Ionization Rates of Holes and Electrons in Silicon

C. A. LEE, R. A. LOGAN, R. L. BATDORF, J. J. KLEIMACK, AND W. WIEGMANN
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 18 July 1963; revised manuscript received 15 January 1964)

The ionization rates of charge carriers in silicon have been measured and fit to the recent theoretical calculations of Baraff; in contrast, none of the existing published data could be fit to these theoretical curves. The study has been made using microplasma-free junctions of demonstrably high, uniform local multiplication. A new and considerably simplified approach to the problem of extracting the ionization rates from the multiplication data has been used. By employing much more precise control of the electron and hole currents used to initiate the multiplication process, the hole ionization rate at electric fields less than 300 kV/cm is found to be more than an order of magnitude smaller than any previously published measurements. Hole and electron ionization rates have been measured in the same junction and consequently in the identical scattering environment. The threshold energy is determined to be $E_g \leq E_i \leq 1.5E_g$, and the mean free path for scattering of high-energy electrons is 50 Å$\leq\lambda_e\leq$70 Å and for energetic holes 30 Å$\leq\lambda_h\leq$45 Å. Measurement of ionization rates at various temperatures substantiates the assumption that the energy-loss mechanism is the emission of optical phonons. In addition, significant differences of the electrical breakdown characteristics of microplasma-free junctions are discussed as well as their preparation.

INTRODUCTION AND SUMMARY

UPON the completion of a new theoretical study of the field dependence of the ionization rates of mobile charge carriers in semiconductors,[1] it was disconcerting to discover that the existing experimentally determined rates in silicon[2–5] could not be fitted to the theory with any reasonable values of the adjustable parameters. The three adjustable parameters are the threshold energy E_i, the total mean free path for scattering λ, and a constant energy loss associated with scattering. This last parameter is simply the energy of the transverse optical phonon E_R, when emission of optical phonons is the dominant scattering mechanism for particle energies greater than E_R.

Although Baraff's analysis represents a considerable advance over the previous investigations of Wolff[6] and Shockley[7] in that no assumptions are made concerning the shape of the distribution function and in the appropriate limits the field dependences are the same as the Wolff and Shockley theories, it is well to remember that there still remain important simplifying assumptions. These simplifications assume that scattering processes are only dependent on energy and are isotropic and further, they neglect the multivalley anisotropic band structure of most semiconductors by assuming simple parabolic bands. It was not, of course, expected that these simplifications would cause severe departures from a reasonable description of the physical situation, but only comparison with experiment could decide this question in a rather simple way.

The first encouraging experimental comparison with Baraff's analysis was published by Logan *et al.*[8] for the ionization rates of electrons in gallium arsenide *p-n* junctions. The field dependence was found to be characteristic of the result obtained by Wolff and the fit to the new curves determined a threshold energy of 1.7 eV and a total mean free path for scattering of 15 Å. This fitting of the GaAs data stood in contrast to the intractability of the published silicon data. A subsequent review of the literature on silicon revealed several areas for improvement both experimentally and theoretically. These considerations have resulted in the study reported here.

The new ionization-rate data, which can be fitted satisfactorily to the curves calculated by Baraff, showed significant differences upon comparison with earlier data.[2–5] For electrons, the ionization rate versus the reciprocal field exhibits a steeper slope than the old data and is somewhat greater in magnitude; it is this steeper slope which is crucial in obtaining a fit to the theoretical curves. Further, the new data for holes, in addition to exhibiting a steeper slope of the ionization rate versus reciprocal field, is approximately a factor of 50 less than the corresponding electron ionization rate measured in the same junction. The largest previously reported difference[3] in the ionization rates was a factor of 5 and in all of the previous measurements the comparison of rates is obscured by the possibility of different scattering environments in the complimentary junctions used.

It is estimated that the success of these measurements lies in three areas of improvement: Firstly, considerable technological improvements have been achieved which permit the study of *p-n* junctions free of microplasma defects[9]; secondly, a new approach to extracting ionization rates from the multiplication data has been used

[1] G. A. Baraff, Phys. Rev. **128**, 2507 (1962).
[2] A. G. Chynoweth, J. Appl. Phys. **31**, 1161 (1960).
[3] A. G. Chynoweth, Phys. Rev. **109**, 1537 (1958).
[4] S. L. Miller, Phys. Rev. **105**, 1246 (1957).
[5] J. L. Moll and R. Van Overstraeten, Solid-State Electron. **6**, 147 (1963).
[6] P. A. Wolff, Phys. Rev. **95**, 1415 (1954).
[7] W. Shockley, Solid-State Electron. **2**, 35 (1961); or Czech. J. Phys. **11**, 81 (1961).
[8] R. A. Logan, A. G. Chynoweth, and G. B. Cohen, Phys. Rev. **128**, 2518 (1962).
[9] There has been a recent attempt to study ionization rates in silicon junctions free of microplasmas. This paper (Ref. 2) did not properly account for the capacitance of a guard-ring structure and thus erred in the calculation of the junction fields.

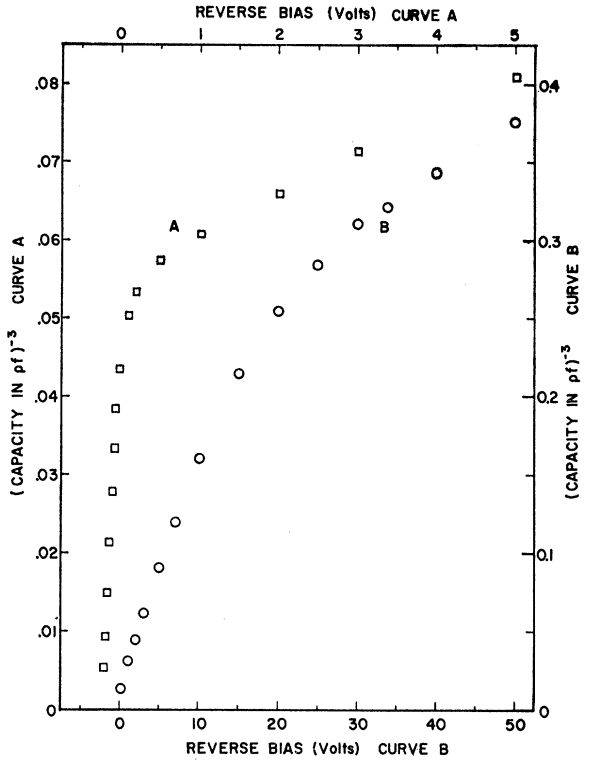


FIG. 1. Variation of capacitance versus reverse bias for two $p\nu n^+$ diodes with different resistivities of the epitaxially grown ν region.

that permits one to obtain an exact solution of the integral equations [see Eqs. (7a) and (7b)] when the multiplication as initiated by electrons and holes is measured in the same junction; and finally, much more precise control of the purity of hole and electron currents which initiate the multiplication process has been employed.

The paper is divided roughly into four parts. The first section describes the impurity distributions of the junctions and the corresponding electric-field distributions. Standard fabrication techniques are indicated and significant departures from these methods are discussed. The second part describes the multiplication properties of microplasma-free junctions with particular attention devoted to the local multiplication as a function of bias. Significant differences in the electrical behavior at breakdown are also discussed. The third section describes the calculation of the ionization rates from the multiplication; this calculation uses a new and considerably simplified approach to the problem. A comparison and discussion of the ionization rates with respect to the earlier published data is also given. The fourth and last section discusses the fitting of the ionization rate data to the theoretical curves calculated by Baraff and the interpretation of the parameters resulting from that fit. In addition, it is shown that the measured temperature variation of the ionization rate leads to a temperature variation of the mean free path for scattering that substantiates the assumption of a constant energy-loss mechanism associated with the emission of optical phonons by energetic carriers.

UNIFORM FIELD STRUCTURES

The field in a $p\nu n^+$ junction is uniform if (1) the middle ν region is sufficiently lightly doped that a negligibly small field change occurs across it and (2) the end regions are sufficiently heavily doped and sharply stepped in concentration. To achieve this device, a high resistivity n-type epitaxial layer, approximately 7 μ thick, was grown on an n^+ substrate of resistivity, 0.005 Ω-cm ($N_D = 1.4\times10^{19}$ cm^{-3}). The surface was then oxidized using a masking technique that left windows, approximately 10 mils in diameter through which boron was diffused to a depth of approximately 3 μ, to form the p-type contact. Electrical contact to the p-layer was made by the evaporation and alloying of a concentric ring of aluminum ~7 mils in diameter and 2 mils in width to which a gold wire was bonded. Contact to the n^+ base was acheived by gold bonding the wafer to a conventional header.

It is noted that the oxide protection to the junction edges at the surface was not disturbed by the subsequent formation of the contacts and that Ohmic series resistance effects were negligible throughout the range of currents used in this study ($I \leq 10^{-3}$ A).

The bias dependence of the capacity was measured to within 10 volts of breakdown. Curves A and B of Fig. 1 are plots of C^{-3} versus V for two typical junctions with epitaxial layer resistivities of 30 and 1.5 Ω-cm, respectively. It is seen that after an initial linear rise a change in slope develops which is caused by the space charge barrier moving from the ν region into the n^+ material. The capacitance at this bias (which may be more easily located on a log C versus log V plot) corresponds to a space-charge width equal to the ν region, the latter being equal to that directly measured with conventional bevelling and staining techniques on an adjacent portion of the structure. This result was verified on many units of varying epitaxial layer resistivities for which applied biases at punch-through V_p varied from almost zero for curve A to values in excess of 100 V. Moreover, the width of the space charge barrier at V_p as estimated from the epitaxial layer resistivity, was in good agreement with that obtained from other techniques.

The built-in voltage V_i for junctions with large V_p (Fig. 1, curve B) was ~0.6 V; for junction A, however, $V_i = 0.3$ V. The concentration gradient inferred from the linear slope of the forward bias portion of curve A is 2×10^{18} cm^{-4}. For a linearly graded junction, this gradient would imply a higher value of V_i. The smaller measured value, however, is a consequence of the asymmetry of the doping profile. Qualitatively one may understand the difference by noting [with the aid of Eq. (2)] that the less rapid termination of the electric field on the lightly doped side is the dominating contri-

bution leading to a slightly smaller junction width and potential at zero bias.

After each epitaxial layer growth and each complete diffusion cycle, the layer depths were determined and the sheet resistivity measured. The donors in the n^+ substrate diffused about 1 μ into the epitaxial layer during the cumulative heat treatments used in processing the structure. The ν layer was therefore about 3 μ thick. The p-type boron layer was approximately Gaussian in shape with a surface concentration of $\sim 5\times 10^{18}$ cm^{-3}.

Typical semilogarithmic plots of the reverse current versus applied bias are shown in Fig. 2. Curve C, one of the diodes used in the multiplication studies to be described in a later section, shows a sharp breakdown at 98 V. The prebreakdown current is estimated to be due to imperfection in the oxide protection of the junction surface and to recombination generation current in the space-charge region. The reverse characteristic shown in curve D of Fig. 2 is that of a junction similar to that of curve C; the prebreakdown current, however, is about an order of magnitude smaller. This variation is typical of that encountered in these units and is ascribed to surface effects. Semilogarithmic plots of the forward current versus bias are linear in the current range 10^{-7}–10^{-3} A, with a slope of $q/1.7\ kT$. The slope is indicative of carrier generation in the transition region. The reverse-bias breakdown voltage decreased with decreasing temperature as expected for avalanche multiplication.

At biases higher than V_p, the space-charge region penetrates very slowly into the p and n^+ regions. In fact, the width changes less than 5% over the bias range $V_B/2<V<V_B$ where the multiplication is measured. The $p\nu n^+$ structure with $V_p\leq 1$ V, therefore, closely approximates a uniform field junction; the field in this case is given by $\mathcal{E}-\mathcal{E}_p=(V_a+V_i-V_p)/W_a$, where W_a is the space charge width corresponding to V_a, and $\mathcal{E}_p$ is the maximum field corresponding to V_p.

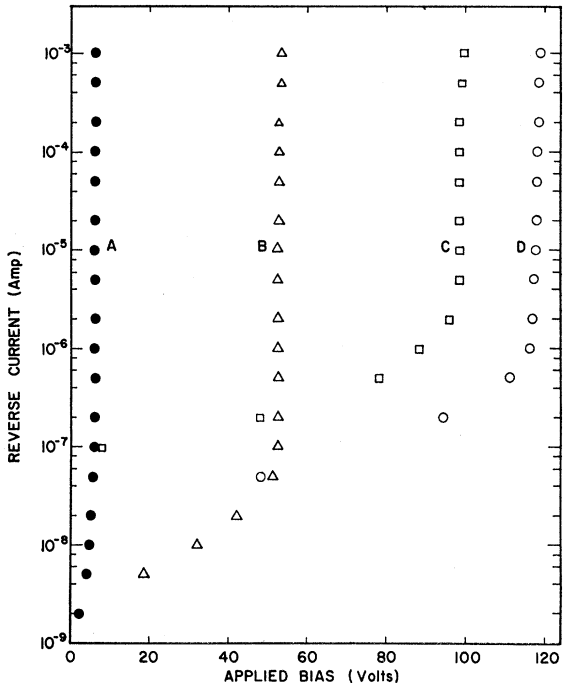


FIG. 2. Semilogarithmic plots of reverse current versus applied bias for typical units studied; curves A and B are graded junctions and curves C and D are $p\nu n^+$ uniform field structures.

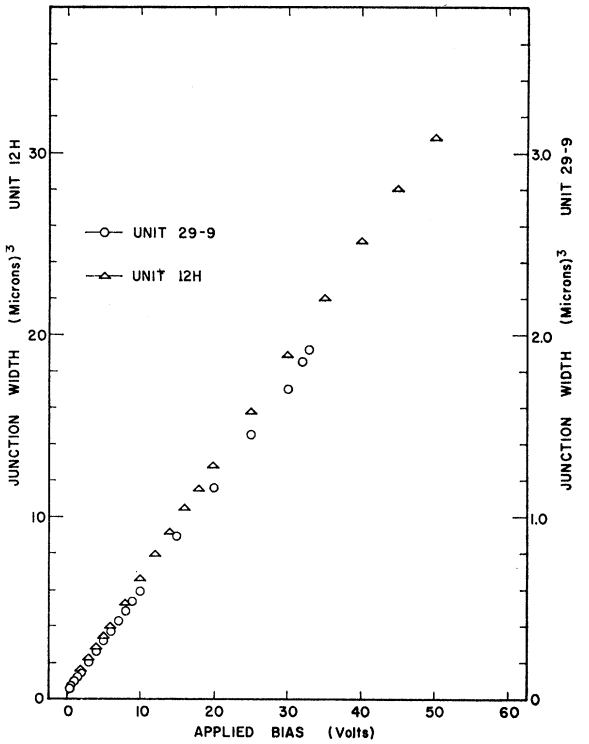


FIG. 3. Junction transition width versus reverse bias for two of the exponentially graded diodes used to determine ionization rates.

GRADED JUNCTIONS

The graded junctions studied were prepared by vacuum diffusion in an apparatus whose characteristics have been described elsewhere.[10] The characteristic feature of this system is the exponential distribution of the diffused inpurities generated by the particular boundary conditions prevailing when equilibrium at constant vapor pressure of the impurities is reached. This distribution is simply the result of the silicon surface evaporating at a rate equal to the quotient of the diffusion current of impurities and the equilibrium concentration at the surface. The shape of the distribution, as will be shown, is an important consideration in computing the electric-field distribution of the junction and also in calculating the ionization rates.

The capacity measurements as a function of the reverse bias, which are used in the calculation of the elec-

[10] R. L. Batdorf and F. M. Smits, J. Appl. Phys. **30**, 259 (1959).

tric fields of the junctions studied, are plotted in Fig. 3. Unit 12H is a gallium-diffused layer of 5.5 μ thickness on n-type material, and unit 29-9 is a phosphorous-diffused layer of 3 μ thickness on p-type crystal. The material of unit 12H was a pulled crystal of moderate dislocation density ($\approx 10^4/\text{cm}^2$), while unit 29-9 was made from a crystal pulled using an after heater to straighten out the liquid-solid interface so that the resulting dislocation density was $\leq 10/\text{cm}^2$.

As we have already mentioned, the impurity distribution in these diffused junctions is exponential so that in the neighborhood of the junction the net chemical impurity has the following distribution

$$N(x) = \pm N_b(1 - e^{-x/L}), \tag{1}$$

where the positive sign is for donors and the negative sign for acceptors, N_b is the background doping of the crystal, and the origin is at the position of exact compensation. The symbol L represents the characteristic decrement of the diffusant that is established at a particular diffusion temperature. Then employing elementary considerations of fixed space charge and sharp boundaries of the exhaustion region one arrives at simple expressions for the potential difference and the maximum electric field as a function of the width w of the space-charge region

$$\Delta\psi = \frac{qN_bL^2}{\epsilon}\omega\left[\frac{\omega}{2}\coth\frac{\omega}{2} - 1\right], \tag{2}$$

and

$$\mathcal{E}_m = \frac{qN_bL}{\epsilon}\left[1 - \frac{\omega}{1-e^{-\omega}} + \ln\frac{\omega}{1-e^{-\omega}}\right], \tag{3}$$

where $\omega \equiv w/L$. Then from previous measurements of the distribution of impurities in these diffused layers formed at 1300°C we find that

$$L(\text{phosphorus}) \approx 0.7\ \mu,$$
$$L(\text{gallium}) \approx 1.4\ \mu.$$

The dimensionless width ω for units 12H and 29-9 does not exceed a value of 3. For this limiting width the corrections of the maximum field of such an exponentially graded junction from that of a linearly graded junction amount to only a few percent. The effect of these field corrections upon the calculation of the ionization rate will be considered later.

The problem of making microplasma-free junctions is a complex one and cannot be discussed in a completely satisfactory way; because in spite of the recipes in hand for preventing the formation of microplasmas and for reintroducing them in a controlled fashion, the details of their formation have not yet been understood. The techniques we have used, however, clearly indicate that microplasmas are associated with the precipitation of unknown impurities at defects in the silicon crystal. The evidence for this conclusion is that under conditions where the contamination during processing of the junctions is held to an extremely low level ($\lesssim 10^{13}\text{cm}^{-3}$) and where material with a very low degree of crystal i nperfection is used we observe the lowest incidence of microplasmas. Further, when a microplasma-free junction is heat treated under conditions known to enhance the precipitation of impurities, microplasmas may be introduced into the junction if such impurities are present during the heat treatment.

The vacuum diffusion system described is particularly suited to achieving a low level of contamination. All parts of the system to which the silicon is exposed during diffusion are bakeable *in situ* at temperatures in excess of 2000°C. Tantalum is the only material directly contacting the silicon; and, obviously within the context of these experiments, it does not introduce any deleterious contamination.

Although the vacuum diffusion system may possess some inherent advantages in the degree of cleanliness obtainable, the more conventional open-tube diffusion techniques are also capable of producing microplasma-free junctions. The first success in making such junctions was reported by Batdorf *et al.*[11] and they described a guard-ring structure made by open-tube diffusion and oxide-masking techniques. These techniques along with a considerable number of refinements and additions have been used to make the $p\nu n^+$ uniform-field structure described in the first section. This $p\nu n^+$ structure comprises the collector junction of the silicon "planar" transistor. Indeed, the techniques and results of the methods used in fabricating the $p\nu n^+$ structure are quite similar to those described by Goetzberger *et al.*[12] for the guard-ring structure they describe in their paper.

Here too, for these more conventional methods, the progressive and painstaking efforts to reduce contamination, initially undertaken in the interests of greater reliability and process control, have resulted in an environment where there is very little precipitation of impurities on crystalline defects. It should be noted, at this point, that since the $p\nu n^+$ structure was made in an epitaxial layer with a considerable degree of crystal imperfection, it is clear that the occurrence of a high dislocation density *per se* does not prevent the fabrication of microplasma-free junctions; Goetzberger *et al.*[12] cite a similar experience in this respect. Our experience with vacuum diffusion, however, would allow us to make the further observation that the density of microplasmas qualitatively correlates with the density of dislocations in materials processed in a given diffusion run.

ELECTRICAL BREAKDOWN AND MULTIPLICATION CHARACTERISTICS

The techniques used to observe the multiplication and electrical breakdown noise are similar in principle to those used by previous investigators but they differ in the degree of refinement of the apparatus. For ex-

[11] R. L. Batdorf, A. G. Chynoweth, G. C. Dacey, and P. W. Foy, J. Appl. Phys. **31**, 1153 (1960).

[12] A. Goetzberger, B. McDonald, R. H. Haitz, and R. M. Scarlett, J. Appl. Phys. **34**, 1591 (1963).

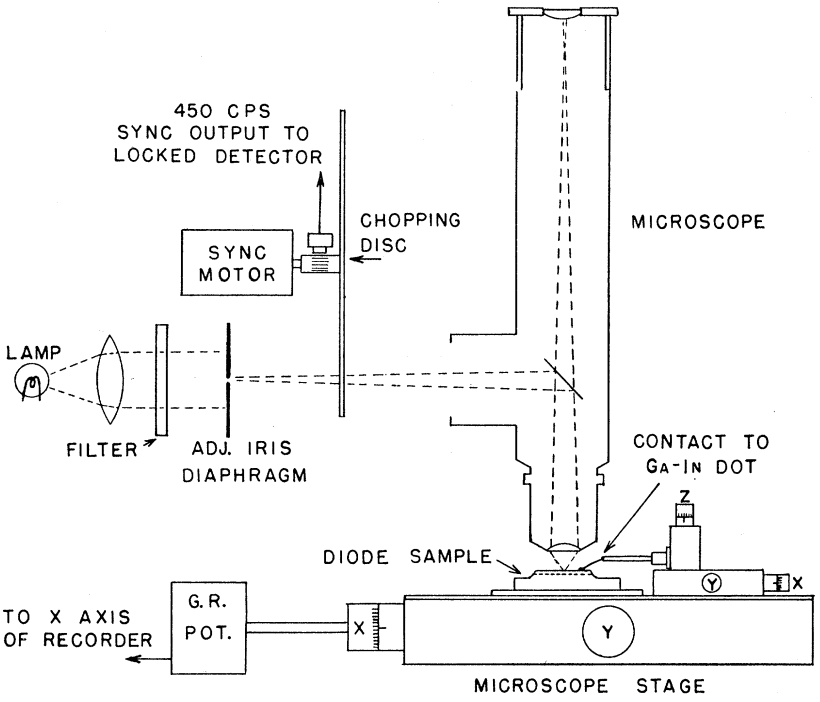


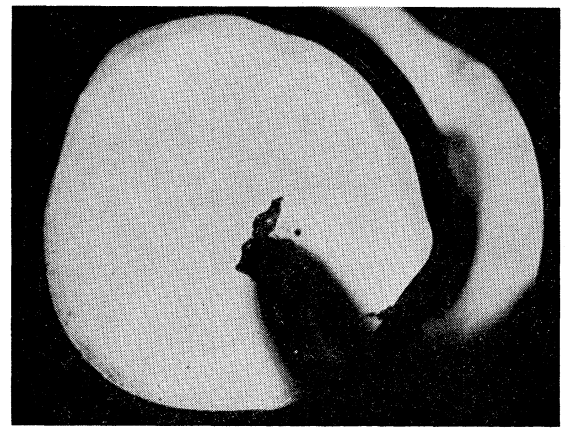
FIG. 4. Schematic of the experimental arrangement for measurement of the local multiplication.

ample, in conjunction with the usual chopped light source and synchronous detector we have found it essential to precisely control the illumination of the junction with regard to both the wavelength of the radiation and control of the scattered radiation. Towards this end we have used the objective of a Zeiss Standard GFL microscope to image an iris diaphragm on the surface of the diode (see Fig. 4). A micromanipulator was constructed so as to permit use of objectives with working distances as short as 0.5 mm. For fine control of movement in the x,y plane, the bed of a toolmaker's microscope was utilized in such a fashion that the junction could be moved under the light spot imaged by the microscope objective. This arrangement permitted simultaneous viewing and measurement of the multiplication of areas of the junction from 1 mm in diameter down to 20 μ. Lastly, provision was made for simultaneous observation of the electrical noise; the signal across a 100-ohm resistor in series with the diode was amplified by three cascaded Spencer-Kennedy distributed amplifiers whose output was fed into a Model-545 Tektronix oscilloscope.

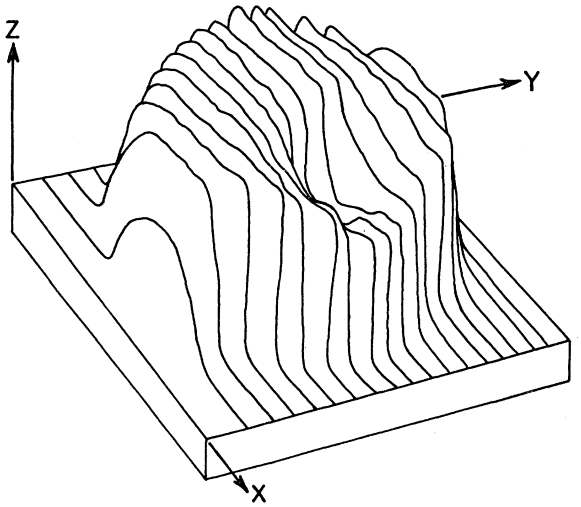
With the intent of obtaining ionization rates of holes and electrons in silicon from junction multiplication measurements, it is of primary importance to assess the uniformity of the local multiplication over the junction area. With the exception of the work of Kikuchi and Tachihawa[13] this problem has been neglected in favor of the more accessible observations concerning the uniformity (or nonuniformity) of the light emission and electrical noise at breakdown. The uniformity of multiplication that we have been able to acheive is illustrated in Fig. 5. Figure 5(a) is a photograph of an etched junction, approximately 16 mils in diameter, showing the periphery of the mesa and the phosphor-bronze point contacting the surface layer in the lower right-hand corner. In the accompanying photograph, Fig. 5(b), the x,y plane is coincident with the plane of the junction and the multiplied photocurrent is plotted on the vertical axis. The orientation is approximately the same as that in Fig. 5(a), and the depression in the photocurrent results from the shadow of the point contacting the layer. Each profile in Fig. 5(b) represents a recorder trace of the multiplied photocurrent from a 1-mil spot of light as it traverses the junction; the separation of the traces is also 1 mil.

The top of the cards represents a multiplication of 100 and is seen to be uniform to a few percent over the area of the junction. At lower multiplications the variations in the local multiplication decrease, and in the range where the ionization rates are calculated they are less than 1%. As one examines the multiplication behavior in the range above 100, however, a rapid increase in the local variations is observed. In fact, at a mean multiplication level of 10^3 the junction illustrated in Fig. 5 had variations of almost an order of magnitude between the areas of minimum and maximum multiplication. The areas of peak multiplication are in general, but not exclusively, associated with minute defects which can just be barely observed on the surface; and in the breakdown region these areas are associated with spots of light.

[13] M. Kikuchi and K. Tachihawa, J. Phys. Soc. Japan **14**, 1830 (1959); **15**, 835 (1960); M. Kikuchi, *ibid.* **15**, 1822 (1960).



(a)



(b)

FIG. 5. (a) Photograph of etched mesa of exponentially graded junction. (b) Local multiplied photocurrent of junction pictured in Fig. 5(a). The x, y plane is coincident with the plane of the junction. Each profile represents a recorder trace of a 1-mil light spot traversing the junction. The separation of the traces is also 1 mil. The mean multiplication is 100 and the depression arises from the shadowing by the point contacting the surface.

The accompanying electrical characteristics of the breakdown region provide evidence for the microplasma-free nature of these junctions. Further, they may be interpreted in a manner consistent with the observed multiplication characteristics. A typical current-voltage characteristic is shown in Fig. 6. The rise of current in the breakdown region is seen to be exponential without abrupt rises characteristic of microplasmas. Much more direct evidence, however, is obtained from direct observation of the pulses generated in the breakdown region. The set of photographs in Fig. 7 show oscillograph traces of these pulses as a function of junction voltage. It is characteristic of a microplasma that the pulse height saturates and the period of time it remains switched on increases progressively with the voltage until the microplasma remains switched on all the time. The pulses of Fig. 7, however, show that the pulse height increases with the bias until the decreasing junction impedance produces an overriding attenuation of the signal. The pulse width is determined by the associated circuitry and is not representative of the junction itself. We know from the measurements of the local multiplication that for very high values the peak multiplication is localized at only a few places in the junction. It is then a consistent interpretation to regard the observed pulses as due to a single minority carrier entering one of these localized regions when the multiplication of that region is of the order of 10^6 or 10^7. Were it not for the localized nature of these regions of high multiplication, the time resolution of our circuitry would probably be insufficient to resolve the pulses arising from the saturation current of the entire junction.

The multiplication data from which the ionization rates are determined are shown in Figs. 8, 9, and 10. The photocurrent of the uniform field structure ($p\nu n^+$) in Fig. 8 shows an initial sharp rise at low bias but this bias range is still in excess of that required for extending the space charge region through the ν layer. A parallel observation is found with other p-i-n structures in that the photocurrent increases until the electric fields in the intrinsic layer are sufficiently high to bring the drift velocity of carriers well into the region of saturation. The remaining increase of photocurrent in the middle region of bias may be attributed to several small changes in the collection efficiency; the multiplication in this

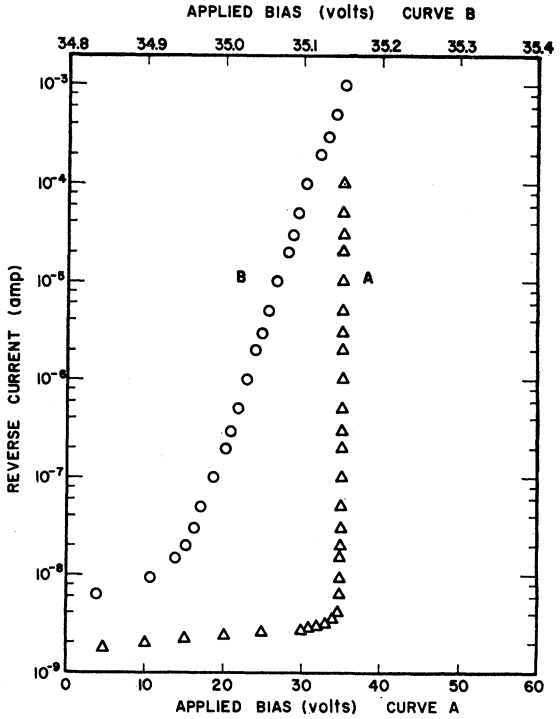


FIG. 6. Expanded plot of reverse current versus applied bias to illustrate the sharp exponential rise of current in breakdown.

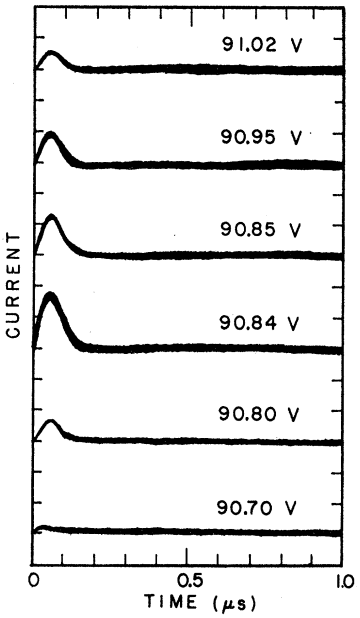


FIG. 7. Pulses observed in a microplasma-free junction for a bias range starting before breakdown and extending into the breakdown region. The vertical scale is approximately 50 μA/div.

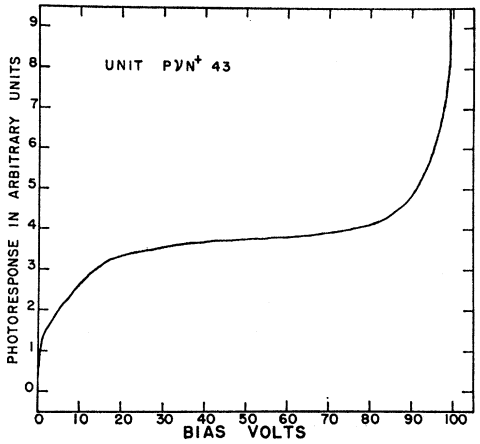


FIG. 8. Photoresponse of the uniform field structure versus bias.

case is determined by using a sloping base line and the final errors on the ionization rate reflect the error introduced by this procedure.

The photoresponse of unit 12H is shown in Fig. 9 for constant electrical gain and progressively decreasing light intensity. The illumination in this case was the green line of mercury (5641 Å); the light was also restricted to the surface layer which was gallium doped and 5.5 μ thick.

Figure 10 shows the multiplication data of unit 29-9 the surface layer of which was phosphorous doped and 3 μ thick. The curve taken with penetrating radiation (tungsten lamp plus silicon filter 1 mm thick) gives the multiplication characteristic of that initiated by electrons. The other curve taken with nonpenetrating radiation (4360 Å) is the multiplication characteristic as initiated by holes. Since the attenuation of the 4360 Å radiation in a distance of 3 μ is considerably in excess of 10^3 it is obvious that the control of scattered radiation which may strike the periphery of the junction is quite important.

We have stated that the curve taken with penetrating radiation in Fig. 10 was characteristic of electrons but it must be noted that the saturation photocurrent is contaminated with holes. A rough estimate of the ratio of the hole and electron currents is the ratio of the surface-layer thickness to the diffusion length of electrons in the p-type body. It is evident from a comparison

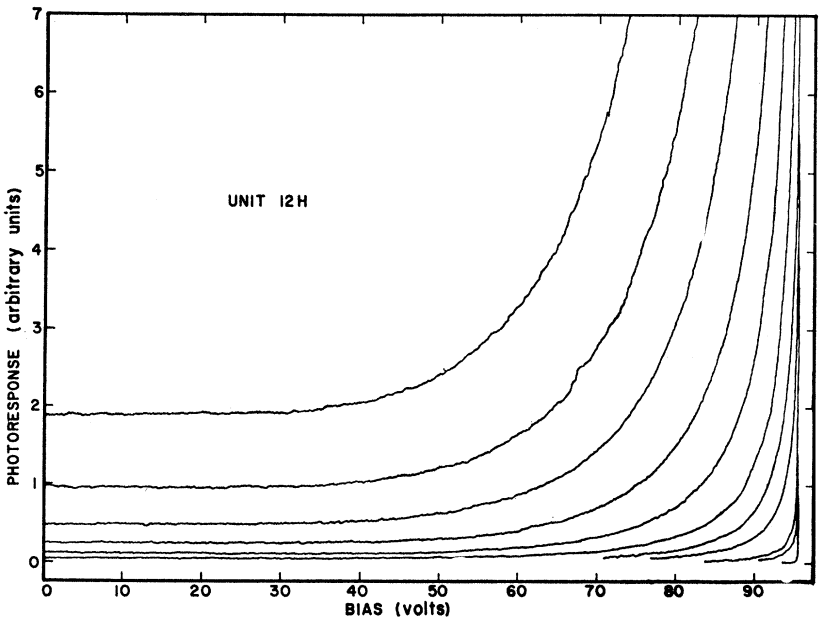


FIG. 9. Photoresponse of an exponentially graded junction versus bias for progressively decreasing light intensity.

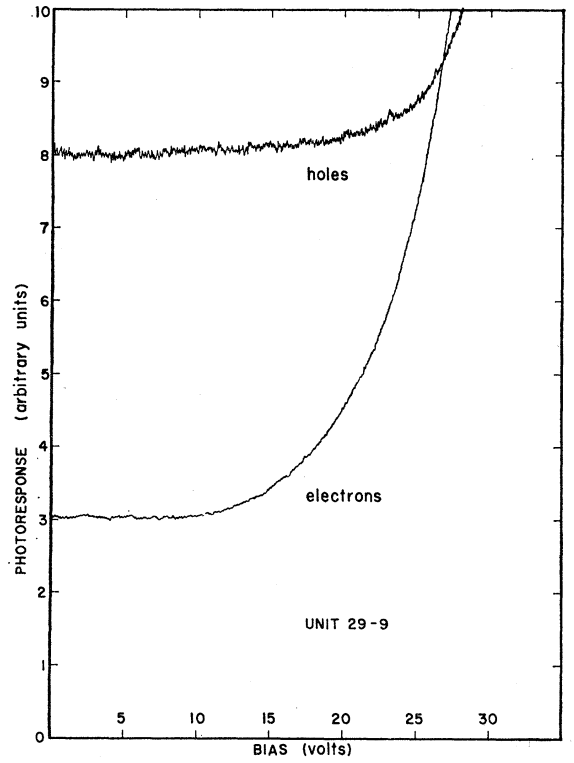


FIG. 10. Photoresponse of junction in which the ionization rates of holes and electrons were measured.

of the electron and hole curves of Fig. 10 that the ionization rate of holes is considerably smaller than that of electrons and in this case the hole contamination of the electron current leads essentially to a base-line correction. Estimates of the body lifetime from diode storage times indicate with the aid of Eq. (6) of the next section an error in the ionization rates of less than 15% and this is included in the final error estimate.

IONIZATION RATES OF ELECTRONS AND HOLES

The problem of calculating the ionization rates from the multiplication has always centered on the difficulties of solving the integral equations relating these quantities. The case of equal ionization rates is a tractable one and McKay[14] first gave solutions for junctions with linear and parabolic field distributions. The subsequent realization that the ionization rates were substantially different led to the generalization presented by Miller.[15] In the same paper Miller also presented a series solution for a linear field distribution that is characteristic of a step junction. There remained, however, the much more difficult cases of the parabolic field distribution and distributions intermediate between this and a linear field distribution. Many investigators[3–5,14,16] have introduced various simplifications, but in light of the heretofore unrealized magnitude of the difference in the ionization rates it can be shown that these simplifications introduce undesirably large errors.

In order to effect some simplification of the problem it is instructive to consider a slight generalization of the situation considered by Miller.[15] Instead of just considering that a pure electron or hole current initiates the multiplication, let us examine the consequence of using a mixture of holes and electrons. This situation is illustrated in Fig. 11. The accompanying differential equation takes the form[17]

$$-\partial J_n/\partial x=\partial J_p/\partial x=\alpha|J_n|+\beta|J_p|\,. \tag{4}$$

If we then define

$$k\equiv J_{ps}/J_s\,, \tag{5a}$$

$$M\equiv(J_p+J_n)/J_s\,, \tag{5b}$$

$$J_s\equiv(J_{ns}+J_{ps})\,, \tag{5c}$$

and use the boundary conditions that $J_n(0)=J_{ns}$, $J_p(w)=J_{ps}$, the solution of Eq. (4) is

$$1-\frac{1}{M}=\frac{(1-k)}{1-k+k\exp\left[-\int_0^w(\alpha-\beta)dx\right]}\times\int_0^w\alpha\exp\left[-\int_0^x(\alpha-\beta)dx'\right]dx+\frac{k\exp\left[-\int_0^w(\alpha-\beta)dx\right]}{1-k+k\exp\left[-\int_0^w(\alpha-\beta)dx\right]}\times\int_0^w\beta\exp\left[\int_x^w(\alpha-\beta)dx'\right]dx. \tag{6}$$

When a pure electron or hole current ($k=0,1$, respec-

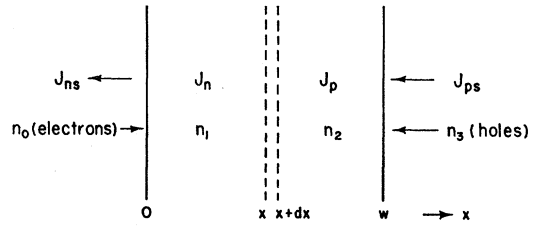


FIG. 11. Schematic of junction quantities used to derive the multiplication when the saturation current is a mixture of holes and electrons.

[14] K. G. McKay, Phys. Rev. **94**, 877 (1954).
[15] S. L. Miller, Phys. Rev. **99**, 1234 (1955).
[16] B. M. Wul and A. P. Shotov, *Solid-State Physics and Electronics Telecommunications* (Academic Press Inc., New York, 1960), Vol. I, p. 491.
[17] The derivations of Miller (Ref. 15) and of McKay (Ref. 14) neglected the different velocities of holes and electrons by writing the differential equation (4) in terms of densities. The slight inaccuracy is immaterial since the end result Eq. (6) is the same in either case.

tively) is used to initiate the multiplication, Eq. (6) yields the pair

$$1-\frac{1}{M_n}=\int_0^w \alpha \exp\left[-\int_0^x(\alpha-\beta)dx'\right]dx, \quad (7a)$$

$$1-\frac{1}{M_p}=\int_0^w \beta \exp\left[\int_x^w(\alpha-\beta)dx'\right]dx. \quad (7b)$$

Equation (7a) is, of course, identical to that given by Miller. To obtain the equation equivalent to (7b), however, Miller stated, quite correctly, that one need only interchange the α and β of Eq. (7a) to obtain

$$1-\frac{1}{M_p}=\int_0^w \beta \exp\left[\int_0^x(\alpha-\beta)dx'\right]dx. \quad (8)$$

The difference between Eqs. (7b) and (8) resides in the fact that, tacitly, the field has been turned around in Eq. (8).

Although this may seem to be a trivial point, it has proved nevertheless an elusive one, for when the forms of Eqs. (7a) and (7b) are considered, in contrast to that of Eqs. (7a) and (8), it becomes obvious upon multiplying (7b) by $\exp[-\int_0^w(\alpha-\beta)dx]$ that the former pair may be combined to give the result

$$\int_0^w(\alpha-\beta)dx=\ln(M_n/M_p). \quad (9)$$

This integral equation may be solved to yield the incomplete integral in the square brackets of Eqs. (7a) and (7b). When the kernels are known, the resulting mathematical simplification of the solution of Eqs. (7a) and (7b) for α and β strongly recommends making the measurement of M_n and M_p in the same junction.[18] Indeed, when the physical reasons for comparing the electron and hole ionization rates in the same scattering environment are appreciated, it is eminently worth the additional experimental effort required.

The graded junctions we have examined are close approximations to a parabolic field distribution and for this case Eqs. (7a) and (7b) take the form

$$1-\frac{1}{M_n}=2\exp\left[-\int_0^{\varepsilon_m}(\alpha-\beta)dy\right]\times\int_0^{\varepsilon_m}\alpha\cosh\left[\int_\varepsilon^{\varepsilon_m}(\alpha-\beta)dy'\right]dy, \quad (10a)$$

[18] Since the authors first reported on this work [C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, Bull. Am. Phys. Soc. **7**, 536 (1962)], N. R. Howard [J. Electron. Control **13**, 537 (1962)] has independently arrived at Eqs. (7a). (7b) and (9). Unfortunately the multiplication data used by Howard are shown by us in this paper to be deficient in the establishment of the proper boundary conditions; thus, it is difficult to assess the meaningfulness of Howard's results.

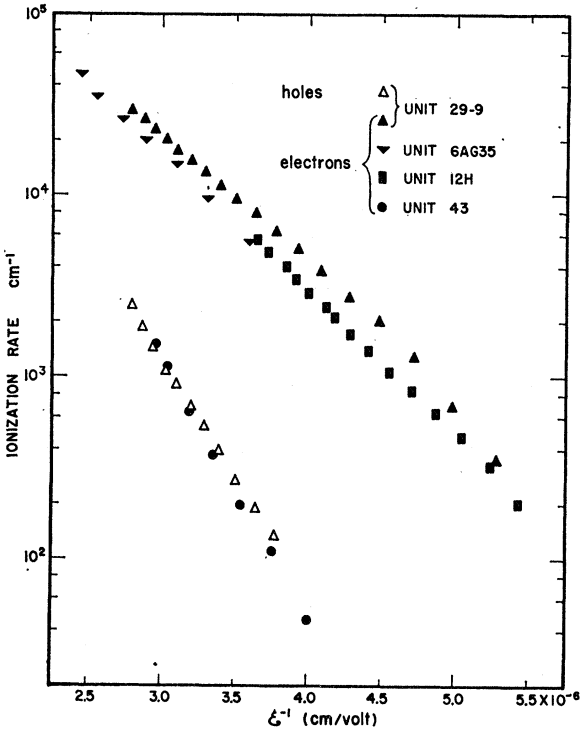


FIG. 12. Ionization rates of holes and electrons versus reciprocal field for the various units studied.

$$1-\frac{1}{M_p}=2\exp\left[\int_0^{\varepsilon_m}(\alpha-\beta)dy\right]\times\int_0^{\varepsilon_m}\beta\cosh\left[\int_\varepsilon^{\varepsilon_m}(\alpha-\beta)dy'\right]dy, \quad (10b)$$

where $dy\equiv d[-(w/2)(1-\mathcal{E}/\mathcal{E}_m)^{1/2}]$. Deviations from this parabolic field distribution that arise from the exponential distribution of impurities in our graded junctions and the consequent effect on the ionization rate are considered in Appendix A. Also, the numerical solutions of Eqs. (10a) and (10b) have been kept quite simple so that the accuracy is consistent with the experimental uncertainties. A description of the explicit procedure may be found in Appendix B.

It should be noted that the stratagem of restricting the calculation of the ionization rates to values of M very close to unity cannot be used with any reasonable accuracy to compute the smaller ionization rate if it is appreciably smaller. As is the case in silicon, the feedback effects of electrons [represented by the term $\int(\alpha-\beta)$] are appreciable throughout the range where multiplication that is initiated by a hole current may be measured.

The ionization rates of the junctions we have described are shown in Fig. 12. The breakdown voltages range from 6 to 98 V. For the most part, the electron ionization rates of the different units are in relative

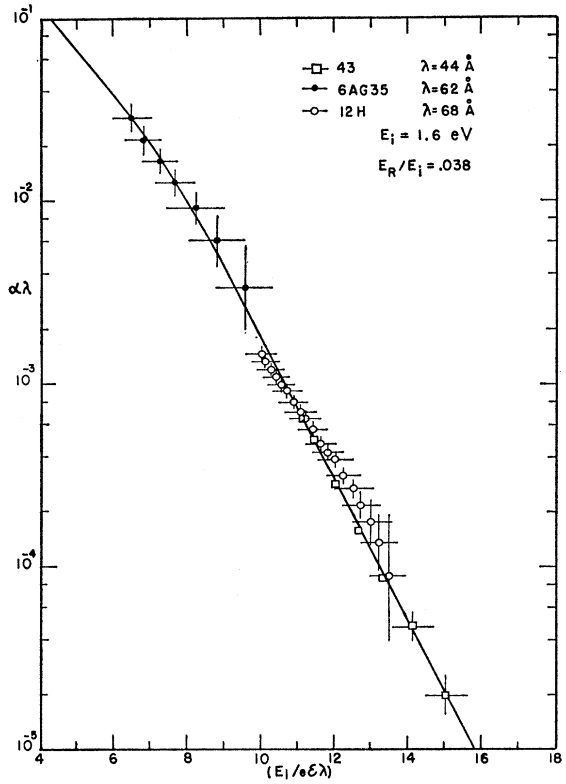


FIG. 13. Fit of the electron ionization rate to one of Baraff's curves for junctions with breakdown voltages ranging from 6 to 95 V.

agreement except for unit 43. This last unit is the uniform field structure ($p\nu n^+$) and the coincidence of the electron ionization rate of this unit with the hole data of unit 29-9 is fortuitous.

In comparing the electron ionization rates of units 12H and 29-9 with earlier data[3,5] we find the magnitude is some 50–60% greater than the older data. Further, the slope of the ionization rate versus reciprocal field is significantly greater than that of the older data. As we shall discuss in more detail in the next section, when the slope of the experimental data is less than a certain critical value, a fit to the theoretical curves is impossible for any reasonable values of the adjustable parameters. It is felt that the more satisfactory slope of the present data may stem from the microplasma-free nature of the junctions used in this study.

The most significant departure from previously published data is the hole ionization rate of unit 29-9. Previous work on silicon[3–5] has cited a factor ranging from 4 to 5 in comparing the ionization rate of electrons to that of holes. The hole data of unit 29-9 is smaller than the older data[3] by another order of magnitude. Although the comparison of rates of the previous data may suffer some uncertainty because holes and electrons were measured in complementary junctions, we feel, rather, the main difference of the present data lies in the experimental establishment of a purer hole current to initiate the multiplication process. The comparative smallness of the ionization rate of holes requires that the hole current used to initiate the multiplication be pure to a fraction of a percent. The conclusion is inescapable that previous experiments covering this range of electric field have not met this requirement.

Lastly, a qualitative explanation of the markedly different ionization rate of unit 43 may be found in the significantly different scattering environment the electrons encounter in this structure; the near intrinsic region was an epitaxially grown layer with a considerable degree of crystal imperfection. A measure of this imperfection is the degradation of the conductivity mobility reported in similar layers of germanium.[19] In addition, it has been found that the imperfections are pronounced in the region near the interface.[20]

The circumstance that unit 43 has such a markedly lower ionization rate is an exaggerated case of a problem that has plagued these measurements from the beginning. A unique ionization rate is attainable only in the limit of measurements of a perfect crystal. We cannot assert, of course, how close these measurements are to this limit, but we have chosen to present measurements of crystals ranging from those with moderate to those with very low dislocation densities; and in addition the processing has introduced a minimal degree of contamination. From this sort of comparison one may at least obtain a qualitative assessment of the effect of these physical variables on the ionization rate.

COMPARISON OF IONIZATION RATES TO THEORY

The issue of primary importance is the comparison of the observed field dependence of the ionization rates with that of the recent calculation by Baraff.[1] The central feature of this calculation is the derivation of an integral equation involving a space and energy-dependent collision density from the Boltzmann equation under the assumption that the scattering probability depends only on the energy. With a further assumption concerning the fraction of ionizing collisions as a function of energy one can obtain, by integrating the product of this fraction and the collision density over all energies, the density of ionizing collisions as a function of distance. The usual ionization rate per unit path length is then defined as the reciprocal of the average distance a particle travels down the field (from the point of release until the first ionizing collision) subject to this density of ionizing collisions. The set of universal curves to which we shall fit the experimental data have been calculated numerically using the assumption of simple parabolic bands and a constant mean free path. In fitting these curves there are three physical parameters of importance: The total mean free path for scattering, λ;

[19] R. P. Ruth, J. C. Marinace, and W. C. Dunlap, Jr., J. Appl. Phys. **31**, 995 (1960).
[20] D. Kahng, C. O. Thomas, and R. C. Manz, J. Electrochem. Soc. **109**, 1106 (1962).

the ionization threshold energy, E_i; and a constant energy loss associated with the scattering which is assumed to be equal to the transverse optical phonon energy, E_R. This last assumption, of optical phonon emission, is substantiated by measurements of the temperature dependence of the ionization rate which will be discussed at the end of this section.

Using a transverse optical-phonon energy of 0.06 eV[21] and a threshold energy of 1.6 eV, we have fitted both the electron and hole ionization-rate data by adjusting the mean free path for scattering as shown in Figs. 13 and 14. Although we have chosen a threshold energy of one and a half times the energy gap in these figures, the uncertainties of the experimental data would give an equally good fit for any choice of threshold energy in the range $E_g<E_i<1.5E_g$ with an appropriate adjustment of λ. A way of assessing the quality of the fit is to plot the value of λ which matches an experimental point to the appropriate member of the family of theoretical curves versus E_i and compare this curve to a similar curve which gives the λ chosen to correctly match the slope of the experimental data to that of the appropriate theoretical curve as a function of E_i. The two resulting curves are approximately parallel in the range $E_g<E_i<1.5E_g$ and diverge for $E_i>1.5E_g$.

This process of adjusting the position and slope of the experimental data reveals the explicit difficulty of fitting the earlier published data on silicon.[3,5] When an experimental point of that published data is matched to a theoretical curve for a given E_i and λ, then the slope of the experimental data is sufficiently smaller than that of the theoretical curve so that the departure of the other points greatly exceeds the experimental error. This situation results for any choice of threshold energy in the range $E_g<E_i<4E_g$. Further on, we shall discuss the effect of reducing the ratio of the ionization cross section to the total cross section in relation to improving the fit of the older data.

For a particular choice of threshold energy, as in Figs. 13 and 14, the uncertainty in adjusting λ is only a few angstroms. If, however, we had chosen a threshold energy close to the energy gap, λ would have been somewhat smaller; e.g., in the case of electrons λ would have been reduced from 70 Å to about 50 Å. Keeping in mind this larger uncertainty in λ, it is instructive to compare the λ's of the different units for a given choice of threshhold energy. Unit 6AG35 which had a 6-V breakdown[22] and unit 12H which had a breakdown of 95 V have mean free paths for scattering of 62 and 68 Å, respectively. The electron data of unit 29-9 of Fig. 14 indicates a λ of 69 Å. It appears that the varying degrees of crystal imperfection of these units do not significantly alter the total

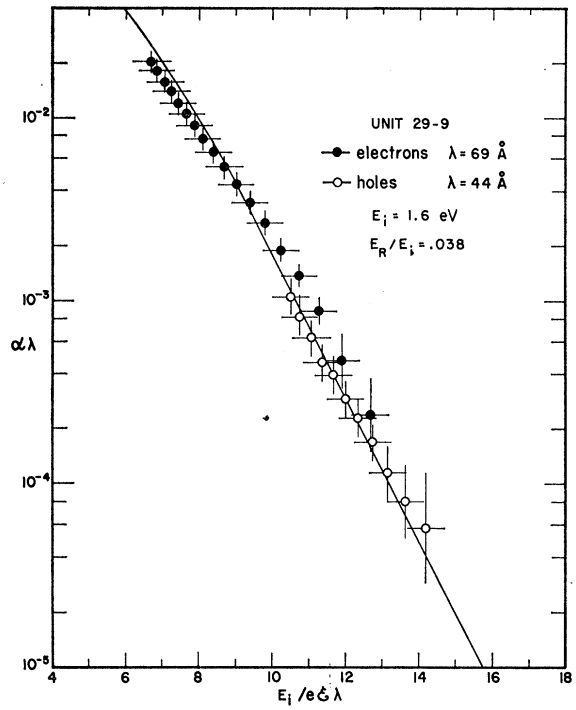


FIG. 14. Fit to theoretical curve of electron and hole ionization rate measured in the same junction for the same threshold energy but different mean free paths for scattering.

mean free path for scattering more than the expected variations in conductivity mobility. If the contribution to the total scattering by these crystal imperfections and impurities is small, as we have reason to believe, then this total mean free path for scattering closely approaches the mean free path for optical phonon scattering.

The electron data of unit 43 is an exception, as we have mentioned before, because the large degree of crystal imperfection in the epitaxial layer is observed to affect λ to qualitatively the same degree as it affects the conductivity mobility. Indeed, one might not expect that the characteristic energy loss of an energetic carrier interacting with such crystal imperfections would be the emission of optical phonons, but the ability to fit this data to the same member of the theoretical curves by merely adjusting λ allows such an interpretation.

The hole ionization rate data of unit 29-9 may also be fitted to the same threshold energy with a mean free path of 44 Å. It would seem logical to interpret this result for the hole data as an indication that holes are more tightly coupled to the lattice. On this point a comparison to germanium becomes interesting because there, conversely to the case in silicon, the hole ionization rate is greater than that of electrons.

Fitting the ionization rates obtained at various temperatures to the curves of Baraff gives the temperature variation of the mean free path $\lambda(T)$. This temperature variation should correspond to the dependence of the

[21] B. N. Brockhouse, Phys. Rev. Letters 2, 256 (1959).

[22] Width corrections for unit 6AG35 corresponding to the distance required to reach the ionization threshold energy have not been applied to these data. Such corrections are less than the indicated experimental error and do not significantly alter the interpretation of these data.

probability of optical phonon emission upon temperature. Similar considerations by Maeda and Suzuki[23] did not enjoy the advantage of the more exact theoretical treatment of Baraff, and further, they only considered the temperature variation of the breakdown voltage rather than that of the ionization rate; imperfections in a junction will produce much greater variations in the breakdown voltage than in the ionization rate which may be measured with some accuracy at only two-thirds of the breakdown voltage.

Using the procedures outlined above, the ionization rate of electrons versus electric field was obtained for a junction with a low breakdown voltage (6 V) at three fixed temperatures: 100, 213, and 300°K. The mean free paths obtained from fitting these data to Baraff's curves are: $\lambda(100°K)=75$ Å, $\lambda(213°K)=72$ Å, and $\lambda(300°K)=68$ Å. The corresponding phonon-emission probabilities $[1+N=1+\exp(-\hbar\omega/kT)]$ are 1.001, 1.043, and 1.106 at 100, 213, and 300°K, respectively, and are inversely proportional to the mean free paths as expected. Indeed, the excellence of the agreement between the ratios of the mean free paths and the optical phonon-emission probability raises the question of whether it is fortuitous in view of the uncertainty in determining E_i and λ. It should be noted that the agreement remains good when comparing the ratios of a set of λ's derived from any threshold energy in the range $E_g<E_i<1.5E_g$. In order to expect good agreement one would require that E_i be nearly temperature-independent, that λ represent optical phonon scattering only, and that the ionization cross section not be a strong function of the temperature. In regard to this last point, Baraff's work[1] shows that the ionization rate is relatively independent of r (the ratio $\sigma_i/\sigma_{\text{total}}$); changes of σ_i are found to be compensated by more or less carriers diffusing to higher energies. The apparent likelihood of these restrictions on E_i, λ, and σ_i imply that the agreement obtained for the temperature variation should be considerably better than the uncertainty of determining E_i and λ.

In view of its practical interest, we have calculated for this junction the temperature coefficient of the breakdown field and obtained the value $\mathcal{E}_B^{-1}(\partial\mathcal{E}_B/\partial T)\approx 7\times10^{-4}/°K$. This value is considerably larger than values given in earlier work. The difference arises from an unexpectedly large temperature variation of the junction transition width at breakdown. The variation in junction width seems to be due to carrier freeze-out at lower temperatures; this variation in junction width was previously thought to be negligible.

Apart from the fact that all of the electron and hole ionization-rate data may be fitted to the theoretical curves within the experimental error, the slopes are consistently smaller than the theoretical curve. It cannot, of course, be asserted that this difference in slope is significant but it is sufficiently suggestive to warrant considering a change in one of the assumptions in the theoretical calculations which would improve the fit. In his calculation of the ionization-rate curves, Baraff has assumed that at the threshold energy the cross section for ionizing collisions was equal to one-half the total cross section; but, in addition, he also considers the effect on the collision density of varying this fraction of the cross sections from 0.05 to 1. The lower values correspond to the indications of the work of Shockley[7] and of Bartelink, Moll, and Meyer[24] that a very energetic carrier will experience on the average several phonon collisions before having an ionizing collision. Thus if the cross section for an ionizing collision were appreciably smaller than the total cross section, the ionization rate would be decreased and a larger mean free path would be required to fit the experimental data to a given theoretical curve. This larger λ for a given E_i would bring the slope of the experimental data and the theoretical curve into closer agreement.

The possibility of improving the fit of the present data by altering the relative cross section for ionizing collisions is also pertinent to previously published data.[3,5] Before it may be discussed adequately, however, it is necessary to outline more quantitatively the magnitude of the adjustments that can be reasonably made. In the previous discussion we have touched upon Baraff's estimation of the relatively weak dependence of the ionization rate upon the ratio $\sigma_i/\sigma_{\text{total}}$. A variation of r of almost two orders of magnitude results in a change of α by only a factor of several. Thus if α is lowered by a reduction in r (a reduction that may considerably exceed the estimates of Shockley[7] and of Bartelink *et al.*[24]) given ionization-rate data may be fit to these new curves by readjusting λ by less than 20%. Since the slope of the fitted experimental data is inversely proportional to λ, one cannot reasonably expect to correct errors in slope of more than 20%. With this magnitude of adjustment available, it should be possible to fit the electron ionization-rate data of Moll and Overstraeten[5] and of Chynoweth[3] within a reasonable experimental error.[25] The hole ionization-rate data of both these papers, however, cannot be fitted by such an adjustment because the slopes are too small by 50–100%. The hole data of Moll and Overstraeten is distorted by the assumption of the proportionality of the hole and electron ionization rates and the hole data of Chynoweth was not obtained with a sufficiently pure saturation current of holes.

ACKNOWLEDGMENTS

We wish to acknowledge our appreciation of the continued interest and helpful discussion given by Dr. G.

[23] K. Maeda and K. Suzuki, Japan. J. Appl. Phys. **1**, 193 (1962).

[24] D. Bartelink, J. L. Moll, and N. Meyer, Phys. Rev. **130**, 972 (1963).

[25] There are reservations that must be stated for each of these papers: (1) Moll and Overstraeten have assumed a parabolic field distribution in solving the integral equation relating M and α whereas their capacitance data indicates a step junction behavior; the correction of this approximation (see our Appendix A) might adversely affect the fitting of their data. (2) A portion of Chynoweth's data towards higher fields must be excluded.

Baraff. We also wish to thank G. Kaminsky and S. Sumski for considerable help with the measurements and sample preparation.

APPENDIX A

In order to estimate the effects of deviations from a strictly parabolic field distribution, we first consider the electric field distribution of our exponentially graded junctions:

$$y=1-\frac{\mathcal{E}}{\mathcal{E}_m}=B(\xi-1+e^{-\xi}), \tag{11}$$

where

$$\xi\equiv x/L, \quad \text{and} \quad B\equiv qN_bL/\epsilon\mathcal{E}_m.$$

Then expanding ξ in powers of y, we obtain for $\xi>0$

$$\xi_+=\left(\frac{2}{B}\right)^{1/2}y^{1/2}+\frac{1}{3}\left(\frac{2}{B}\right)\frac{y}{2!}+\frac{1}{6}\left(\frac{2}{B}\right)^{3/2}\frac{y^{3/2}}{3!}+\cdots, \tag{12}$$

and for $\xi<0$

$$-\xi_-=\left(\frac{2}{B}\right)^{1/2}y^{1/2}-\frac{1}{3}\left(\frac{2}{B}\right)\frac{y}{2!}+\frac{1}{6}\left(\frac{2}{B}\right)^{3/2}\frac{y^{3/2}}{3!}-\cdots. \tag{13}$$

Next we wish to evaluate the integral

$$\int_{x_a}^{x_b}\alpha dx=L\int_{\xi_a}^{\xi_b}\alpha d\xi=L\int_0^1\alpha\left[\left(\frac{2}{B}\right)^{1/2}y^{-1/2}+\frac{1}{12}\left(\frac{2}{B}\right)^{3/2}y^{1/2}\right]dy. \tag{14}$$

For our purpose of estimating errors, it is sufficient to take

$$\alpha=A\mathcal{E}^r=\alpha(\mathcal{E}_m)(1-y)^r. \tag{15}$$

Then

$$\int_{x_a}^{x_b}\alpha dx=\alpha(\mathcal{E}_m)L\left\{2\left(\frac{2}{B}\right)^{1/2}\frac{1}{2}\int_0^1(1-y)^ry^{-1/2}dy+\frac{1}{6}\left(\frac{2}{B}\right)^{3/2}\frac{1}{2}\int_0^1(1-y)^ry^{1/2}dy\right\}, \tag{16}$$

and we observe that

$$m=\frac{1}{2}\int_0^1(1-y)^ry^{-1/2}dy=\frac{\Gamma(1/2)\Gamma(r+1)}{2\Gamma(r+3/2)}, \tag{17}$$

this last integral being an Eulerian integral of the first kind. We may further reduce this result to the form[26]

$$\int_{x_a}^{x_b}\alpha dx=\alpha(\mathcal{E}_m)L\left[2\left(\frac{2}{B}\right)^{1/2}\right]m\left(1+\frac{1}{12}\frac{2}{B}\frac{1}{2r+3}\right). \tag{18}$$

Now, for the restriction we have imposed on these junctions that $\omega=(w/L)<3$which implies that $B>1$, and with the observation that $r\gtrsim 8$, we see that the error due to deviations from a strict parabolic field distribution is less than 1%.

APPENDIX B

It will be sufficient to give the procedure for obtaining the hole ionization rate. From Eq. (10b) we have

$$1-\frac{1}{M_p}=2\exp\left[\int_0^{\mathcal{E}_m}(\alpha-\beta)dy\right]\times\int_0^{\mathcal{E}_m}\beta\cosh\left[\int_{\mathcal{E}}^{\mathcal{E}_m}(\alpha-\beta)dy'\right]dy. \tag{10b}$$

Now if we restrict the range of hole multiplication such that

$$\int_0^{\mathcal{E}_m}(\alpha-\beta)dy=\tfrac{1}{2}\ln\frac{M_n}{M_p}<0.5, \tag{19}$$

then we may expand the right side of Eq. (10b) to

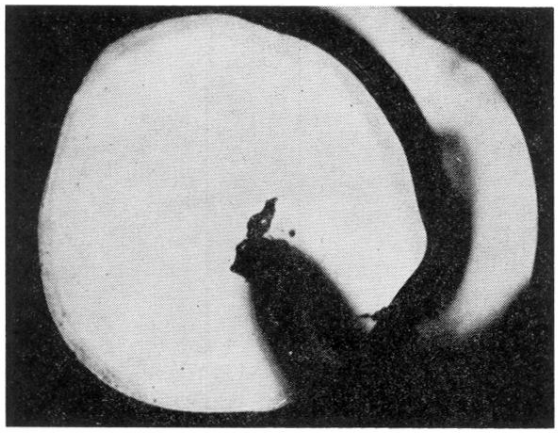
$$1-\frac{1}{M_p}\simeq 2\left(\frac{M_n}{M_p}\right)^{1/2}\int_0^{\mathcal{E}_m}\times\beta\left\{1+\frac{1}{2}\left[\int_{\mathcal{E}}^{\mathcal{E}_m}(\alpha-\beta)dy'\right]^2\right\}dy \tag{20}$$

and with some manipulation the second term under the integral can be shown to constitute an error of less than 5%. The remaining integral then becomes

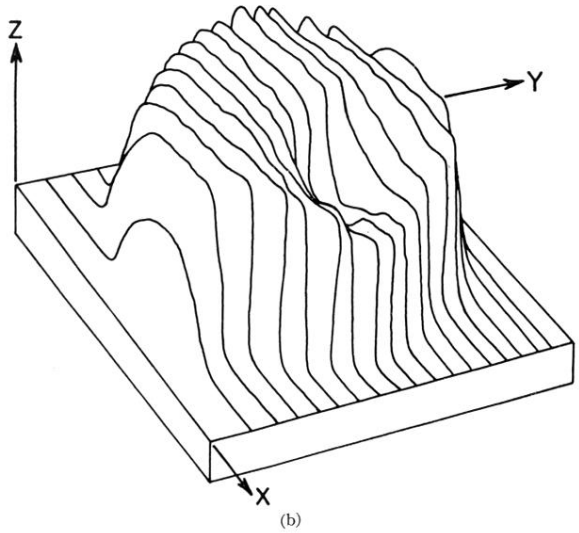
$$1-\frac{1}{M_p}\simeq 2\left(\frac{M_n}{M_p}\right)^{1/2}\frac{w}{4}\int_0^{\mathcal{E}_m}\beta\frac{d(\mathcal{E}/\mathcal{E}_m)}{(1-\mathcal{E}/\mathcal{E}_m)^{1/2}}, \tag{21}$$

and this can be solved by Abel's integral equation as shown by McKay.[14]

[26] Note from Eq. (3) that for $\omega\ll 1$ (linear gradient) $2(2/B)^{1/2}\to\omega$ and $\int\alpha dx\to\alpha(\mathcal{E}_m)wm$, an approximation used by K. G. McKay (unpublished work) and later by Chynoweth (Ref. 3). The analytical form of Eq. (17), however, has not been given previously.



(a)



(b)

FIG. 5. (a) Photograph of etched mesa of exponentially graded junction. (b) Local multiplied photocurrent of junction pictured in Fig. 5(a). The x, y plane is coincident with the plane of the junction. Each profile represents a recorder trace of a 1-mil light spot traversing the junction. The separation of the traces is also 1 mil. The mean multiplication is 100 and the depression arises from the shadowing by the point contacting the surface.