63	−1.86	−2.09	−2.32	−2.55	−2.78	−3.01	−3.25
175	1.8	−1.79	−2.02	−2.24	−2.47	−2.69	−2.92	−3.15	−3.39
175	3.0	−1.97	−2.20	−2.42	−2.64	−2.86	−3.09	−3.32	−3.55
175	5.0	−2.21	−2.43	−2.65	−2.86	−3.09	−3.31	−3.54	−3.77
175	8.0	−2.47	−2.69	−2.91	−3.12	−3.35	−3.57	−3.80	−4.03
175	12.0	−2.75	−2.97	−3.18	−3.40	−3.62	−3.84	−4.07	−4.30
200	0.0	−1.24	−1.52	−1.78	−2.04	−2.29	−2.54	−2.79	−3.05
200	0.4	−1.43	−1.68	−1.93	−2.17	−2.42	−2.66	−2.91	−3.16
200	1.0	−1.59	−1.84	−2.08	−2.32	−2.56	−2.80	−3.04	−3.29
200	1.8	−1.76	−2.00	−2.23	−2.47	−2.70	−2.94	−3.18	−3.43
200	3.0	−1.94	−2.18	−2.41	−2.64	−2.88	−3.11	−3.35	−3.59
200	5.0	−2.18	−2.41	−2.64	−2.87	−3.10	−3.33	−3.57	−3.81
200	8.0	−2.45	−2.67	−2.90	−3.13	−3.36	−3.59	−3.83	−4.07
200	12.0	−2.73	−2.95	−3.18	−3.40	−3.63	−3.86	−4.10	−4.34

Table 3—Reduced data for CD4007 NMOS transistor showing extrapolated threshold voltage V_t, K-factor, and error coefficient $1\text{-}R$ for the first least square fit and channel concentration N, effective surface state density N_{ss}, and error coefficient $1\text{-}Q$ for the second least square fit as a function of temperature. A matrix of 8 drain circuits and 8 substrate voltages has been used as described in the text.

Temperature (°C)	V_t (volts)	K-Factor (cm^2/Vs)	$1\text{-}R$ (ppm)	$N \times 10^{16}$ (cm^{-3})	$N_{ss} \times 10^{10}$ (cm^{-2})	$1\text{-}Q$ (ppm)
25	1.374	278	142	1.31	5.0	23
50	1.301	246	176	1.31	4.8	24
75	1.238	221	185	1.30	4.3	38
100	1.165	197	192	1.30	4.1	49
125	1.080	171	310	1.30	3.9	54
150	0.995	151	182	1.30	3.7	58
175	0.927	137	288	1.28	3.1	78
200	0.839	122	304	1.28	2.8	102

Here, w_d is the depth of carrier depletion in the channel region. Ion implantations are characterized by a range R_p and a straddle $^7\Delta R_p$

$$N(x) = N_{max} \exp\left[\frac{-(x - R_p)^2}{2(\Delta R_p)}\right]. \qquad [13]$$

The integration can be done in analytical form. For a very shallow implant, the maximum voltage shift for a dose N is qD/C_{ox}. This is the easiest case because the addition or subtraction can be done prior to going to second least square fit. Fortunately, most implants for this purpose are rather shallow.

Discussion

The described technique is a quick and powerful way of determining experimentally the factors that influence threshold voltage. By monitoring a particular MOS fabrication line, line control can be measured and fabrication tolerances can be narrowed. This, in turn, allows reduced tolerances in circuit design and ultimately the fabrication of denser integrated circuits.

In the last few years, the effective surface state density N_{ss} has become very small, as MOS fabrication technology has advanced, and is presently in the low 10^{10} cm^{-2} range. The errors in N_{ss} become large as N_{ss} becomes smaller than the 10^{11} cm^{-2}. The qN_{ss}/C_{ox} term for 1000 Å of oxide and a value of $N_{ss} = 3 \times 10^{10}$ cm^{-2} is 142 millivolts. This term is getting smaller with the thinner oxides that are now being used. The reduction in N_{ss} in Table 1 with temperature for the NMOS transistor may be a real effect. However, any residual mobile charge in the oxide, such as sodium ions, will contribute as will any hole traps that are charged. N_{ss} should, in this context, only be regarded as an effective surface state

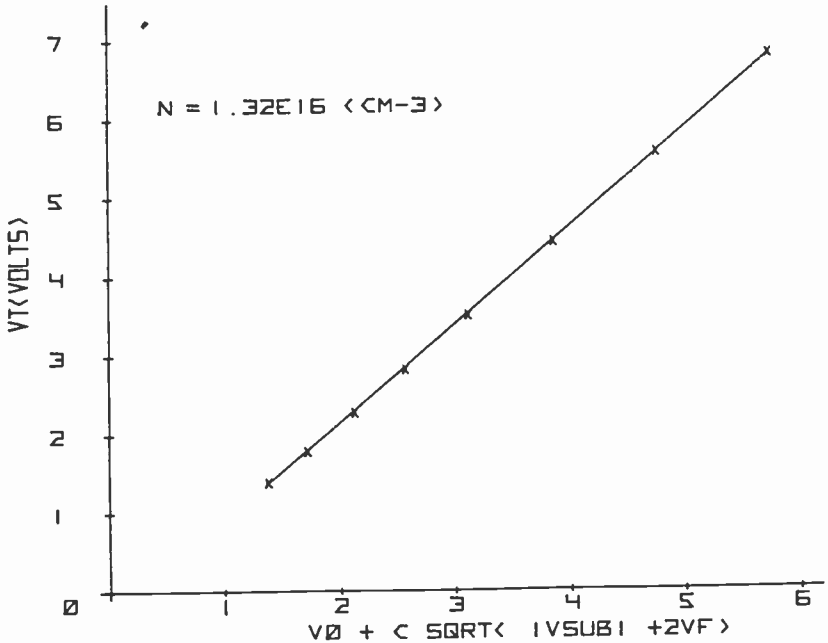


Fig. 4—Plot of Eq. [2] for NMOS transistor at room temperature. The least square fit is excellent.

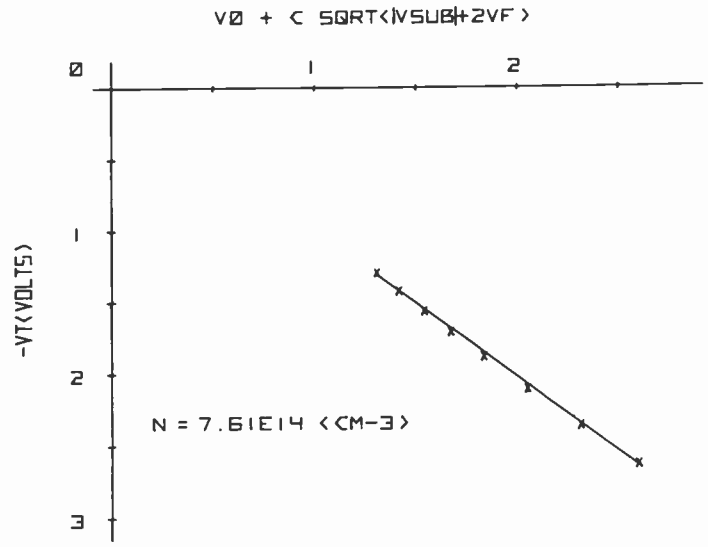


Fig. 5—Plot of Eq. [3] for PMOS transistor at room temperature.

Table 4—Reduced data for CD4007 PMOS transitor showing extrapolated threshold voltage V_t, K-factor, and error coefficient 1-R for the first least square fit and channel concentration N, effective surface state density N_{ss}, and error coefficient 1-Q for the second least square fit as a function of temperature. A matrix of 8 drain currents and 8 substrate voltage has been used.

Temperature (°C)	V_t (volts)	K-Factor (cm^2/Vs)	1-R (ppm)	$N \times 10^{14}$ (cm^{-3})	$N_{ss} \times 10^{10}$ (cm^{-2})	1-Q (ppm)
25	−1.305	315	58	7.61	4.9	755
50	−1.264	278	69	7.61	5.0	527
75	−1.217	244	23	7.60	5.0	534
100	−1.157	211	32	7.67	4.9	597
125	−1.109	191	44	7.66	4.9	451
150	−1.048	169	13	7.71	4.9	428
175	−0.990	152	80	7.67	4.9	486
200	−0.916	137	124	7.77	4.7	514

density that, for experimental purposes, includes the above mentioned charges.

Acknowledgments

The measurements were skillfully taken by J. Brady. Technical discussions were held with C. Davis regarding the comparison of his nonlinear regression approach with the dual linear regression used here. D. S. Woo and J. Yeh programmed this approach for another dc tester.

References

[1] J. E. Meyer, "MOS Models and Circuit Simulation," *RCA Review*, **32**, p 42, March (1971).

[2] P. Richman, *MOS Field Effect Transistors and Integrated Circuits*, Wiley-Interscience, New York (1973).

[3] Hewlett-Packard Calculator Stat-Pac, Vol. 1, Part #09100-70800, Program IV-1.

[4] L. Vadasz and A. S. Grove, "Temperature Dependence of MOS Transistor Characteristics Below Saturation," *IEEE Trans. on Electron Devices*, **ED-13**, p. 863 (1966).

[5] V. L. Rideout, F. H. Gaensslen, and A. LeBlanc, "Device Design Considerations for Ion Implanted n-Channel MOSFET's," *IBM J. Research and Development*, p. 50 (1975).

[6] R. R. Troutman, "Ion-Implanted Threshold Tailoring for Insulated Gate Field-Effect Transistors," *IEEE Trans. on Electron Devices*, **ED-24**, p. 182 (1976).

[7] J. Lindhard and M. Scharff, "Energy Dissipation by Ions in the kev Region," *Phys. Rev.*, **124**, p. 128 (1961).

Low Frequency Excess Noise in SOS MOS FET's

S. T. Hsu

RCA Laboratories, Princeton, NJ 08540

Abstract—Noise in SOS field-effect transistors has been investigated in the 20 Hz to 1 MHz frequency range. The measurements were made with frequency, drain bias voltage, gate bias voltage, and temperature as parameters. The low-frequency excess noise source was identified as electron trap states in the thin silicon film. The electron trap states are located at 0.65 eV below the edge of the conduction band. The effective barrier height of the electron trap state was found to decrease with increasing drain bias voltage. The experimental results also indicate that in thin silicon films grown on sapphire substrates, the densities of hole trap states and generation recombination centers are very small compared to that of electron trap states.

1. Introduction

SOS MOS field-effect transistors are thin-film field-effect devices, employing a thin layer of silicon film grown on a sapphire substrate. With this structure, air isolation can be easily applied, and the parasitic capacitances at the substrate and the isolation regions are eliminated. Consequently, SOS devices have better frequency response than devices made on a bulk semiconductor substrate. Noise has been a major problem for SOS linear circuits applications, however. In this paper, we present the results of our recent study on noise properties of SOS MOS field-effect transistors, in which we identified the low-frequency noise sources of the devices. The results also provide information for the low-noise application of existing devices.

2. Experiments

The thickness of the silicon film of the SOS MOS FET's used in this study was 1 μm. The impurity concentration of both p- and n-channel devices are of the order of 10^{15} cm^{-3}. The SiO_2 film of the device was 0.1 μm thick. The source-to-drain distance of the devices used varied from 5 μm to 25 μm. Both linear and circular geometry devices were used. The general noise characteristics of the devices were found to be independent of the device geometry.

2.1 Noise in N-Channel SOS MOS FET's

It is well understood that the drain-current/voltage characteristics of an n-channel SOS MOS FET exhibits current kinks.[1] The onset of the current kink increases as the drain bias voltage is increased. In general, for a given transistor the voltage between the onset of the current saturation and the onset of the current kink is approximately independent of the gate bias voltage. Noise spectra of the device were measured at several drain bias voltages. The results are plotted in Fig. 1. In this experiment the gate bias voltage was kept constant at 6 V. The figure

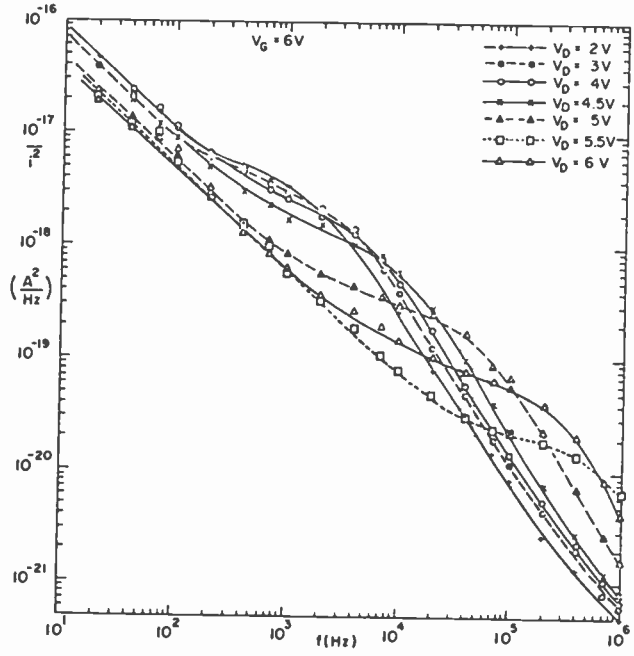


Fig. 1—Noise spectra with drain bias voltage as a parameter of a typical n-channel SOS MOS FET.

clearly shows that at very low frequencies the noise spectrum follows a $1/f$ frequency variation. In the intermediate frequency region, the noise power decreases only very slightly with increasing frequency. At high frequency, the noise spectrum has $f^{-\alpha}$ frequency dependence, with α nearly equal to 2. The limiting noise[2] of the device is below 10^{-22} A^2/Hz, which is negligibly small compared to the plateau of the noise spectrum. The low-frequency noise of the device, therefore, can be considered as a superposition of a $1/f$ noise component and a generation-recombination (g-r) noise component. As we will see later, in general, the g-r noise has a distributed time constant. In this particular case, however, a single time constant approximation is appropriate. This is consistent with our early observations.[3] We will refer to the g-r noise component as excess noise.

We define the frequency where the amplitude of the g-r noise decreases 3 dB from the amplitude at the low-frequency plateau as the corner frequency of the g-r noise. It is seen from Fig. 1 that the corner frequency increases very rapidly as the drain bias voltage is increased. The time constant of the g-r noise is, therefore, strongly dependent on the field intensity caused by the drain voltage. From Fig. 1, it is obvious that at larger drain bias voltages no excess noise can be observed. This fact strongly suggests that the excess noise is due to a trapping effect rather than to recombination or generation of electron–hole pairs. This will be seen more clearly when we present additional experimental results.

The corner frequency of the trapping noise taken from Fig. 1 is plotted as a function of drain bias voltage in Fig. 2. At small biases the corner frequency increases slightly with increasing drain bias voltage. At large bias, the corner frequency increases almost exponentially with drain bias voltage. Mathematically, the experimental result can be written

$$\omega_C = \omega_{C10} + \omega_{C20} \exp(V_D/V_{00}), \qquad [1]$$

where ω_{C10}, ω_{C20} and V_{00} are constants and V_D is the drain bias voltage.

The low-frequency plateau of the trapping noise as a function of drain bias voltage is also plotted in Fig. 2. The noise plateau decreases with increasing drain bias voltage. The product of the plateau and the corner frequency of the trapping noise increases slightly with increasing drain bias voltage. This product is proportional to the number of the effective trap or generation-recombination centers.[2,4–6]

The gate voltage dependence of the low-frequency noise is plotted in Figs. 3(a) and 3(b) for drain voltages of 3 and 5 V, respectively. The shape of the family of curves in Fig. 3(a) is different from that in Fig. 3(b) due to the field intensity difference in these two experiments. In general, the

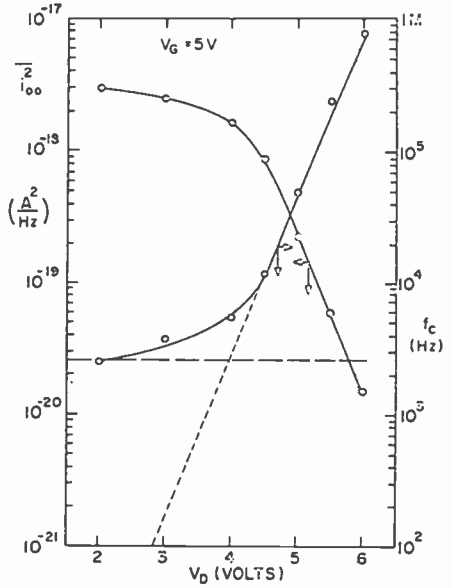


Fig. 2—Drain bias dependence of the corner frequency and the low-frequency plateau of g-r noise in Fig. 1.

trapping noise exhibits a distributed time constant. Thus, the high-frequency tail of the trapping noise has $f^{-\alpha}$ frequency dependence, with α less than two. This is particularly true when the gate voltage is large. At small gate biases the trapping noise may be approximated as having a single time constant.

Generation-recombination and trapping noise has a large temperature dependence. The corner frequency of the trapping noise possesses a temperature activation energy. To investigate the temperature dependence of the trapping noise, noise spectra of several SOS MOS FET's were measured as a function of temperature. Noise spectra of a typical device are plotted in Fig. 4 with temperature as a parameter. As expected, the corner frequency of the trapping noise spectrum increases very rapidly with increasing temperature. On the other hand, the low-frequency plateau decreases as the temperature is increased.

In order to obtain the activation energy of the trap states, the corner frequency of the trapping noise spectrum was plotted, as shown in Fig. 5, as a function of $1/T$ for devices with different channel lengths. The channel length of the device was found to have no effect on the corner frequency as long as the drain bias voltage is small. In Fig. 5a the drain bias voltage is 2 V and the channel lengths of the devices used are 5, 10, and 25 μm. The experimental points follow a straight line. From the slope

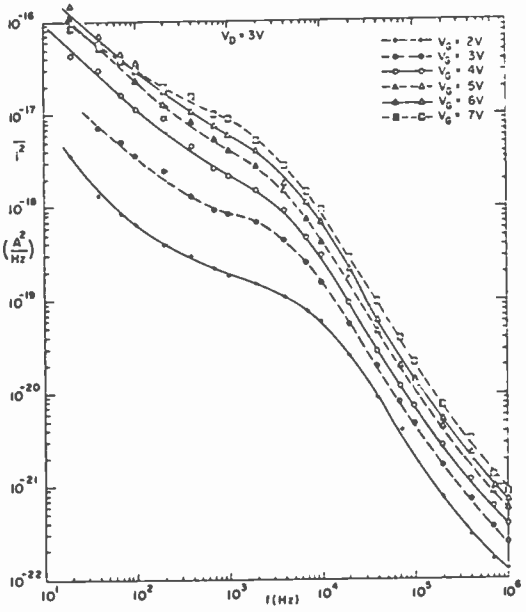


Fig. 3(a)—Noise spectra with gate bias voltage as a parameter of a typical n-channel SOS MOS FET (V_D = 3 V).

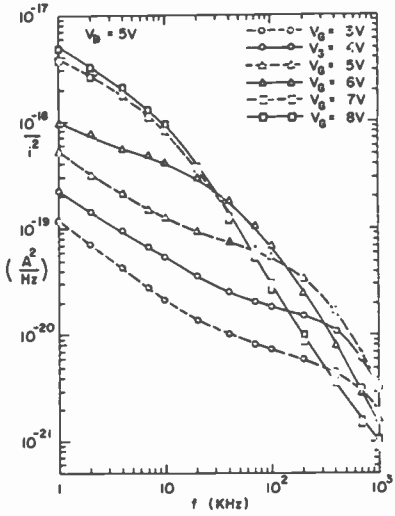


Fig. 3(b)—Noise spectra with gate bias voltage as a parameter of a typical n-channel SOS MOS FET (V_D = 5 V).

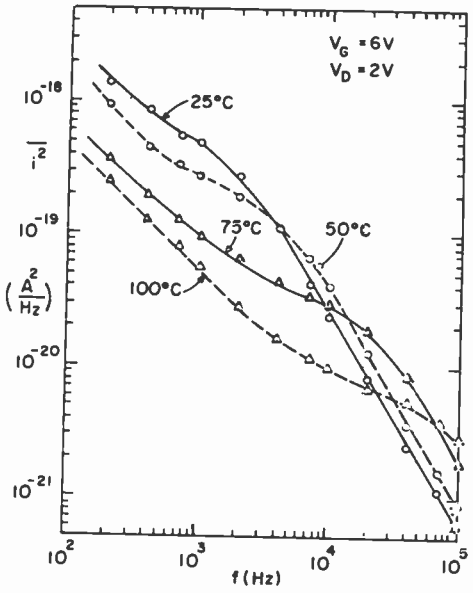


Fig. 4—Noise spectra of a typical n-channel SOS MOS FET with temperature as a parameter.

of the straight line, the activation energy was found to be equal to 0.65 eV. It is found that this activation energy decreases with the increasing drain bias voltage. Fig. 5(b) shows the corner frequency as a function of $1/T$ of one of the device used in Fig. 5(a). The channel length of the device is 5 μm and the drain bias voltage in this experiment is 4 V. The activation energy was found to be equal to 0.45 eV. This demonstrates that the effective energy barrier of the trap states decreases with increasing drain bias voltage.

2.2 Noise in P-Channel SOS MOST's

The drain-current/voltage characteristic of p-channel SOS MOS FET's exhibits no current kink. This difference in dc characteristics also leads to a difference in noise properties between n-channel and p-channel SOS MOS FET's. The I-V characteristic is very similar to that of a MOS FET made on a bulk semiconductor. Fig. 6 shows noise spectra of the FET with drain bias voltages of -3, -10 and -13.5 V. The gate bias voltage was kept constant at -4 V. The figure clearly shows that a p-channel SOS MOS FET also exhibits low-frequency excess noise. The low-frequency noise spectrum can be represented by the superposition of $1/f$ noise and g-r or trapping noise with distributed time constant. In Fig. 7 noise

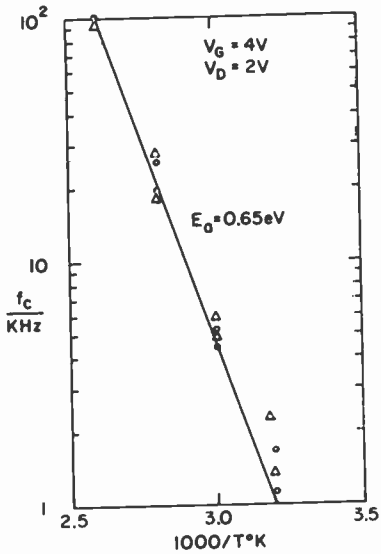


Fig. 5(a)—Corner frequency of the trap noise spectrum as a function of 1000/T°K for V_D = 2 V.

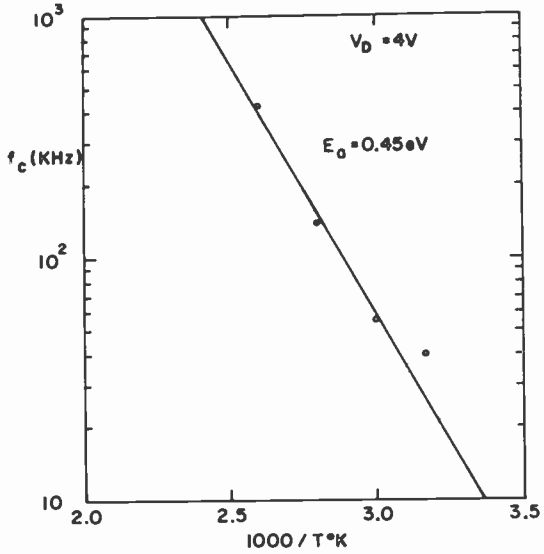


Fig. 5(b)—Corner frequency of the trap noise spectrum as a function of 1000/T°K for V_D = 4 V.

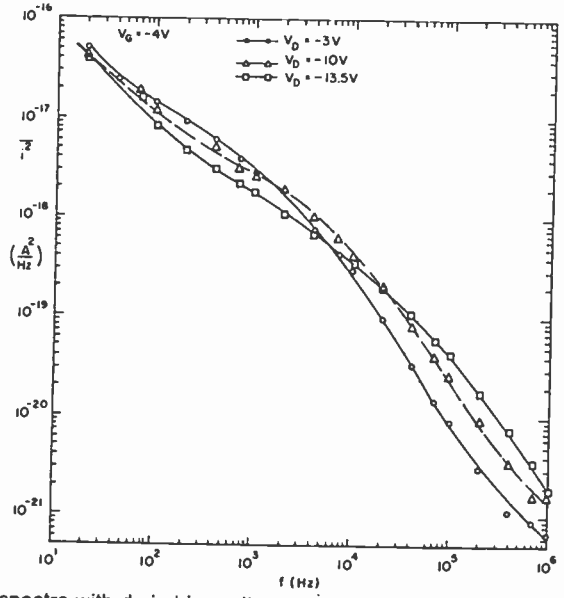


Fig. 6—Noise spectra with drain bias voltage as a parameter of a typical p-channel SOS MOS FET.

spectra of the p-channel SOS FET are plotted with gate bias voltage as a parameter.

The curves in Figs. 6 and 7 have three distinguishing regions. If we write $\overline{i^2} \propto f^{-\alpha}$, $\alpha \simeq 1$ at very low frequencies. In the intermediate frequency range, $\alpha < 1$; and at high frequencies, $1.5 < \alpha < 2$. If we extrapolate the straight lines of the intermediate and high-frequency regions and call the frequency at the intersect point of these two straight lines the corner frequency, f_c, we find that f_c increases slightly with increasing drain bias voltage. In this particular case, f_c is equal to 3 KHz and 25 KHz for drain bias voltages of -3 and -13.5 V, respectively. In the case of n-channel devices when the drain bias voltage is larger than 10 V, trapping noise has not been observed at room temperature in the frequency region below 0.1 MHz. Fig. 7 clearly indicates that the corner frequency is independent of the gate bias voltage.

The temperature dependence of low-frequency excess noise in p-channel SOS MOS FET's is also expected to be large. Fig. 8 shows noise spectra of a p-channel device with temperature as a parameter. The corner frequency does increase very rapidly with increasing temperature. In order to obtain the activation energy of the trap states, the corner frequency of the trapping noise spectrum was plotted as a function of $1/T$, as shown in Fig. 9. The experimental points follow a straight line.

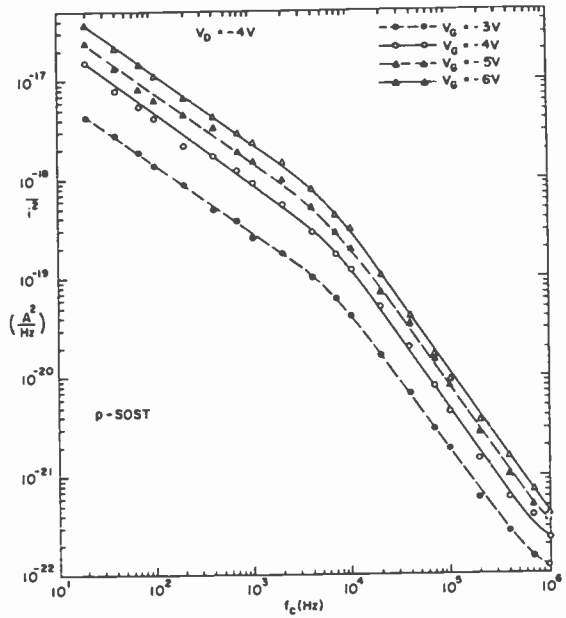


Fig. 7—Noise spectrum with gate bias voltage as a parameter of a typical p-channel SOS MOS FET.

From the slope of this straight line, the activation energy was found to be equal to 0.65 eV.

3. Discussion

The experimental results presented in the previous section clearly indicate that low-frequency excess noise in SOS MOS FET's is due to electron trap states in the thin silicon film. It is the electron trap state that produces excess noise in both n-channel and p-channel SOS MOS FET's. The same electron trap states are also responsible for current kinks in n-channel SOS MOS FET's.

It has been shown independently by Sah[7] and Laurizen[8] that the fluctuation of charge occupancy of a generation center in the depletion layer of a field-effect transistor modulates the channel voltage and, hence, produces noise in the device. Van der Ziel[9] shows that g-r centers most subject to fluctuation are half occupied. Hsu[10] showed that recombination noise is dominated by the nearly half occupied recombination centers if the center has equal capture cross-section for electrons and holes and that, for a first-order approximation, the recombination noise may be considered as caused by the nearly half occupied states only.

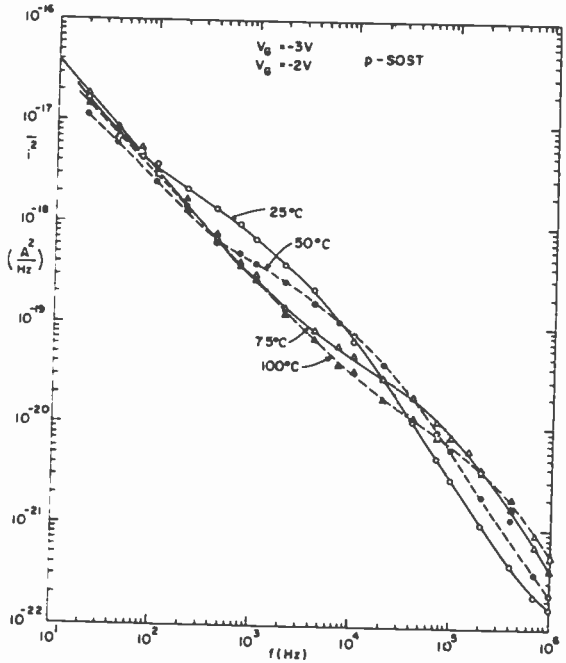


Fig. 8—Temperature dependence of low-frequency excess noise of a p-channel SOS MOS FET.

Since a trapping process is a special case of a generation-recombination process, the theories for g-r noise are directly applicable to trapping noise. However, an ideal electron trap state has zero hole capture cross-section. As a result the electron trap states most subject to fluctuation are not half occupied. From g-r noise theory[4] it can be shown that

$$\tau = \frac{1 - f_t}{C_1 n_1}, \qquad [2]$$

where C_1 is a constant, n_1 is the free electron density when the electron Fermi level is located at the trap energy level, and f_t is the probability for trap states to be occupied by electrons. The variance of the trap state charge density fluctuation is

$$\text{Var } n_t = N_t f_t (1 - f_t). \qquad [3]$$

The spectral intensity, $S_{n_t}(\omega)$, is therefore given by

$$S_{n_t}(\omega) = \frac{4\tau \text{ Var } n_t}{1 + \omega^2\tau^2} = \frac{\frac{4N_t}{C_1 n_1} f_t (1 - f_t)^2}{1 + \frac{\omega^2}{C_1^2 n_1^2}(1 - f_t)^2}. \qquad [4]$$

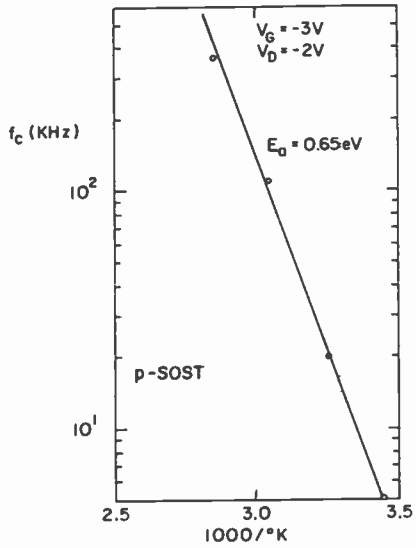


Fig. 9—Corner frequency of the trap noise spectrum of Fig. 8 as a function of $1000/T$ °K.

Differentiating Eq. [4] with respect to f_t and setting the result equal to zero yields $f_t = 1/3$. Thus, the electron trap states most subject to fluctuation are 1/3 occupied. From Figs. 5 and 9 we immediately identify that the trap state is located at 0.65 eV below the edge of the conduction band. If the noise is dominated by the trap states with a 0.65 eV energy below the edge of the conduction band located in the channel depletion region, one expects the output noise of n-channel devices to have a very small voltage dependence. However, at the drain depletion region, the 1/3 occupied electron trap states are located very close to the $Si\text{-}SiO_2$ interface for n-channel devices. The field intensity at the sites of those trap states increases with increasing drain bias voltage.

For p-channel devices, the 1/3 occupied electron trap states in the drain depletion region are far away from the $Si\text{-}SiO_2$ interface. The field intensity at those trap states can hardly be changed by the drain bias voltage. Figs. 1, 6, and 7 strongly suggest that the excess noise of the devices investigated is due to electron trap states in the drain depletion region. This is also supported by the fact that a small change of the voltage at the channel depletion region will produce a very small change in output current. However, for the same amount of voltage change at the pinch-off point, much larger change in output current can be obtained. The output noise of n-channel SOS MOS FET's is, therefore, due to electron trap states near the pinch-off region or in the drain-depletion region. The voltage at the pinch-off point fluctuates as free electrons hop

through the electron trap states from the conducting channel to the drain diffusion region.

The corner frequency, f_c, of the trapping noise spectrum is given by

$$\omega_c = \frac{1}{\tau_1} + \frac{1}{\tau_2}, \qquad [5]$$

where τ_1 is the trap empty time and τ_2 is the trap occupation time. The time constants τ_1 and τ_2 are given by[11]

$$\frac{1}{\tau_1} = \upsilon S n \text{ and } \frac{1}{\tau_2} = \mu_0 \exp\left(-\frac{qE_t}{kT}\right), \qquad [6]$$

where υ is the thermal velocity of electrons, S is the capture cross section of the trap states, μ_0 is the vibration frequency, and E_t is the energy of the trap state. The time constant τ_1 is therefore decreased with increasing gate bias voltage. The time constant τ_2 increases exponentially with the trap state energy. If

$$\tau_2 \ll \tau_1 \text{ or } \mu_0^{-1} \exp\left(\frac{qE_t}{kT}\right) \ll \upsilon S n, \qquad [7]$$

then

$$\omega_c = \mu_0 \exp\left(-\frac{qE_t}{kT}\right). \qquad [8]$$

This condition can be best satisfied at higher temperature. Thus, when Eq. [4] is satisfied, a plot of $\log\omega_c$ versus $1/T$ yields a straight line. The energy of the trap state can be obtained from the slope of this plot. These relationships bear out the experimental results of Figs. 5 and 9.

The results in Fig. 2 clearly show that the trap energy is strongly dependent on drain bias voltage. This can be explained as a lowering of the trap energy barrier by the applied field. The electric field intensity at the trap states increases as the drain bias voltage is increased. This field intensity reduces the trap state energy barriers such that[12,13]

$$E_{t_{eff}} = E_{t_0} - \sqrt{qF/\epsilon\pi}\,, \qquad [9]$$

where $E_{t_{eff}}$ is the effective barrier height, E_{t_0} is the trap barrier height at zero field intensity, and F is the field intensity at the trap site. The trap barrier lowering effect is best shown in Fig. 5. Thus, when the drain bias voltage is increased from 2 to 4 V, the trap energy barrier is lowered from 0.65 to 0.45 eV.

The field intensity, F, is a complex function of voltage. We will not try to obtain an accurate expression of F here, but will accept the experimental fact that the trap occupation time decreases exponentially

with drain bias voltage. In fact, the experimental results show that $1/\tau_2$ is proportional to $\exp(V_D/V_\infty)^\beta$ where $0.5 \leq \beta \leq 1$. An accurate value of β cannot be determined from our noise data, however. Fig. 2 also provides the information that the trap Fermi level decreases with increasing drain bias voltage. Let N_t be the trap state density and n_t the density of trapped electrons. From Fig. 2,

$$n_t = N_t \frac{\tau_1}{\tau_1 + \tau_2} = N_t \frac{\tau_{20} \exp(-V_D/V_\infty)}{\tau_1 + \tau_{20} \exp(-V_D/V_\infty)}. \quad [10]$$

If E_t is the trap energy and E_{F_t} is the trap quasi-Fermi level, it can be shown that

$$E_{F_t} = E_t + kT[\log(\tau_{20}/\tau_1) - (V_D/V_\infty)]. \quad [11]$$

The trap quasi-Fermi level is, therefore, decreased with the increasing drain bias voltage. At small drain biases, most of the trap states are occupied by electrons, while at large biases, most of the electron trap states are empty.

Strictly speaking, no single time constant for the trap noise spectrum exists in the present case. The field intensity and electron density are not uniformly distributed. The characteristic noise time constant increases as the distance between the point under consideration and the source is decreased. However, only those nearly ⅓-occupied electron trap states located near the pinch-off point are important. The dispersion of the time constant is, therefore, small. When the gate voltage is small, charge density in the channel is small. The free electrons flow very close to the SiO_2-Si interface. The time constant dispersion is therefore very small. If the gate bias voltage is increased, the density of free electron is also increased. There is a large gradient of electron density normal to the Si-SiO_2 interface. The actual channel width spreads away from the Si-SiO_2 interface, and the time constant dispersion is increased. Consequently, a nearly single time constant for the trapping noise spectrum can be observed only when the gate bias voltage is small. As the gate bias voltage is increased, the free electron density increases and the field intensity near the pinch-off region decreases. Accordingly, the time constant is larger and the corner frequency of the trap noise spectrum is smaller, as shown in Fig. 3.

In a p-channel SOS MOS FET the current carriers are holes. Since electron trap states cannot trap holes, the electron trap states located near the pinch-off region close to Si-SiO_2 interface contribute no noise to the device. The fluctuation of charge density of the trap states in the space-charge region near the pinch-off point may produce a large amount of noise in the device. As the gate bias voltage is unable to change the time constant of these electron trap states, the corner frequency of the

trap noise spectrum is independent of the gate bias voltage. If the silicon film is very thin, the entire film may be depleted and there may be no 1/3-occupied electron trap states existing in the channel depletion region. The drain bias voltage is able to increase field intensity to the nearly 1/3-occupied electron trap states in the depletion region. However, this effect is expected to be small. The corner frequency is, therefore, increased only slightly with increasing drain bias voltage.

Neither g-r noise nor hole trapping noise has been observed in our devices, which indicates that near the Si-SiO_2 interface the density of both hole trap states and g-r centers are very small compared to the density of electron trap states. However, one may not therefore draw the conclusion that the hole lifetime in SOS film must be long. The explanation is that electron trap states may act as g-r centers. An electron in the conduction band may be trapped by an electron trap state and subsequently released to the valence band to recombine with a hole. This process may cause free holes in a SOS film to have a very short lifetime.

The imperfect states in silicon films of SOS structures with 0.65 eV energy was discovered by Heiman in 1967.[14] He investigated the photoconductivity of SOS films and found that there was a large photocurrent for an incident photon energy of 0.65 eV. The trapping noise we observed is most likely due to the same set of imperfect centers. The noise data indicate that the imperfect centers in our devices are electron trap states, instead of acceptor traps as found by Heiman.

4. Conclusions

The experimental results indicate that low-frequency excess noise in investigated SOS MOS FET's is due to electron trap states located inside the thin silicon film. The energy of the electron trap states is 0.65 eV below the edge of the conduction band. The trap energy barrier can be lowered by the applied electric field.

To minimize the low-frequency excess noise in SOS MOS FET's, a high quality, low-trap-density silicon film is essential. As the trap density is reduced, the low frequency noise would be dominated by surface noise. The low-frequency noise of SOS MOS FET's is, therefore, expected to be the same as that of MOS FET's built on bulk semiconductor substrates. Although the existing SOS MOS FET's exhibit a large amount of low frequency excess noise, they may be used as low-noise amplifiers. The device noise can be as low as that of bulk MOS FET's in a given frequency range. This is particularly true for n-channel devices as is clearly shown in Fig. 2. If the drain bias voltage is larger than the onset voltage of the current kink, the trap noise is not important at low

frequencies. If the drain bias voltage is smaller than the onset voltage of the current kink, the trap noise is important only in the frequency range below a few tens of kilohertz.

It is also concluded that in the SOS films, the densities of g-r centers and hole trap states are negligibly small compared to the density of electron trap states. Although an electron trap state cannot trap a hole, it may act as a g-r center by trapping an electron and subsequently releasing the captured electron to the valence band. This process may cause a very short hole lifetime.

Acknowledgment

The author is deeply grateful to J. H. Scott, Jr., and G. B. Herzog for their interest and encouragement of this work. He is also grateful to W. E. Ham for providing the samples used and L. C. Lucas for the noise measurements.

References:

[1] S. T. Hsu, "Drain Characteristic of Thin-Film MOS FET's," *RCA Rev.*, **38**, p. 139, March 1977.

[2] A. van der Ziel, *Noise: Sources, Characterization, Measurement*, Chapter 5, p. 69, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1970).

[3] S. T. Hsu, "Electron Trapping Noise in SOS MOS Field-Effect Transistors Operated in the Linear Region," *RCA Rev.*, **38**, p. 226, June 1977.

[4] K. M. van Vliet and J. R. Fassett, *Fluctuation Phenomena in Solids*, Ed. by R. E. Burgess, Chapter 7, Academic Press, N.Y. (1965).

[5] S. T. Hsu, "Surface State Related 1/*f* Noise in MOS Transistors," *Solid-State Electronics*, **13**, p. 1451 (1970).

[6] S. T. Hsu, "Surface State Related 1/*f* Noise in p-n Junctions," *Solid-State Electronics*, **13**, p. 843 (1970).

[7] C. T. Sah, "Theory of Low-Frequency Generation Noise in Junction-Gate Field-Effect Transistors," *Proc. IEEE*, **52**, p. 795 (1964).

[8] P. O. Lauritzen, "Low-Frequency Generation Noise in Function Field-Effect Transistors," *Solid-State Electronics*, **8**, p. 41 (1965).

[9] A. van der Ziel, *Fluctuation Phenomena in Semiconductors*, p. 25, Butterworth, London (1959).

[10] S. T. Hsu, "Low-Frequency Excess Noise in Metal-Silicon Schottky Barrier Diodes," *IEEE Trans. Electron Devices*, **ED-17**, p. 496 (1970).

[11] A. van der Ziel, *Solid-State Physical Electronics*, 2nd Ed., Prentice-Hall, Inc., Englewood Cliffs, N.Y., Chapter 6, p. 122 (1968).

[12] J. Frenkel, "On Pre-Breakdown Phenomena in Insulators and Electronic Semiconductors," *Phys. Rev.*, **54**, p. 647 (1938).

[13] S. M. Sze, *Physics of Semiconductor Devices*, Chapter 9, p. 496, John Wiley and Sons, Inc., N.Y. (1969).

[14] F. P. Heiman, "Donor Surface States and Bulk Acceptor Traps in Silicon-on-Sapphire Films," *Appl. Phys. Lett.*, **11**, p. 132 (1967).

Review of Factors Affecting Warpage of Silicon Wafers

D. Thebault and L. Jastrzebski

RCA Laboratories, Princeton, NJ 08540

Abstract—One of the most serious and difficult problems in silicon device technology is warpage or bowing of silicon wafers during device processing. Warpage of silicon wafers is a result of plastic deformation which takes place at elevated temperatures due to the presence of stresses. The origin of stress forces are reviewed, and ways of minimizing them by optimizing processing are discussed. The mechanical properties of silicon wafers (their resistance to warpage) depend on surface preparation, resistivity, dislocation density and especially oxygen content in the wafer. The effect of these parameters on material properties is reviewed in this work.

One significant conclusion evident from published data is that the oxygen content and its chemical form are important factors affecting the mechanical properties of silicon substrates. Silicate precipitates or low oxygen content reduce resistance to warpage. It seems that a certain state of oxygen and range of concentration exist for which oxygen has a beneficial effect on the mechanical properties. Published data on the effect of oxygen on mechanical properties at the present time are outlinec

1. Introduction

Warpage or bowing of silicon wafers observed after high temperature device processing is a common and undesirable phenomenon. Warped wafers are difficult to handle in automatic processing equipment. The generation of complex circuit patterns by photolithography on a warped or bowed wafer is difficult to perform due to a varying plane of focus. This problem is especially serious with projection printing, during which

wafer flatness is no longer improved by application of mechanical pressure to a wafer with a mask, as occurs during contact printing. To print a 3 μm line with 0.5 μm tolerance with projection printing, wafer bowing or warpage can not exceed 10 μm.[1] VLSI technology imposes new, severe demands on flatness of processed wafers to print micron and submicron dimensioned patterns. Also, the necessity to process four- and five-inch diameter wafers makes warpage and bowing a "new" rediscovered topic of interest. In addition to mechanical and photolithographic problems, warpage can also affect the electric characteristics of devices and circuits. Dislocations introduced into processed wafers as a result of plastic deformation are known to cause excessive leakage current in processed circuits.[2] Due to a change of lateral dimensions on the wafer by warpage, a device geometry (e.g., a channel length) can be distorted which can cause degradation of device characteristics.[3]

Warpage or bowing of silicon wafers is a result of plastic deformation which occurs at an elevated temperature during device processing. This processing may lead to stress causing forces generated by (1) thermal stress introduced during fast heating or cooling (too fast pulling and pushing rates),[4] (2) strain introduced into wafers by diffused regions[5] (caused by lattice mismatch), or (3) strain introduced by polysilicon or nitride layers deposited on a silicon wafer.[6,7] These stress forces can be minimized by an optimization of processing conditions and methods, although warpage can be observed even with the most careful processing. Appearance of warpage in some wafers when it is not present in the others subject to the same processing raises the question of material resistance to warpage. Our understanding of this problem is not complete, but the role of material parameters such as the presence of dislocations, surface preparation, and the amount of oxygen in silicon has been investigated in the recent years.

Although contradictory information about the effect of oxygen on mechanical properties of silicon wafers has been published in the literature,[8–21] in general it can be concluded that the presence of oxygen precipitation in the silicon wafers has an adverse effect on the mechanical properties of silicon.[10–14] On the other hand, the presence of only subppma amounts of oxygen in float-zone wafers also decrease their resistance to bowing or warpage.[21] It seems that a certain range of oxygen concentration exists for which oxygen has a beneficial effect on the mechanical properties of silicon. Information available on this subject in the literature is summarized in this paper. We will outline our present understanding of the warpage problem and point out where knowledge is still lacking and where comprehensive understanding is necessary to quantitatively predict the mechanical properties of silicon wafers during processing.

2. Mechanical Properties

The mechanical properties of a material are described by a deformation curve (i.e., stress versus displacement). In Fig. 1 the tensile stress σ is plotted against the percent of the crosshead displacement divided by the initial gauge length of the specimen. For small displacements, the stress-strain curve is essentially linear until the stress value reaches σ_y. In this region, a specimen deforms elastically obeying Hooke's law, and when the applied forces are removed, the specimen returns to the previous shape. When the stress reaches σ_y, however, the stress-strain curve is no longer linear and the specimen starts to deform plastically. Dislocations are introduced into the sample and the sample will not return to its previous shape after removal of the external forces. After plastic deformation occurs, the stress reaches a maximum value σ_m (also called the upper yield point) after which it drops to σ_f, the flow stress. The larger the value of σ_m, the more stress is required to warp a silicon wafer. The value of σ_m is a strong function of material parameters such as dislocation density in the specimen,[8–12] oxygen level,[10,11,20] resistivity,[22,23] or surface preparation.[24] This will be discussed in more detail later. Also, the value of σ_m decreases with increasing temperature as is shown in Fig. 2.

It is known that in silicon, dislocation generation and propagation take place in certain crystallographic planes, called slip planes.[25] The parameter that describes directly the susceptibility of a silicon wafer to warp is τ_c the critical resolved shear stress. When a shear stress acting in the slip plane exceeds τ_c, dislocations will be generated in silicon in the (111) family of planes in a [110] direction. The relationship between

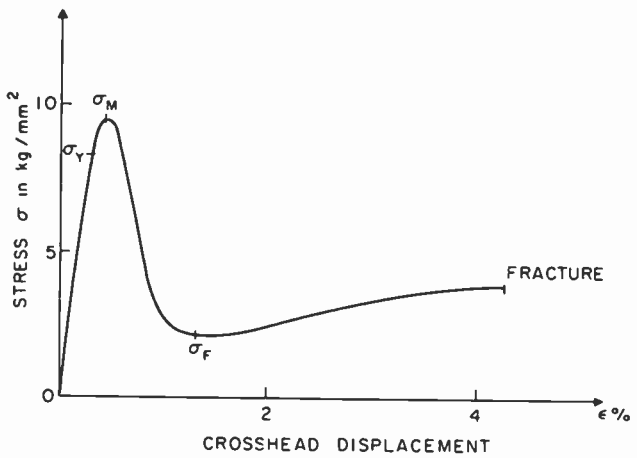


Fig. 1—Typically stress-strain curve for silicon crystal.

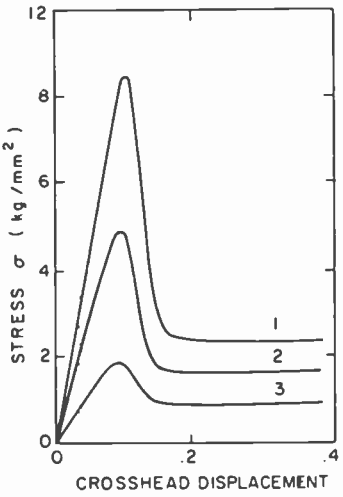


Fig. 2—Stress-strain curve at various temperatures (1) $T = 800$, (2) $T = 850$, (3) $T = 950$°C (after Ref. [8]).

shear stress acting along the plane of slip and stress applied to a specimen is given by Schmid's law:[25]

$$\tau = \sigma \cos \phi \cos \lambda,$$

where λ is an angle between the slip direction and the axis of tension and ϕ is the angle between the normal to the slip plane and the axis of tension (Fig. 3). To improve resistance to warpage, wafers with as high a value of τ_c as possible should be used.

3. Stress Forces in Device Processing

3.1 Temperature gradient across the wafer

When closely spaced silicon wafers are pulled out from a furnace at typically 900°C the wafer periphery cools faster than the wafer center and a radial temperature gradient develops.[4,16] The thermal stress introduced by this radial temperature gradient can be large enough to cause a plastical deformation.[4,24,26] The degree of thermal stress will be affected by the spacing between the wafers, the type of furnace container or boat used, the heating or cooling rates, and the furnace temperature. Obviously, temperature gradients are larger for wafers with larger diameters. For example, for the same cooling conditions, warpage of 75-mm wafers is a factor of two larger than for 57.2-mm wafers, as shown in Fig. 4. The figure shows data for wafers withdrawn rapidly in two seconds

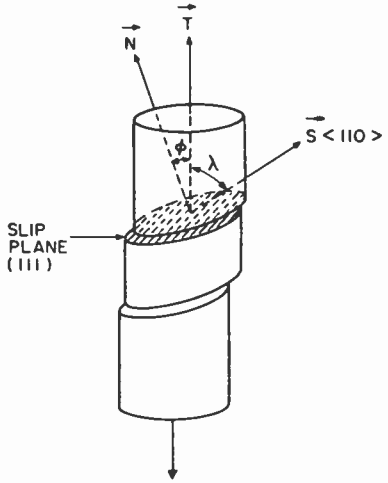


Fig. 3—Illustration of the geometry of slip in silicon crystal. $\vec{S}$ is the slip direction, $\vec{T}$ the axis of tension, $\vec{N}$ the slip plane normal.

from a 1200°C temperature furnace.[24] A detailed model has been developed for rapidly cooled wafers,[16,27] and it is interesting to note that for this model, the warped area on a concave side is always larger than on a convex side. Because the concave side is the most compressed region, dislocations are generated there first to relieve stress.

Of the boats used, the smallest warpage has been observed in those with a top and a metal reflector (see Fig. 4). In such a boat, the presence of the reflector reduces temperature gradient across the wafer. Other procedures commonly used to eliminate stress introduced by temperature gradients are slow pushing and pulling of wafers into and out of a furnace during high temperature processing and slow ramping of a furnace to its maximum temperature.

3.2 Interfacial stress, polysilicon or nitride films

Most solid-state integrated circuits require a polysilicon film for electrical connections and a silicon nitride film to improve reliability. Deposition of these films on a wafer generates stress in the wafer that can lead to bowing or warping.[6,7] This stress is interfacial stress induced during deposition and is not necessarily caused by the difference in the thermal expansion coefficient of the film and the substrate.[6,7] The magnitude of the stress depends on deposition conditions and it has been found to be quite different for polysilicon deposited at atmospheric pressure and for the same material at low pressure with similar substrate temperatures.[28] When the magnitude of the stress of the deposited film exceeds

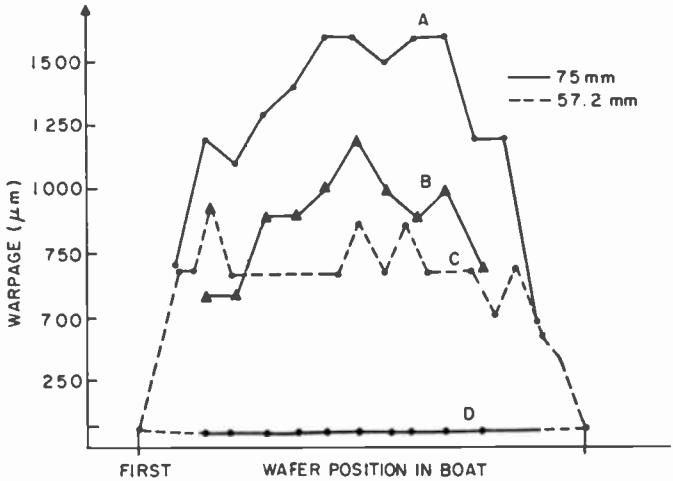


Fig. 4—Warpage measurements for 75 mm and 57.2 mm diameter wafers with different boats: (A) open boat, (B) boat with roof, (C) open boat, (D) boat with roof and metal reflector (after Ref. [24]).

the magnitude of the critical yield stress for the silicon substrate, dislocations will occur in the slip plane (111), as shown in Fig. 5. These dislocations can be generated to the depth of 20 μm below the surface.[7] In addition to the initial generation of dislocations, the interfacial stress from polysilicon or nitride films and stress from the difference in thermal expansion coefficients can further generate dislocations during subsequent high temperature material processing steps. Note that the yield stress of silicon is lower at high temperature (Fig. 2).

To prevent warpage caused by interfacial stress from polysilicon and nitride layers, the stress from the deposited films should be minimized. This can be done by optimizing deposition conditions and minimizing the thickness of the deposited films. Also, stress applied by the film deposited on one side of the wafer can be balanced by depositing a film of the right thickness on the other side of the wafer.[29] The practicality

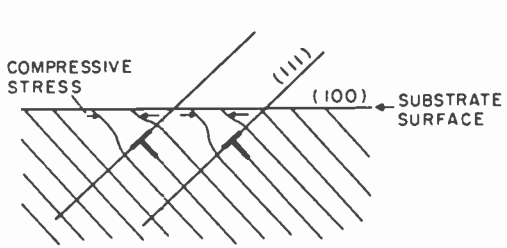


Fig. 5—Schematic illustration of the dislocation configuration on inclined {111} planes.

and effectiveness of this approach has to be estimated for each device process.

3.3 Impurity diffusion

As is well known, diffusion of dopants into silicon can result in bowing or warping of the wafer. Stress is generated due to a change in the lattice constant caused by doping. A heavy phosphorus diffusion (above 7×10^{20} cm^{-3}), which is often used for gettering, is of particular interest because the degree and direction of bow depends on the cooling conditions; the wafer bow for slow cooling is in an opposite direction from that observed for fast cooling.[5] This unusual phenomena is explained as follows. In silicon heavily doped with phosphorus occupying substitutional sites, the crystal lattice constant is smaller than that of the undoped part of the silicon wafer and a contractile stress is introduced by the diffused phosphorus layer (Fig. 6a). When the contractile stress exceeds a critical value, dislocations are generated in slip planes to relieve stress. By the end of the phosphorus diffusion process, the crystal lattice becomes deformed, as shown in Fig. 6b, and the wafer will bow. With rapid quenching to room temperature, the phosphorus which at low temperature is supersaturated does not have enough time to precipitate and remains frozen in the lattice in substitutional positions. The wafer is bowed as shown in Fig. 6b. When the wafer is cooled down slowly after diffusion the supersaturated phosphorus starts to precipitate. The silicon

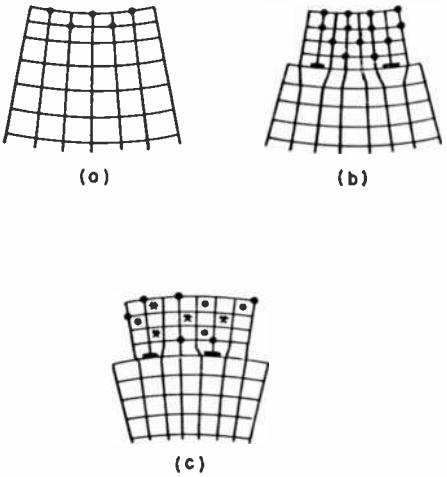


Fig. 6—Mechanism of anomalous bending of silicon wafers with heavy phosphorus diffusion. Solid dots represent P atom and asterisks a precipitate formed by P atoms removed from substitutional sites.

lattice constant is larger with the phosphorus atoms at interstitial sites or aggregated into precipitates to form a second phase. Now the diffused layer develops stress in the opposite direction than when precipitates are absent. If the tensile stress induced by the movement of the substitutional phosphorus atoms to interstitial positions exceeds the contractile stress present in the layer the wafer will bend in the opposite direction, as shown in Fig. 6c.

The obvious approach to avoid warpage during diffusion is to balance forces generated by the diffused regions. This can be achieved by the simultaneous diffusion of dopants into both faces of the wafer. Because diffused areas on the opposite faces of the wafer are usually different (e.g., the back side does not have a defined masking oxide), a different type or concentration of dopants should be used on each face to balance stress.[30]

4. Material Factors

4.1 Role of oxygen

Oxygen is believed to be the most important impurity in terms of effect on the mechanical properties of silicon. Its effect on the plastic deformation of silicon has been thoroughly investigated, but the observations are often contradictory and its role is still not fully understood. Before discussing the effect of oxygen on the mechanical properties of silicon, it is necessary to briefly outline its behavior in silicon.

The two types of silicon crystals used by the electronic industry, which lead to significant differences in oxygen content, are grown by the Czochralski and float zone methods. In Czochralski grown silicon, oxygen is introduced in the growing crystals due to dissolution of the quartz crucible by melted silicon during crystal growth. Depending on growth arrangements, the concentration of oxygen in the grown crystal can be as high as 2×10^{18} atoms/cm^3, which corresponds to the solubility limit of oxygen at the growth temperature.[31–33] By adjusting the molten silicon flow pattern in the crucible, the oxygen concentration can be reduced to a few $\times 10^{17}$ atoms/cm^3.[34] With the float zone growth process, no crucible is required to support the molten silicon, and the oxygen concentration in float zone grown crystals usually is on the order of 10^{15} atoms/cm^3.[34]

In either float zone or Czochralski grown silicon, oxygen can be in different forms such as interstitials,[35] substitutionals,[36] complexes,[37] or precipitations.[10] The state in which oxygen is present in the crystal will depend on the total oxygen concentration and the thermal history of the silicon crystal.

Infrared absorption is commonly used to detect the form and concentration of oxygen in silicon wafers used for fabrication process. Room temperature measurements allow the detection of oxygen on the level of 10^{17} to 10^{18} atoms/cm^3. At liquid helium temperature this method can be used to detect oxygen as low as $\sim 2 \times 10^{15}$ atoms/cm^3.[38]

Kaiser[35,39] demonstrated that the intensity of 1106 cm^{-1} band at 300°K is directly related to concentration of oxygen, as shown in Fig. 7. Initially, it was postulated that this band was caused by a single vibration mode of oxygen in interstitial positions.[31] Low temperature measurements show that this band is usually a result of superimposition of two absorption bands, one at 1128 cm^{-1} and the other at 1135 cm^{-1}, which can be clearly separated at temperatures below liquid nitrogen, as shown in Fig. 8.[40] It has been suggested that these bands correspond to different vibration modes of silicon isotopes with oxygen,[41] but changes of the band intensities with heat treatments[37] indicate that they are related to the different states of oxygen. The 1128 cm^{-1} band has been attributed to interstitial oxygen in the silicon lattice and the 1135 cm^{-1} band to oxygen forming second phase particles.[37] Oxygen precipitates are also responsible for appearance of at least one band at 1225 cm^{-1}.[40]

To understand the relation between oxygen and the mechanical

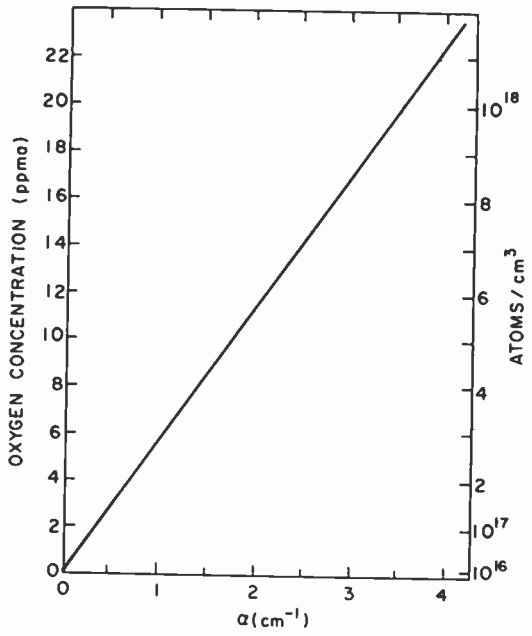


Fig. 7—Oxygen concentration versus absorption at 9 μm at 300°K (after Ref. [15]).

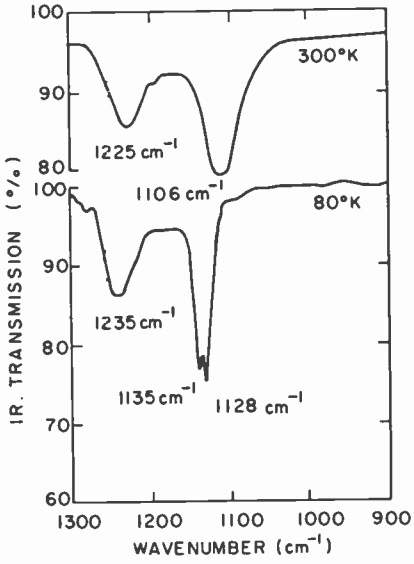


Fig. 8—IR absorption bands due to oxygen (after Ref. [40]).

properties of silicon we must consider the mechanism of dislocation generation. Plastic deformation takes place when dislocations are introduced. They are usually nucleated in the weak spots of a crystal, where the presence of an internal stress field decreases the energy required to nucleate dislocations. The internal stress field can be introduced by mechanical damage[42] or second phase precipitations. After nucleation under the influence of mechanical stress, dislocations can move from the place of nucleation to other areas.[42] During movement, dislocations can multiply by the Frank–Read mechanism, which is explained in detail in Appendix 1. The oxygen concentration in the crystal affects dislocation nucleation and movement. Presence of the oxygen precipitations induced by heat treatment increases the amount of nucleation sites and decreases the upper yield point, as shown in Fig. 9. The threshold energy required to start dislocation movement in the presence of oxygen is much higher than in an oxygen-free crystal.[21] At a low stress field, which is below the threshold value, the dislocations will remain locked to oxygen atoms, thereby exhibiting the so-called oxygen pinning effect. When the stress field exceeds the threshold value, which increases with increasing oxygen concentration, dislocations will move and will multiply by Frank–Read mechanism. The velocity with which they move is not influenced by the presence of oxygen, as shown by measurements of dislocation velocity in float zone and Czochralski crystals.[21] The influence of oxygen on dislocation propagation is significantly different[23] from

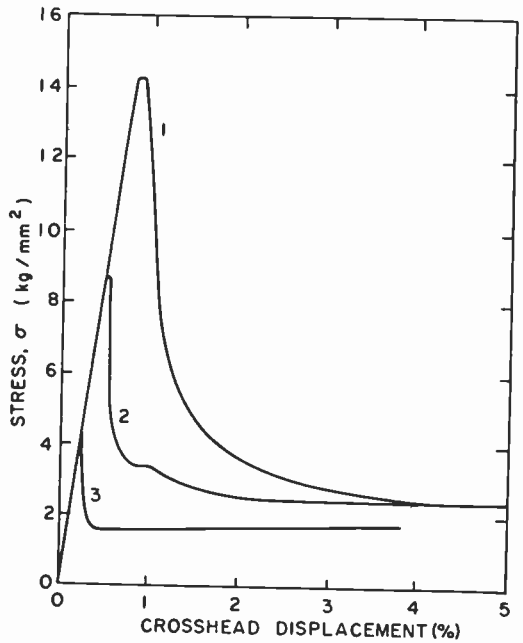


Fig. 9—Stress-strain curves with different concentrations of oxygen precipitates induced by various heat treatments; (1) dislocation free, no heat treatment, $\alpha(9\ \mu m) = 3.9\ cm^{-1}$, (2) 3 h at 1000°C, $\alpha = 3.8\ cm^{-1}$, (3) 16 h at 1000°C, $\alpha = 2.7\ cm^{-1}$ (after Ref. [11]).

that of acceptor or donor impurities. These affect the dislocation velocity, but do not exhibit a pinning effect.

The oxygen content of float zone silicon is two orders of magnitude lower than that of Czochralski silicon. Therefore, correspondingly lower values of the threshold strain field required for dislocation movement and multiplication in float zone silicon accounts for the better mechanical properties, i.e., the resistance to warpage of Czochralski wafers relative to float zone wafers.

The state of oxygen that gives the dislocation pinning effect is of technical importance but, to date, not well characterized. Work on the propagation of dislocations[15] established that oxygen clusters and not precipitates or interstitial oxygen are responsible for pinning. Oxygen clusters can be induced in Czochralski silicon by a low temperature (700°C) anneal and their appearance has been manifested by broadening of 1107 cm^{-1} absorption band.[15]

Although the chemistry of the oxygen clusters, which are known to improve the mechanical properties of silicon,[15] has not been clearly identified, an empirical procedure based on what is known of their

chemistry has been designed to improve material resistance to warpage.[43] Oxidation at 925°C for 100 minutes in dry oxygen followed by annealing for not less than five hours at 780°C–820°C (at 800°C for 15 hours is preferred) has been found to result in a significant increase in the threshold value of the initial shear stress during subsequent device processing at 920°C. Although this procedure is thought to introduce oxygen clusters into a beneficial state prior to a heat treatment, it is not obvious how long oxygen will remain in this state during subsequent heat treatments. It is still not known how oxygen can be maintained in the appropriate beneficial state during processing steps that have the largest potential for the introduction of plastic deformation (highest temperature and largest external forces). The oxygen state in silicon is a function of the wafer thermal history and total oxygen concentration. Therefore, it can be expected that the optimum range of oxygen concentration and a preheat treatment procedure can be established for a given device manufacture process which will result in the best mechanical properties of silicon wafers.

4.2 Other factors

(a) *Surface Preparation*

Gettering is usually applied to the back of a wafer to remove from active device regions metallic contamination that deteriorates device performance. One of the oldest methods of gettering was to deliberately introduce mechanical damage on the back side of a wafer.[50] This can, however, deteriorate the mechanical properties of the wafer. Fig. 10 shows the effect of mechanical damage that was left after polishing of the wafer surface.[12] When mechanical damage is present on the wafer surface, internal stresses are introduced and act as nucleation sides for dislocations.[44] Therefore, the value of the upper yield point on the stress-strain curve is significantly reduced compared to polished wafers and wafers with backside mechanical damage are more susceptible to a plastic deformation. Consequently, the benefits that gettering by mechanical damage provide during processing have to be very carefully evaluated against the deterioration of wafer mechanical properties.

(b) *Dislocations*

It has to be pointed out that the presence of dislocations in the wafer prior to processing also influences its mechanical properties. Fig. 11 shows the effect that dislocations have on the stress-strain curve.[12] The dislocations initially present in the crystal act as nucleation sites for new dislocations (Frank-Read mechanism) and, therefore, the resistance to warpage of a crystal with dislocations is lower than that of a crystal free of dislocations.

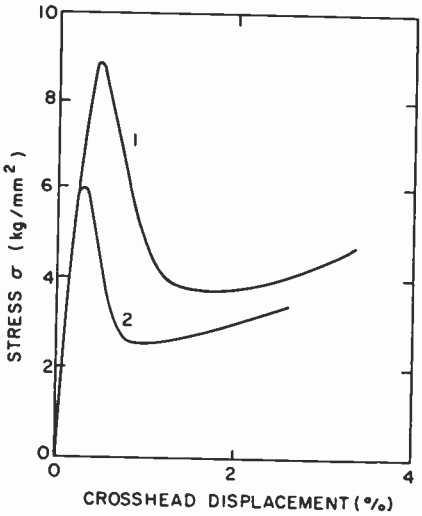


Fig. 10—The effect of the mechanical damage on the surface on stress-strain curves: (1) damage free, (2) with mechanical damage (after Ref. [12]).

(c) *Impurity density*

In addition, the mechanical properties of silicon wafers also depend on impurity density (measured by resistivity), although available data dealing with this subject are extremely limited. Fig. 12 presents the hardness measured at room temperature as a function of resistivity in n- and p-type silicon.[22] A relationship between hardness H and yield point Y is given by[45] $H = CY$, where C is the constraint factor (approximately equal to three). The room temperature mechanical properties improve in p-type wafers and deteriorate in n-type wafers with increasing concentrations of impurities. Contradictory data reported on microhardness of n-type Si as a function of resistivity[46] can be related to the difference in the starting material used in such experiments, e.g., state and concentration of oxygen.

The explanation of the dependence of hardness on resistivity (Fig. 12) is not straight-forward. It can be postulated that for n-type Si the observed decrease of hardness with increasing impurity concentration is related to an increase in the number of nucleation sites caused by impurities precipitation or by an increase of the dislocation velocity observed with decreasing resistivity (Fig. 13). But the sharp change in hardness (Fig. 12), which takes place for a dopant concentration of about 10^{16} cm^{-3}, is difficult to explain. Significant impurity precipation and an increase in dislocation velocity do not occur at this concentration. In addition, the data for p-type Si indicate that other mechanisms not

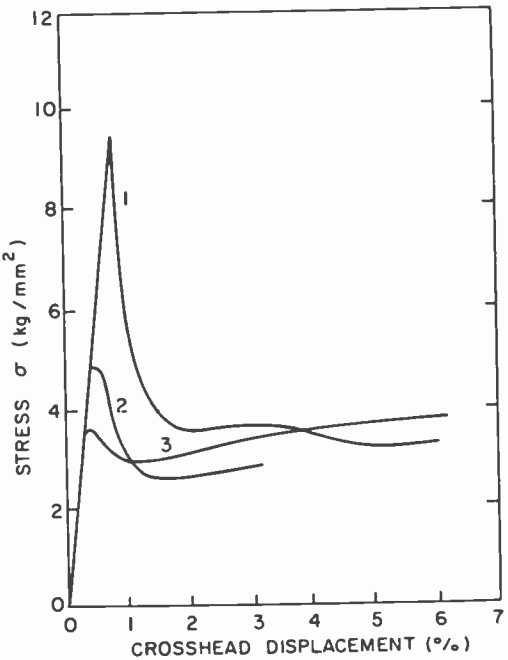


Fig. 11—Stress-strain curve with different dislocation density: (1) dislocation free, (2) $5 \times 10^4 cm/^2$, (3) $10^6/cm^2$ (after Ref. [12]).

understood so far are responsible for the observed dependence of mechanical properties on resistivity. The lack of understanding of the mechanisms that control the relationship between resistivity and mechanical properties makes it impossible to predict the mechanical behavior of wafers, with different resistivity, during device processing. It

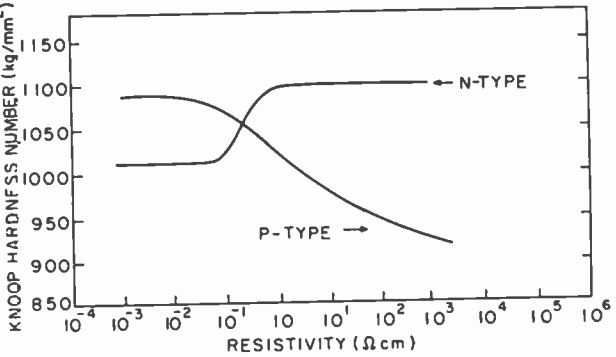


Fig. 12—Microhardness of p- and n-type silicon versus the resistivity (after Ref. [22]).

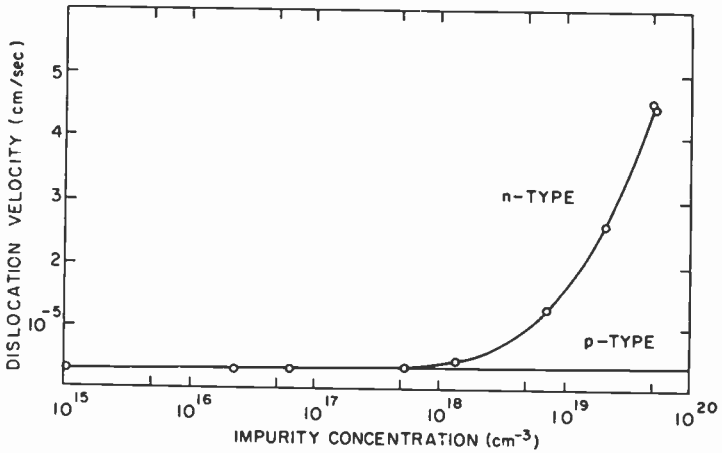


Fig. 13—Dislocation velocity versus impurity concentration for n and p type silicon for T = 600°C, τ = 3 kg/mm² (after Ref. [23]).

can be speculated that if the mechanical properties are a function of the Fermi level position and not chemical properties of impurities,[23] no difference in the mechanical properties between n- and p-type high resistivity material should exist. A thermal, band-to-band carrier generation in silicon will establish a high temperature carrier concentration on the order of $10^{18}cm^{-3}$ at 1200°C, which will override the differences in the room temperature resistivity in these wafers. If the room temperature data from Fig. 12 can be extrapolated to high temperatures, it can be expected that p-type low resistivity wafers will have higher resistance to warpage than their n-type counterpart.

5. Summary

Although it causes persistent yield losses in the device fabrication processes, warping or bowing of silicon wafers has received relatively little attention so far. The effect that dislocations, surface roughness, or oxygen have on the mechanical properties has been reviewed, although our understanding of these important problems is still pretty limited. The effect of these properties can be minimized by optimization of the process. The increasing diameter of silicon wafers and decreasing device dimensions make the yield of the processed circuits more and more susceptible to warpage and, therefore, the issue of material resistance to warpage is of primary importance.

Published data reviewed here show that there is an oxygen state in silicon that improves the wafer resistance to warpage. On the other hand,

oxygen in other forms, e.g., precipitates, will adversely affect mechanical properties. To date, the beneficial form of oxygen is not fully identified and our knowledge of how to introduce oxygen into this state is based mostly on empirical methods. Without understanding the fundamental nature of this state, and its behavior during subsequent heat treatments, the behavior of wafers during high temperature steps is difficult to predict. Therefore, it is difficult to choose an optimum oxygen range in silicon and optimum preheat treatment to increase material resistance to warpage during the most important processing steps. In addition to the role of oxygen, the role of other material parameters, such as carbon, point defects, or contaminants, has to be investigated in much more detail. Carbon can influence the mechanical properties by providing additional nucleation sides for dislocations at carbon precipitates. Point defects or contamination can affect dislocation movement by changing the electric state of dislocations after decoration[42,47] and, therefore, change the interaction of dislocations with a stress field.[23] All these factors have to be carefully investigated before a comprehensive model of the warpage of silicon wafers can be developed.

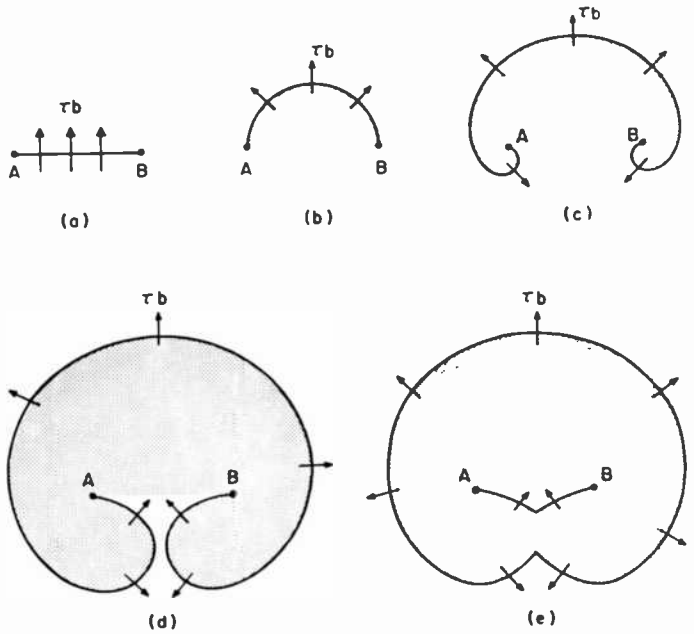


Fig. 14—Representation of the dislocation movement in the Frank-Read source.

Appendix 1

When annealed crystals are deformed, a rapid multiplication of dislocations and a progressive increase of dislocation density takes place. The mechanism believed to be responsible for dislocation multiplication is the Frank-Read mechanism[48,49] and its nature is explained in Fig. 14, which shows an example of a line dislocation situated in a slip plane.

The plane of Fig. 14 is the slip plane of a section of dislocation AB. The dislocation leaves the plane of the figure at the fixed points A and B which provide a barrier for the dislocation movement in the slip plane. An applied stress produces a normal force τ_b on the dislocation and makes the dislocation bulge in the slip plane (the A and B points are immobile), as show in Fig. 14b. If the stress is raised above the critical value, then the dislocation becomes statistically unstable (Fig. 14c) and starts to expand indefinitely. The unstable loop begins to double back on itself (Fig. 14d) and two parts of the dislocation are about to join up behind the original position of AB. Finally they join (Fig. 14e), and the sections of the dislocation that meet annihilate one another and leave

(a) a close dislocation loop which continues to expand under the applied stress
(b) a section of dislocation AB, which strengthens out and is ready to go through the same process all over again.

The Frank-Read model described above accounts for multiplication of all types of dislocations in a stress field.

Acknowledgments

The authors would like to thank especially P. J. Zanzucchi for valuable discussions and also G. W. Cullen and L. A. Goodman for a critical review of this paper.

Bibliography

1. Specifications for Perkin-Elmer Model 500 Projection Mask Aligner.
2. O. Paz, E. Hearn, and E. Fayo, "$POCl_3$ and Boron Gettering of LSI Silicon Devices," *J. Electrochem. Soc.*, **126**, p. 1754 (1979).
3. R. E. Gegenwarth and F. P. Laming, "Effect of Plastic Deformation of Silicon Wafers on Overlay," *SPIE*, **100**, p. 66 (1977).
4. K. Morizane and P. S. Gleim, "Thermal Stress and Plastic Deformation of Thin Silicon Slices," *J. Appl. Phys.*, **40**, p. 4104 (1969).
5. Y. Takano, N. Namba, and M. Maki, "Analysis of Anomalous

Bending of Silicon Wafers with Heavy Phosphorus Diffusion," *Semiconductor Silicon 1977*, Electrochem. Soc., Princeton, N.J., p. 695.

6. T. Suzuki, A. Mimura, and T. Ogawa, "The Deformation of Polycrystalline Silicon Deposited on Oxide-Covered Single Crystal Silicon Substrates," *J. Electrochem. Soc.*, **124**, p. 1776 (1977).
7. M. Tamura and H. Sunami, "Generation of Dislocations Induced by Chemical Vapor Deposited Si_3N_4 Films on Silicon," *Japan. J. Appl. Phys.*, **11**, p. 1097 (1972).
8. M. Suezawa, K. Sumino, and I. Yonenaga, "Dislocation Dynamics and the Plastic Deformation of Silicon Crystals. II," *Phys. Stat. Soli.* (*a*), **51**, p. 217 (1979).
9. I. Yonenaga and K. Sumino, "Dislocation Dynamics and the Plastic Deformation of Silicon Crystals. I," *Phys. Stat. Soli.* (*a*), **50**, p. 685 (1978).
10. J. R. Patel and A. R. Chaudhuri, "Macroscopic Plastic Properties of Dislocation-Free Germanium and Other Semiconductor Crystals: I. Yield Behavior," *J. Appl. Phys.*, **34**, p. 2788 (1963).
11. J. R. Patel, "Impurity Clustering Effects on Dislocation Generation in Silicon," *Dis. Faraday Soc.*, No. 38, p. 201 (1964).
12. J. R. Patel and A. R. Chaudhuri, *J. Appl. Phys.*, **34**, 2788 (1963).
13. S. Lederhandler and J. R. Patel, "Behavior of Oxygen in Plastically Deformed Silicon," *Phys. Rev.*, **108**, p. 239 (1957).
14. J. R. Patel and B. W. Batterman, "Impurity Clustering Effects on the Anomalous Transmission of X-Rays in Silicon," *J. Appl. Phys.*, **34**, p. 2716 (1963).
15. J. R. Patel, "Oxygen in Silicon," *Semiconductor Silicon 1977*, Electrochem. Soc., Princeton, N.J., p. 521.
16. B. Leroy and C. Plougonven, "Warpage of Silicon Wafers," *J. Electrochemical Soc.*, **127**, p. 961 (1980).
17. S. M. Hu, "Dislocation Pinning Effects of Oxygen Atoms in Silicon," *Appl. Phys. Lett.*, **31**, p. 53 (1977).
18. W. D. Sylwestrowicz, *Phil. Mag.*, **7**, p. 1825 (1962).
19. G. L. Pearson, W. T. Read, and W. L. Reldmann, "Deformation and Fracture of Small Silicon Crystals," *Acta Metal.*, **5**, p. 181 (1957).
20. S. M. Hu and W. J. Patrick, "Effect of Oxygen on Dislocation Movement in Silicon," *J. Appl. Phys.*, **46**, p. 1869 (1975).
21. K. Sumino, H. Harada, and I. Yonenaga, "Origin of the Difference in the Mechanical Strengths of Czochralski-Grown Silicon and Float-Zone-Grown Silicon," *Japan. J. Appl. Phys.*, **19**, p. L49 (1980).
22. A. H. Tong and M. R. Poponiak, IBM Technical Report, 22-507 (1967).

23. J. R. Patel, L. R. Testardi, and P. E. Freeland, "Electronic Effects on Dislocation Velocities in Heavily Doped Silicon," *Phys. Rev.*, **B13**, p. 3548 (1976).
24. E. W. Hearn, E. H. Tekaat, and G. H. Schwuttke, "The Closed Boat: A New Approach for Semiconductor Batch Processing," *Microelectr. Reliab.*, **15**, p. 51 (1976).
25. D. Hull, *Introduction to Dislocations*, Pergamon Press (1976).
26. S. M. Hu, "Temperature Distribution and Stresses in Circular Wafers in a Row During Radiative Cooling," *J. Appl. Phys.*, **40**, p. 4413 (1969).
27. L. D. Dyer, H. R. Huff, and W. W. Boyd, "Plastic Deformation in Central Regions of Epitaxial Silicon Slices," *J. Appl. Phys.*, **42**, p. 5680 (1971).
28. A. M. Goodman, private communication.
29. A. K. Gaind, E. W. Hearn, and G. H. Schwuttke, *IBM Technical Discussion Bull.*, **21** (3), p. 1001 (1978).
30. D. C. Lepiane, U.S. Patent 3,445,320 (1969).
31. H. J. Hrostowski and R. H. Kaiser, "Infrared Absorption of Oxygen in Silicon," *Phys. Rev.*, **107**, p. 966 (1957).
32. H. J. Hrostowski and R. H. Kaiser, "Infrared Spectra of Group III Acceptors in Silicon," *J. Phys. Chem. Solids*, **4**, p. 148 (1958).
33. F. A. Trumbore, "Solid Solubilities of Impurity Elements in Germanium and Silicon," *The Bell System Tech. J.*, **39**, p. 205 (1960).
34. Wacker specifications for commercially available crystals.
35. W. Kaiser, P. H. Keck, and C. F. Lange, "Infrared Absorption and Oxygen Content in Silicon and Germanium," *Phys. Rev.*, **101**, p. 1264 (1956).
36. D. Helmreich and E. Sirtl, "Oxygen in Silicon: A Modern View," *Semiconductor Silicon 1977*, Electrochem. Soc., Princeton, N.J., p. 626.
37. E. M. Ryzhkova, I. T. Trapeznikova, V. E. Chelnokov, and A. A. Yabovenko, "Optical Properties of Oxygen in Silicon," *Sov. Phys. Semicon.*, **11**, p. 628 (1977).
38. B. Pajot, *Analyst*, **9** (7), p. 293 (1977), and L. Jastrzebski and P. Zanzucchi, submitted to Fourth International Symp. on Silicon Science and Technology, May 10–15, 1981.
39. W. Kaiser and P. H. Keck, "Oxygen Content of Silicon Single Crystals," *J. Appl. Phys.*, **28**, p. 882 (1957).
40. W. J. Patrick, "The Precipitation of Oxygen in Silicon and its Effect on Surface Perfection," *Silicon Device Processing*, Charles P. Marsden, Editor, p. 442 (1970).
41. R. C. Newman, "Infrared Absorption due to Localized Modes of

Vibration of Impurity Complexes in Ionic and Semiconductor Crystals," *Advances in Physics*, 18, p. 545, (1969).
42. L. Jastrzebski, P. A. Levine, W. A. Fisher, A. D. Cope. E. D. Savoye, and W. N. Henry, presented at the 57th Electrochem. Soc. Meeting, St. Louis, May 11–16, 1980, Abstract No. 152. (submitted to the *J. Electrochem. Soc.*).
43. A. Guillot, B. Leroy, and C. Plougonven, *IBM Tech. Dist. Bull.*, 22, p. 3 (1979).
44. S. M. Hu, "Defects in Silicon Substrates," *J. Vacuum Science Tech.*, 14 (1), p. 17 (1977).
45. *Hardness*, J. R. Westbrook, Editor, American Soc. for Metals, Metals Park, Ohio, (1973).
46. N. A. Kalnev and I. N. Magden, *Electron. Technology*, 9 (1), p. 23 (1976).
47. P. M. Petroff, D. V. Lang, J. L. Strudel, and R. A. Logan, *Scanning Electron. Microscopy*, Vol. 1, p. 325, SEM Inc. AMF O'Hare (1978).
48. W. T. Read, *Dislocations in Crystals*, McGraw-Hill Book Co., Inc. (1953).
49. F. C. Frank and W. T. Read, Jr., "Multiplication Processes for Slow Moving Dislocations," *Phys. Rev.*, 79, p. 722 (1950).
50. E. J. Mets, "Poisening and Gettering Effects in Silicon Junctions," *J. Electrochem. Soc.*, 112, p. 420 (1965).

Patents Issued to RCA Inventors—Third Quarter 1980

July

A. A. Ahmed Transistor Saturation Control (4,213,068)
L. J. Bazin, G. R. Peterson, and D. M. Schneider Synchronizing Apparatus for Remote Television Apparatus (4,214,261)
W. Bohringer Regulated Deflection Circuit (4,214,189)
D. Botez Single Filament Semiconductor Laser (4,215,319)
J. K. Clemens Recording/Playback Apparatus Using a Single Pilot Tone for Active Tracking and Control of a VideoDisc Pickup Device (4,213,148)
J. W. Daniel, Jr. Phaselock Receiver with Phaselock Detector (4,213,096)
R. G. Ferrie Signal Selecting System (4,214,213)
R. E. Flory and M. R. Nelson Memory Addressing System for Automatic Setup TV Camera System (4,215,368)
M. T. Gale Projector for Reading Out Color-Encoded Focused Image Holograms Employing an Optimum Encoding Scheme (4,213,673)
A. Goldman Aqueous Photoresist of Casein and N-Methylol Acrylamide (4,211,563)
B. Halon and D. M. Hoffman Method for Making Adherent Pinhole Free Aluminum Films on Pyroelectric and/or Piezoelectric Substrates (4,214,018)
J. R. Harford Synchronization and Gain Control Circuit (4,212,032)
J. R. Harford AGC Keying Signal Circuit (4,213,151)
F. Z. Hawrylo Apparatus for the Deposition of a Material from a Liquid Phase (4,214,550)
L. F. Hopen Method for Spot-Knocking the Electron-Gun Mount Assembly of a CRT (4,214,798)
R. S. Hopkins, Jr. Circuit for Displaying Characters on Limited Bandwidth, Raster Scanned Display (4,212,008)
H. P. Kleinknecht Optical Testing of a Semiconductor (4,211,488)
H. P. Kleinknecht and W. A. Bosenberg Photomask Alignment System (4,211,489)
W. F. Kosonocky and D. J. Sauer CCD Gate Electrode Structures and Systems Employing the Same (4,211,936)
J. M. Neilson Gate Turn-Off Triac with Dual Low Conductivity Regions Contacting Central Gate Region (4,214,255)
E. J. Nossen Voltage Measuring Circuit (4,213,088)
J. I. Nubani and F. S. Sawicki Wet Carbon-Dioxide Treatment of Partially-Completed CRT (4,213,663)
M. Packer Sealing Composition (4,214,067)
M. Packer and O. D. Black Water Soluble Adhesive Coating for Mounting Components to Printed Wiring Boards (4,215,025)
M. H. Polhemus and A. L. Alger VideoDisc Player (256,117)
C. G. Rauchfuss, Jr., R. Herman, and H. R. Barton, Jr. Web Position Controller for Web Transport Systems (4,212,422)
W. Rosnowski Diffusion Apparatus (4,211,182)
W. Rosnowski Method of Fabricating Semiconductor Devices (4,213,807)
O. H. Schade, Jr. Integrated Circuitry Including Low-Leakage Capacitance (4,211,941)
V. J. Stakun and W. H. White Testing the Divergence of a Beam From a Laser (4,212,540)
T. O. Stanley Cadence Scanned Flat Image Display Device (4,215,293)
L. A. Torrington VideoDisc Player Having Adjustable End-of-Play Switch (4,211,421)
J. C. Turnbull and B. F. Yoder Method for Making an Indirectly-Heated Cathode Assembly (4,210,988)
R. Williams and Y. Arie Tellurium Schottky Barrier Contact for Amorphous Silicon Solar Cells (4,213,798)
R. Williams Liquid Junction Schottky Barrier Solar Cell (4,215,185)
F. P. Wipff Cadence Suppression System (4,214,126)

August

A. A. Ahmed Voltage-to-Current Converter Apparatus (4,216,435)
J. A. Allen Carriage Translating Apparatus for VideoDisc Player (4,216,969)
A. R. Balaban and S. A. Steckler Sample-Hold Phase Detector (4,216,396)
W. H. Barkow and R. W. Shisler Convergence Adjustment Arrangement Using Magnetic Tabs With Differential Motion (4,218,667)

T. E. Bart and R. A. Mills TV Graphics and Mixing Control (4,218,698)
A. E. Bell Optical Anti-Reflective Information Record (4,216,501)
A. Bloom and R. A. Bartolini Ablatable Medium for Optical Recording (4,218,689)
A. Bloom, W. J. Burke, and D. L. Ross Optical Recording Medium (4,219,826)
D. W. Breithaupt Push-Pull Amplifier (4,218,638)
R. E. Cardinal Method for Resistance Welding of an Electro-Optic Device Housing (4,217,683)
D. E. Carlson Compensated Amorphous Silicon Solar Cell (4,217,148)
J. E. Carnes, R. H. Dawson, and W. F. Kosonocky Comb Filter Employing a Charge Transfer Device with Plural Mutually Proportioned Signal Charge Inputs (4,217,605)
C. A. Catanese and L. S. Cosentino Isolation Busbar for a Flat Panel Display Device (4,217,519)
R. A. Dischert and R. L. Libbey Black Level Control System for a Color Television Camera (4,218,699)
F. C. Easter Astable Multivibrator Circuit (4,216,441)
E. C. Fox Peak Detector Circuit (4,216,502)
J. E. Gillberg and N. Kucharewski Drive Circuit for Controlling Current Output Rise and Fall Times (4,216,393)
J. R. Harford Modular Printed Circuit Board (4,216,523)
J. R. Harford Keyed AGC Circuit (4,218,708)
A. C. Ipri Fabrication of an Integrated Injection Logic Device Incorporating an MOS/Bipolar Current Injector (4,217,688)
H. C. Johnson Scanning Radar (4,219,814)
A. J. Leidich Leakage Current Compensation Circuit (4,216,394)
A. M. Miller and H. J. Richardson Intermittent Prime Film Mover Using Multi-Toothed Wheel Drive (4,215,921)
H. T. Nelson Guide Wing for a Furnace Paddle (4,218,214)
J. J. Piascinski Method for Assembling a Base to an Electron Tube (4,217,014)
R. M. Rast Hangup Corrector Useful in Locked Loop Tuning System (4,218,657)
O. H. Schade, Jr. Relaxation Oscillator Not Restricted by FET Threshold (4,219,787)
H. M. Schwartz Magnetic Transducer Head Core (4,217,613)
F. W. Spong Optical Record Playback Apparatus Employing Light Frequency at Which Alternate Regions of Record Track Exhibit Anti-Reflection Condition (4,219,848)
D. L. Staebler and P. J. Zanzucchi Method of Inducing Differential Etch Rates in Glow Discharge Produced Amorphous Silicon (4,217,393)
R. G. Stewart Level Shift Circuit (4,216,390)
F. E. Vaccaro Flat Panel Display Device with Beam Collector (4,216,407)
C. C. Wang, L. Ekstrom, and H. Wielicki VideoDisc Lubricants (4,216,970)
C. F. Wheatley, Jr. Control Circuit for Multivibrator (4,216,442)
R. E. Wilson Narrow Pulse Eliminator (4,216,388)

September

A. A. Ahmed Quasi-Linear Amplifier with Feedback-Controlled Idling Currents (4,220,930)
A. A. Ahmed Non-Inverting Buffer Circuits (4,221,979)
A. A. Ahmed Relaxation Oscillator Including an SCR and Having Switch Means for Interrupting Current Flow Therethrough (4,223,280)
A. A. Ahmed SCR Relaxation Oscillator with Current Amplifier in its Gate Circuit (4,223,281)
A. E. Bell and R. A. Bartolini Sensitivity Information Record (4,222,071)
S. L. Bendell Focusing System for Color TV Camera (4,223,253)
J. P. Bingham and W. A. Lagoni Image Detail Improvement in a Vertical Detail Enhancement System (4,223,340)
D. W. Breithaupt Horizontal Synchronizing System (4,222,074)
K. K. Chan Two into Three Port Phase Shifting Power Divider (4,223,283)
C. F. Coleman VideoDisc Player Having Modular Construction (4,220,338)
C. F. Coleman and N. L. Farley Caddy-Actuated Declutching Mechanism for VideoDisc Player (4,220,339)
J. Craft AFT Circuit (4,220,974)
G. A. Cutsorgeorge Stabilization of Monolithic Integrated Circuit Output Levels (4,224,536)
W. Denhollander Raster Distortion Correction Circuit (4,220,898)
P. Foldes Rotatable Polarization Duplexer (4,222,017)
R. L. Giordano Temperature Compensated Switching Circuit (4,220,873)
R. L. Giordano Temperature Compensated Switching Circuit (4,220,877)
N. F. Cubitose, R. E. Novak, H. R. Ronan, Jr., and M. R. Schuler Meter and Dispensing System for Abrasive Materials (4,222,502)
P. E. Haferl Deflection Circuit with Retrace Control (4,223,251)

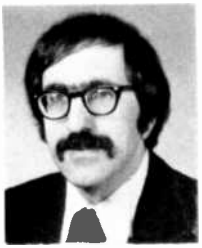
W. J. Hannan and H. J. Woll Patching Tape for Diffractive Subtractive Filter Viewgraphs (4,221,465)
A. C. Ipri Short Channel MOS Devices and the Method of Manufacturing Same (4,225,875)
L. A. Kaplan Overcurrent Protection Circuit for Power Transistor (4,225,897)
H. Kawamoto and E. L. Allen, Jr. Resistivity Measurement System (4,220,915)
W. A. Lagoni and J. S. Fuhrer Video Image Vertical Detail Restoration and Enhancement (4,223,339)
P. A. Levine Reduction of Sparkle Noise and Mottling in CCD Imagers (4,223,234)
J. Lusch Tuning Position Display System (4,222,120)
R. U. Martinelli Semiconductor Device Having Integrated Diode (4,225,874)
R. S. Mezrich Apparatus for Dividing an Arbitrary Waveform into Equal-Amplitude Steps (4,222,008)
E. J. Nossen and C. H. Haber Signal Correlation Means (4,224,679)
Y. Okuno FM Defect Compensation Apparatus (4,221,930)
J. I. Pankove Method and Structure for Passivating a Semiconductor Device (4,224,084)
D. H. Pritchard, W. E. Sepp, and W. A. Lagoni Frequency Multiplier for Use with CCD Comb Filter (4,224,638)
L. C. Ruth Cathode Ray Tube Display Error Measurement Apparatus and Method (4,220,919)
J. R. Sandercock Fabry-Perot Interferometer (4,225,236)
O. H. Schade, Jr. Precision Current Source (4,225,816)
L. N. Schiff Compandor for Group of FDM Signals (4,221,934)
M. A. Spak Etchant Solution Containing HF-HNO_3-H_2SO_4-H_2O_2 for Etching Al-Ti-Cu or Ni Contact Metallurgy on Silicon Substrates (4,220,706)
L. A. Torrington and L. D. Huff Constant Drag Carriage Translating Mechanism for VideoDisc Player (4,220,340)
L. A. Torrington and L. D. Huff Mechanism for Aiding Carriage Return in VideoDisc Player (4,225,141)
J. A. VanRaalte Phosphor Screen for Modular Flat Panel Display Device (4,220,892)
C. T. Wu Encoding Analog Signals into Digital Signals Using a Triangular Reference (4,220,925)

AUTHORS

R. Casanova Alig received the B.A. degree from Wabash College in 1963, and the M.S. and Ph.D. degrees in physics from Purdue University in 1965 and 1967. He did theoretical studies of acoustic transport phenomena in metals for his dissertation. In 1967 Dr. Alig joined RCA Laboratories as a member of the Physics and Chemistry of Solids Group in the Physical Electronics Laboratory. He has recently done theoretical studies of the optical spectra of impurities and color centers in crystals. During the academic year 1970–71 he was a visiting professor at the Instituto de Fisica e Quimica, São Carlos, S.P., Brasil; in addition, since 1967 he has given courses at Mercer County Community College; LaSalle College, and Princeton University. He is a member of Phi Beta Kappa, Sigma Xi, and the American Physical Society.

Wolfram A. Bosenberg attended the University of Stuttgart, Germany. He received his diploma and Ph.D. degrees in physics in 1951 and 1955, respectively. From 1952 to 1957, he worked at the semiconductor plant of IT&T in Germany on germanium diffusion, crystal growth and silicon power rectifiers. In 1957, he joined RCA and has worked at both the Somerville and Princeton locations. Dr. Bösenberg's thirty years of semiconductor experience has spanned a broad range of projects: silicon RF power transistors, gallium arsenide devides, germanium and gallium arsenide tunnel diodes, junction field effect transistors, MOS field effect transistor arrays for core drivers, Schottky barrier circuits, and MOS large scale integration. He has contributed extensively to semiconductor technology and instrumentation.

Sheng T. Hsu received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1958, the M.S.E.E. degree from National Chiao-Tung University, Hsienchu, Taiwan, in 1960 and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 1966. From 1966 to 1970, he was with Fairchild Semiconductor Research and Development Laboratory, Palo Alto, California, as a Member of the Technical Staff. From 1970 to 1972, he was an Assistant Professor of the Department of Electrical Engineering, University of Manitoba, Winnipeg, Man., Canada. He is now a staff member of RCA Laboratories, Princeton, N.J., working on semiconductor devices and integrated circuit technologies.



Alfred C. Ipri received a B.S. degree from Drexel Institute of Technology in 1965 and an M.S. in Electrical Engineering and a Ph.D. from the University of Pennsylvania in 1967 and 1972 respectively. Since joining the technical staff at RCA Laboratories, Princeton, N.J., in 1967, he has been involved in materials utilization, process development, device characterization, circuit design, and systems applications in the field of integrated circuits and this work has led to numerous presentations, publications and patents.

Dr. Ipri is a member of Eta Kappa Nu, the Institute of Electrical and Electronic Engineers, and the National Society of Professional Engineers.

Lubomir Jastrzebski completed his undergraduate studies at Warsaw University in 1970 and he received a M.S. degree in Physics from Warsaw University in 1971 and a Ph.D. degree from the Institute of Physics, Polish Academy of Science in 1974. Between 1975 and 1978 he worked in the Materials Science and Engineering Department at M.I.T. In 1978 he joined RCA Laboratories. His main area of interest is analysis of the relationships between defects affecting performance of solid state devices and defects introduced to material during processing or crystal growth process. His work has resulted in about 45 original papers published in various scientific journals and presented at international conferences. Dr. Jastrzebski is a member of the American Physical Society and American Association of Crystal Growth.

Arnold R. Moore received the B.S. in Chemistry from the Polytechnic Institute of New York in 1942, with thesis work in Chemical Kinetics. Thereafter, he worked on gas discharge and photoelectric surface problems with RCA manufacturing divisions in Harrison, NJ, and Lancaster, Pa. In 1945 he returned to graduate school at Cornell University, where in 1949 he received the Ph.D. degree in Experimental Physics, with a minor in Theoretical Physics and Mathematics, under a National Research Council Fellowship. His work at Cornell consisted chiefly of experimental research into the motion of electrons in alkali and silver halides, along with the growth of the required single crystals. Since starting work at RCA Laboratories in 1949, he has been active in the fields of transistor physics, optical absorption in semiconductors and semiconductor alloys, magnetic susceptibility, acoustoelectric effect in cadmium sulfide and gallium arsenide, and most recently photovoltaic solar converters. In 1970–71, as a visiting professor at Brown University, he taught a graduate course in physics of solid-state devices and participated in research on photovoltaic cells.

Since his return to RCA Laboratories in 1971, Dr. Moore has been working on insulators and photo devices. He has received three RCA Achievement Awards and is a Fellow of the American Physical Society and a member of Sigma Xi.

Daniel Thebault obtained the diploma-engineer degree in physics from the National Institut of Applied Sciences (INSA) of Toulouse, France and received the D.E.A. degree from the University of Toulouse in 1980. He worked seven months this year as an exchange student in the Materials Synthesis Research Group at RCA Laboratories.

Index Volume 41, 1980

Index to Authors, Volume 41, 1980