Solid-State Electronics, 1975, Vol. 18, pp. 161–168. Pergamon Press. Printed in Great Britain

THRESHOLD ENERGY EFFECT ON AVALANCHE BREAKDOWN VOLTAGE IN SEMICONDUCTOR JUNCTIONS*

Y. OKUTO† and C. R. CROWELL

Department of Materials Science, University of Southern California, Los Angeles, California 90007, U.S.A.

(*Received* 22 *April* 1974; *in revised form* 20 *July* 1974)

Abstract—The band bending for avalanche breakdown in semiconductor junctions and its temperature dependence are predicted taking account of threshold energy effects on the ionization process in semiconductors. Where experimental results exist, the theoretical predictions and experimental results are in excellent agreement. In the high electric field region inclusion of both bulk and boundary threshold energy effects is essential. The predictions were based on exact solutions in the nonlocalized ionization coefficient formulation developed by Okuto and Crowell who showed that ionization coefficients as usually understood are functions of both electric field and position in a device. Predictions for abrupt and p–i–n junctions in Ge, Si, GaAs and GaP are presented.

1. INTRODUCTION

A knowledge of the reverse breakdown voltage is necessary for optimum use of any junction device. Avalanche at low through high electric fields and tunneling at ultra high electric fields are the two breakdown mechanisms. The field range at which the transition occurs between the two mechanisms is not known precisely. This is partly because analysis of avalanche phenomena in this field range has been hampered by an inadequate treatment of the effects of the threshold energy for ionization scattering. In reality this affects not only the boundary conditions but also the definition of the ionization coefficient itself. Physically the avalanche breakdown voltage, V_A, is defined as the applied voltage where electron and hole multiplication rates (M_n and M_p, respectively) reach infinity. Here multiplication rates are defined as ratios of number of carriers extracted to number of carriers injected. Experimentally, since the infinite current condition cannot be obtained, one must make some type of extrapolation to define the breakdown voltage[1]. The theoretical estimate, however, can be defined somewhat more precisely as the potential drop across the junction at which the following condition is satisfied:

$$\lim_{M_n\to\infty}\left\{1-\frac{1}{M_n}\right\}=1=\int_0^W \alpha_n(x)\exp\int_x^W(\alpha_p(x')-\alpha_n(x'))\,dx'\,dx, \tag{1}$$

where $\alpha_n(x)$ and $\alpha_p(x)$ are electron and hole ionization coefficients at the point x and electrons are injected at $x = W$. 0 and W denote coordinates where the electric field, $\mathscr{E}$, terminates. The potential difference between $x = 0$ and $x = W$, V_B, is the total band bending at breakdown. The breakdown voltage, V_A, is less than V_B by the zero bias diffusion potential, V_D, i.e.,

$$\int_0^W \mathscr{E}(x)\,dx = V_B = V_A + V_D. \tag{2}$$

Previously V_B has been estimated[2] with the use of equation (1) by assuming α_n and α_p are only functions of the electric field strength. Existing predictions underestimate V_B for high electric field configurations[3, 4]. Recently the importance of the threshold energy correction has been emphasized by many authors[5–7]. Some improvements have been obtained[7] but the threshold energy corrections have been limited to boundary dark spaces for injected carriers. Recently Okuto and Crowell (O–C)[8] have shown that once the full implications of the threshold energy corrections are taken into account the basic understanding of the ionization effect itself has to be modified (e.g. the ionization coefficients as they are commonly understood to be defined become functions of position as well as electric field strength). O–C have suggested that the fundamental field dependent quantity is what they have called a non-localized ionization coefficient.

In this paper we examine the O–C results and show how they can be used to predict V_B. To simplify the calculation we introduce a stylized model for ionization in junctions (in section 2) and an analytical expression to represent

*Work supported by the Joint Services Electronics Program through the Air Force Office of Scientific Research/AFSC under contract F-44620-71-C-0067.

†Present address: Nippon Electric Co., Ltd., Central Research Labs., 1753 Shimonumabe, Kawasaki, Japan.

apparent ionization coefficients via a pseudolocal approximation (in section 3) as outlined earlier by O–C[8]. In the pseudolocal approximation, ionization coefficients do not have their customary meaning as pure electron and hole ionization coefficients but they are dependent solely on the electric field[8]. In section 4, predicted V_B values are presented for abrupt and p–i–n junctions in Ge, Si, GaAs and GaP at various temperatures.

The results show excellent agreement with existing experimental data for Si and GaAs abrupt junctions.

2. STYLIZED MODEL

One example of the exact solution of the formulation proposed by Okuto and Crowell is shown in Fig. 1.* Figure 1 shows room temperature electron and hole concentrations and electron and hole ionization coefficients as functions of position for a Si p–i–n diode with a 0·5 μm i-layer. The conventional definition of the ionization coefficient is stated below. The band bending is 20·9 V and M_p is 196. The horizontal axis is the distance from the p–i junction. D_n and D_p are dark space distances (measured from the p–i and the i–n junctions, respectively) for electrons and holes. These distances are defined as

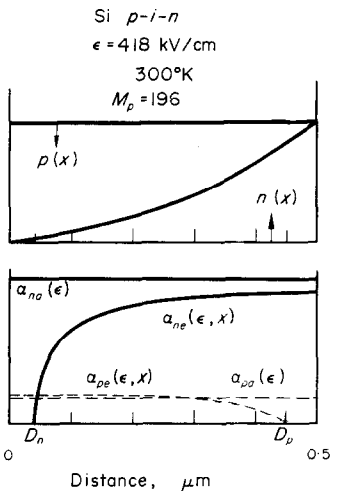

Fig. 1. An exact solution for a Si p–i–n junction with 0·5 μm i-region width at 300°K for a hole injection case. Band bending is 20·9 V and hole multiplication factor is 196. The upper portion shows $n(x)$ and $p(x)$ as functions of position. The lower portion shows position dependent ionization coefficients obtained from the exact formulation $\alpha_{ne}(\mathscr{E}, x)$ for electrons and $\alpha_{pe}(\mathscr{E}, x)$ holes respectively. Position independent ionization coefficients obtained via the pseudolocal approximation $\alpha_{na}(\mathscr{E})$ for electrons and $\alpha_{pa}(\mathscr{E})$ for holes are also shown. D_n and D_p are dark spaces for electrons (measured from $x = 0$) and for holes (measured from $x = 0{\cdot}5$ μm), respectively.

$$D_n \equiv E_{in}/q\mathscr{E} \tag{3}$$

and

$$D_p \equiv E_{ip}/q\mathscr{E} \tag{4}$$

where E_{in} and E_{ip} are ionization threshold energies for electrons and holes[9], respectively and q is the electronic charge. $\alpha_{ne}(\mathscr{E}, x)$ and $\alpha_{pe}(\mathscr{E}, x)$ are position dependent ionization coefficients which are defined as

$$\alpha_{ne}(\mathscr{E}, x) \equiv \frac{1}{n(x)} \frac{dn_n(x)}{dx} \tag{5}$$

and

$$\alpha_{pe}(\mathscr{E}, x) \equiv \frac{1}{p(x)} \frac{dn_p(x)}{dx}. \tag{6}$$

Here $n(x)$ and $p(x)$ are densities of electrons and holes at x and $dn_n(x)/dx$ and $dn_p(x)/dx$ are gradients in the densities of electrons at x originated by electrons and by holes, respectively. $\alpha_{na}(\mathscr{E})$ and $\alpha_{pa}(\mathscr{E})$ are the position independent apparent (or pseudolocal) ionization coefficients for electrons and holes which O–C obtained from a situation in which only bulk threshold energy effects were considered.† The exact result shows the following characteristic features, i.e.

(1) Electrons and holes have zero ionization coefficients within corresponding boundary dark spaces.

(2) Throughout the rest of the region, the ionization coefficients vary slowly with distance and tend towards asymptotic values. The tendency to reach close to asymptotic values is more apparent in wider space charge regions[8].

(3) $\alpha_{ne}(\mathscr{E}, x)$ and $\alpha_{pe}(\mathscr{E}, x)$ are fairly constant and their asymptotic values are close to those of $\alpha_{na}(\mathscr{E})$ and $\alpha_{pa}(\mathscr{E})$ respectively.

These results show why earlier predictions underestimated V_B, especially for high field configurations: the boundary dark spaces introduce a potential drop which does not contribute directly to ionization. The threshold energy effect also tends to suppress the positive feedback effect which is essential for breakdown. In principle, the best estimate of V_B would be obtained by solving the O–C formulation for the case of infinite multiplication. In reality, however, this approach demands appreciable computer time and the formulation is very complicated when an electric field gradient exists. Thus a stylized model to reproduce the exact solution is proposed here. In this model electron and hole apparent ionization coeffi-

*More examples are presented in Ref. [8].

†O–C have shown that $\alpha_{na}(\mathscr{E})$ and $\alpha_{pa}(\mathscr{E})$ are in satisfactory agreement with existing experimental data[9] but that $\alpha_{na}n \neq dn_n/dx$ and $\alpha_{pa}p \neq dn_p/dx$.

cients are assumed to be zero within their corresponding boundary dark spaces and are assumed to be $\alpha_{na}[\mathscr{E}(x)]$ and $\alpha_{pa}[\mathscr{E}(x)]$ through the rest of the region. Figure 2 shows this situation schematically for a p–i–n junction. This stylized model has the following relatively minor differences from the exact solutions:

(1) In the exact solution the ionization coefficients grow slower than the step functions assumed here. This tendency is stronger for high field configurations and produces an underestimate of V_B.

(2) The apparent ionization coefficients are larger for the species with the larger ionization coefficient (e.g. for electrons in Si) and smaller for the species with the smaller ionization coefficient (e.g. for holes in Si). This is due to boundary effects on the feedback which were not taken into account when $\alpha_{na}(\mathscr{E})$ and $\alpha_{pa}(\mathscr{E})$ were deduced.

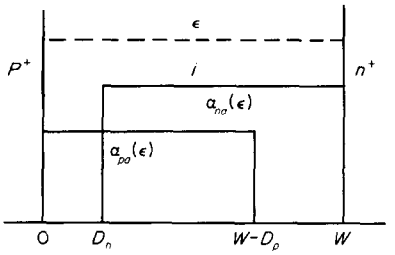

Fig. 2. Stylized model to approximate the distributed ionization effect in a p–i–n junction.

3. ANALYTICAL CURVE FITTING

The precise expressions for the pseudolocal ionization coefficients are too complicated algebraically for easy use in practical device characterization. Thus we have attempted to reproduce the results with a simple analytical expression. The trial function had the form

$$\alpha(\mathscr{E}) = a\mathscr{E}^n \exp\{-(b/\mathscr{E})^m\}. \tag{7}$$

Here a and b are constants and n and m are integers. A conventional least square fitting procedure was used to determine these constants. This method was examined in various materials at various temperatures. The choice of $n = 1$ and $m = 2$ was found to be the best for all cases. a and b were found to be linear functions of the lattice temperature around 300°K, i.e.

$$a = a_{300}\{1 + c(T - 300)\} \tag{8}$$

and

$$b = b_{300}\{1 + d(T - 300)\}. \tag{9}$$

The values of the constants, range of the electric field over which the fitting is valid, and the average errors are listed in Table 1 along with the threshold energy[9], optical phonon energy[10] and also mean free path for optical phonon scattering[8] at 300°K. Figure 3 shows the ratio of the ionization coefficients obtained by the present expression to that obtained via the pseudolocal approximation for electrons in Si at 300°K as a function of the electric field strength. The agreement of the simple expression with the pseudolocal ionization coefficients is

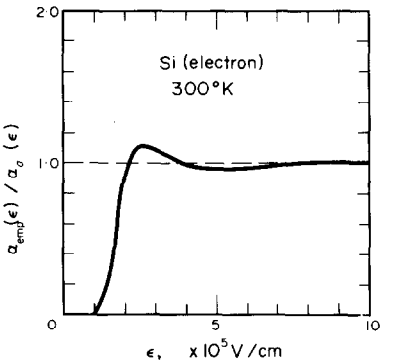

Fig. 3. Ratio of the ionization coefficients obtained by the present empirical expression to that obtained via the pseudolocal approximation for electrons in Si at 300°K as a function of electric field strength.

Table 1. Apparent ionization parameters (300°K)

		Ge		Si		GaAs	GaP
e/h		e	h	e	h	e/h	e/h
E_i	eV	0·8	0·9	1·1	1·8	1·7	2·6
E_r	meV	22		53		22	37
λ	Å	39	51	48	47	33	31
a_{300}	V^{-1}	0·569	0·559	0·426	0·243	0·294	0·191
c	$\times 10^{-4}$ °K^{-1}	6·33	7·87	3·05	5·35	8·50	8·38
b_{300}	$\times 10^5$ V/cm	3·32	2·72	4·81	6·53	5·86	9·91
d	$\times 10^{-4}$ °K^{-1}	9·34	8·82	6·86	5·67	7·17	5·95
$\mathscr{E}$	$\times 10^5$ V/cm	0·7–7		1–10		1–20	3–40
Error	%	3	4	4	4	5	4

Expression $\alpha(\mathscr{E}) = a_{300}\{1 + c(T - 300)\}\mathscr{E} \exp[-[b_{300}\{1 + d(T - 300)\}/\mathscr{E}]^2]$.

excellent over a wide enough range of the electric field to be applicable to most practical devices.

4. BREAKDOWN BAND BENDING ESTIMATE AND DISCUSSION

The V_B estimate has been performed with the use of the basic equation

$$1-\frac{1}{M_n}=\int_0^W \alpha_n(\mathscr{E}(x),x)\exp\int_x^W[\alpha_p(\mathscr{E}(x'),x') -\alpha_n(\mathscr{E}(x'),x')]\,dx'\,dx \tag{10}$$

Here

$$\alpha_n(\mathscr{E}(x),x)\equiv 0 \quad \text{for} \quad 0\le x<D_n \tag{11}$$

$$\alpha_p(\mathscr{E}(x),x)\equiv 0 \quad \text{for} \quad W-D_p<x\le W \tag{12}$$

and in the rest of the region the electron ionization coefficient was assumed to be expressed as

$$\alpha_n(\mathscr{E}(x),x)=a_n\mathscr{E}(x)\exp\{-(b_n/\mathscr{E}(x))^2\}. \tag{13}$$

A corresponding expression was used for α_p (cf. Table 1). Equations (10)–(12) can be combined into a single equation expressible in the following form (cf. Appendix).

$$1-\frac{1}{M_n}\exp\left\{\int_{D_n}^W \alpha_n(x)\,dx-\int_0^{W-D_p}\alpha_p(x)\,dx\right\} \tag{14}$$

$$=\int_{D_n}^{W-D_p}\alpha_p(x)\exp\left[\int_x^{W-D_p}\{\alpha_n(x')-\alpha_p(x')\}\,dx'\right]dx\cdot\exp\int_{W-D_p}^W\alpha_n(x)\,dx$$

$$+\int_0^{D_n}\alpha_p(x)\exp\left[\int_x^{D_n}-\alpha_p(x')\,dx'\right]dx\cdot\exp\left\{\int_{D_n}^W\alpha_n(x)\,dx-\int_{D_n}^{W-D_p}\alpha_p(x)\,dx\right\}$$

Note that when $D_n\to 0$ and $D_p\to 0$, equation (14) becomes the conventional relationship. This relationship is the one which should be used to deduce apparent ionization coefficients from multiplication measurements.

If one has to use a computer, however, equations (11)–(13) are just as good as equation (14). Figures 4 and 5 show results for Ge, Si, GaAs and GaP abrupt and p–i–n junctions respectively. Here the ordinates are the breakdown band bending in units of volts. In Fig. 4 the abscissa is the background impurity density and in Fig. 5 the abscissa is the width of the i-region. In both figures one notices the saturation of V_B. This tendency is more apparent in p–i–n junctions than in step junctions. This is due to the threshold energy effect. This limit can easily be derived as follows for p–i–n junctions when $\alpha_n=\alpha_p$. The conventional breakdown condition provides us with a minimum estimate of the breakdown band bending. This condition is

$$\alpha_a\cdot W\simeq 1 \tag{15}$$

According to Okuto and Crowell[8]

$$\alpha_a\cdot D\to 1/2 \quad \text{as} \quad \mathscr{E}\to\infty \tag{16}$$

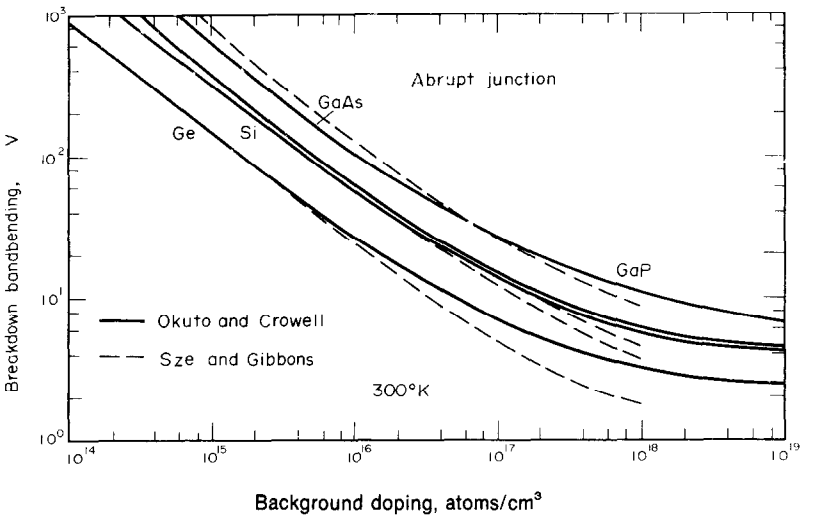

Fig. 4. Breakdown band bending as a function of the background impurity density for abrupt junctions at 300°K in selected semiconductors. Legend: ———— present prediction — — — — — previous estimate[2].

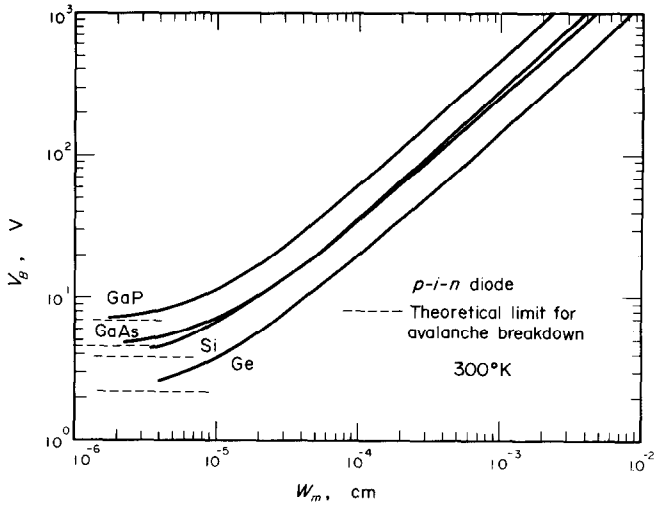

Fig. 5. Breakdown band bending as a function of width of i-region for p–i–n junctions at 300°K in selected semiconductors. Legend: ———— present prediction, — — — — — minimum breakdown band bending obtained by equation (18).

From equations (15) and (16) one obtains

$$\frac{1}{W}\cdot\frac{E_i}{q\mathscr{E}}\approx\frac{1}{2}. \tag{17}$$

Thus

$$V_{B\min}\equiv\mathscr{E}\cdot W\approx 2E_i/q. \tag{18}$$

For the $\alpha_n\neq\alpha_p$ case the situation is more complicated but one can also write as the analogy to equation (18)

$$V_{B\min}\geq(E_{in}+E_{ip})/q. \tag{19}$$

Equations (18) and (19) can be understood physically as follows. When the electric field increases, the apparent ionization coefficients increase. At an infinitely large field they also become infinitely large. Nevertheless, both boundary dark spaces require corresponding threshold potential drops. Thus the minimum band bending cannot become less than the sum of these two potentials [cf. equations (18) and (19)]. This V_B can be used as a measure to identify the breakdown mechanism. Namely once a junction shows less than $V_{B\min}$, the breakdown mechanism is due to tunneling. Note, that the converse of this statement is not necessarily true. The maximum electric

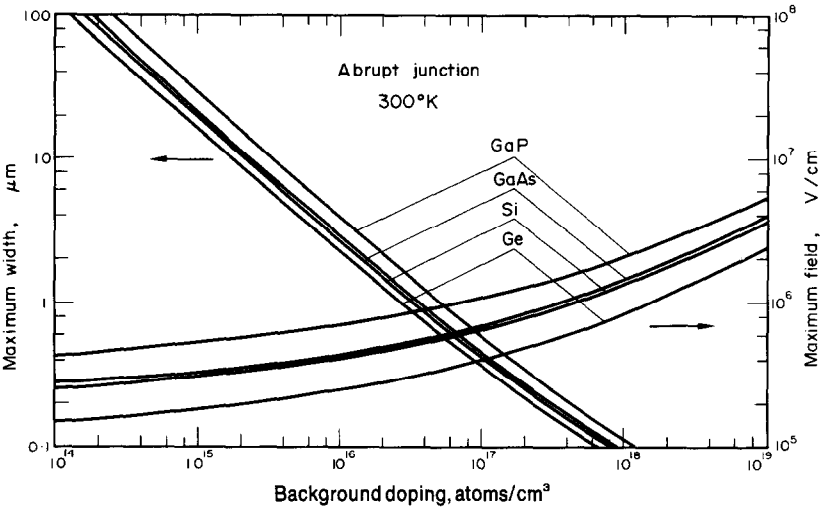

Fig. 6. The maximum electric field strength and width of the space charge region at breakdown for abrupt junctions as a function of the background doping at 300°K for selected semiconductors.

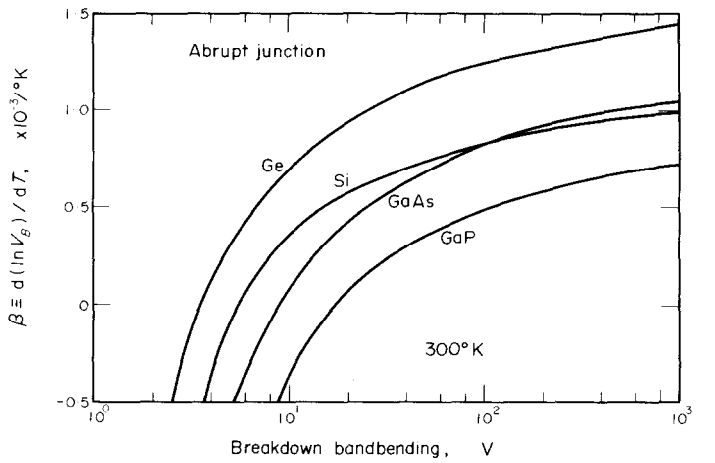

Fig. 7. Logarithmic temperature coefficients of the breakdown band bending for abrupt junctions as a function of the breakdown band bending at 300°K for selected semiconductors.

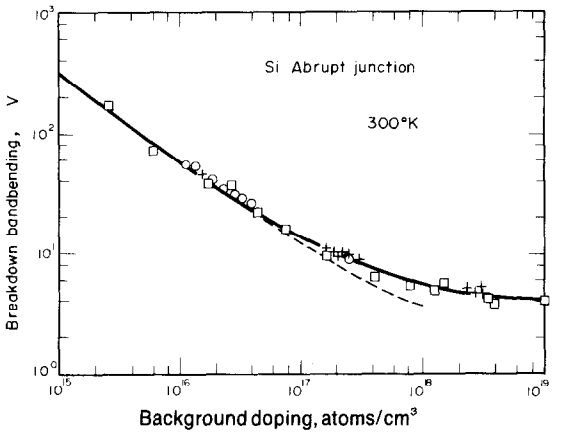

Fig. 8. Comparison of the breakdown band bending obtained experimentally and that obtained theoretically for Si abrupt junctions at room temperature. Legend: ———— present prediction, – – – – – – previous estimate[2], □ experimental results after Miller[11] (diffusion potential excluded); +, ○ after Okuto and Crowell[12].

field strength, $\mathscr{E}_{max}$, and width of the space charge region at breakdown, W, for abrupt junctions are shown in Fig. 6 as a function of the background impurity density. The predicted logarithmic temperature coefficients, β, of the breakdown band bending around room temperature for abrupt junctions are shown in Fig. 7. Here β is defined as

$$\beta = \frac{\mathrm{d} \ln V_B(T)}{\mathrm{d}T} \tag{20}$$

The tendency of the results agrees with existing experimental results[11]. Note that β seems to become negative for high field configurations: for high field configurations, the temperature dependence of the threshold energy (which is assumed to be the same as that of the band gap energy) becomes of major importance.

Figures 8 and 9 show a comparison of experimental data[11–14] and present predictions for Si and GaAs abrupt junctions at room temperature. Here most of the experimental data[11, 13, 14] were obtained from diffused or alloyed junctions and do not include the diffusion potential correction. This correction is important for junctions with low breakdown voltages. The data obtained by Okuto and Crowell[12] include this correction since their data were obtained from Schottky barrier configurations. For both Si and GaAs, the experimental data show excellent agreement with present predictions. This indicates the importance of the threshold energy effect and partially justifies the stylized model proposed here. The experimental results also seem to indicate that the breakdown mechanism is still due to avalanche in Si up to impurity densities of the order of 1×10^{19} atoms/cm^3. This result was not expected previously. The exact transition region is not known, however, since Miller's p^+–n junctions near 10^{19} doping may have had somewhat graded impurity profiles. The data in this region are obviously not sensitive to the exact details of an impurity profile. This is also apparent from the form of equation (7) (when $n = 1$):

$$\alpha \delta x \approx a \delta V \tag{21}$$

in the asymptotic high field limit. In this limit, equation (14) becomes

$$1 - \frac{1}{M_n} \exp\left[(V - E_{in}/g)a_n - (V - E_{ip}/g)a_p\right]$$
$$= \frac{a_p \exp(a_n E_{ip}/g)}{a_n - a_p} \{\exp[(a_n - a_p)(V - E_{ip}/g - E_{in}/g)] - 1\}$$
$$+ \exp[a_n(V - E_{in}/g) - a_p(V - E_{ip}/g)]$$
$$\times \{\exp(a_p E_{in}/g) - 1\}. \tag{22}$$

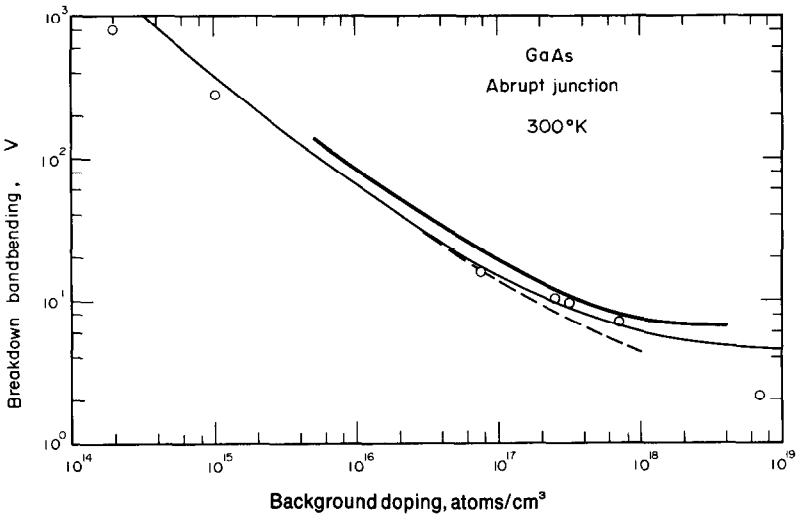

Fig. 9. Comparison of the breakdown band bending obtained experimentally and that obtained theoretically for GaAs abrupt junctions at room temperature. Legend: ———— present prediction, — — — — — previous estimate [2] ————, after Kressel *et al.* [13], ○ after Weinstein and Mlavsky [14].

When $a_n = a_p = a$ and $E_{in} = E_{ip} = E_i$, the criterion for the breakdown bandbending, V_B, becomes

$$V_B = 2E_i/g + (2\exp(-aE_i/g) - 1)/a. \tag{23}$$

Since $a \approx 1/(2E_i)$ in this case [8], then

$$V_B = \frac{4}{\sqrt{e}}\frac{E_i}{g} \tag{24}$$

REFERENCES

1. Y. Okuto, *Jap. J. appl. Phys.* **8**, 917 (1969).
2. S. M. Sze and G. Gibbons, *Appl. Phys. Letts*, **8**, 111 (1966).
3. A. S. Grove, *Physics and Technology of Semiconductor Devices*, p. 195, Wiley, New York (1967).
4. H. Kressel, *RCA Rev.* **28**, 175 (1967).
5. R. van Overstraeten and H. DeMan, *Solid-St. Electron.* **13**, 583 (1970)
6. M. H. Woods, W. G. Johnson and M. A. Lampert, *Solid-St. Electron.* **16**, 381 (1973).
7. W. N. Grant, *Solid-St. Electron.* **16**, 1189 (1973).
8. Y. Okuto and C. R. Crowell, *European Solid State Device Research Conference*, Lancaster, U.K., September 1972; *Phys. Rev.*, **B**, 15 Oct (1974).
9. C. L. Anderson and C. R. Crowell, *Phys. Rev. B.* **5**, 2267 (1972).
10. C. R. Crowell and S. M. Sze, *Appl. Phys. Letts*, **9**, 242 (1966).
11. H. Kressel, A. Blicher and L. H. Gibbons, Jr., *Proc. IRE* **50**, 2493 (1962).
12. Y. Okuto and C. R. Crowell, *Proc Fifth Conf.* (1973 *International*) *on Solid State Devices*, Tokyo, Japan, August 1973, (Suppl) *J. Jap. Soc. appl. Phys.* **43**, 390 (1974).
13. S. L. Miller, *Phys. Rev.* **105**, 1246 (1957).
14. M. Weinstein and A. J. Mlavsky, *Appl. Phys. Letts* **2**, 97 (1963).

APPENDIX

Derivation of equation (14)

In this treatment we assume that n_0 electrons are injected from $x = 0$ and also that electrons and holes have the same saturation velocity.

$$n_t = n(x) + p(x) = \text{constant} \tag{A1}$$

Here n_t is the total density of carriers. One can divide a junction into three regions, i.e.

Region I

$$0 \le x < D_n$$

where

$$E_{in} = g\int_0^{D_n} \mathscr{E}(x)\,dx,$$

$$\alpha_n(x) = 0$$

and

$$\alpha_p(x) = \alpha_{pa}[\mathscr{E}(x)].$$

Region II

$$D_n \le x \le W - D_p$$

where

$$E_{ip} = g\int_{W-D_p}^{W} \mathscr{E}(x)\,dx,$$

$$\alpha_n(x) = \alpha_{na}(\mathscr{E}(x))$$

and

$$\alpha_p(x) = \alpha_{pa}(\mathscr{E}(x)).$$

Region III

$$W - D_p < x \leq W$$

where

$$\alpha_n(x) = \alpha_{na}(\mathscr{E}(x))$$

and

$$\alpha_p(x) = 0.$$

Then the mathematical treatment is straightforward.

Region I

$$\frac{dn(x)}{dx} = \alpha_p(x)p(x) \tag{A2}$$

$$n(D_n) = n_t \int_0^{D_n} \alpha_p(x) \exp - \int_x^{D_n} \alpha_p(x')\,dx'\,dx + n_0 \exp - \int_0^{D_n} \alpha_p(x)\,dx \tag{A3}$$

Region II

$$\frac{dn(x)}{dx} = \alpha_n(x)n(x) + \alpha_p p(x) \tag{A4}$$

Thus

$$n(W - D_p) = n_t \int_{D_n}^{W-D_p} \alpha_p(x) \exp \int_x^{W-D_p} [\alpha_n(x') - \alpha_p(x')]\,dx\,dx + n(D_n) \exp \int_{D_n}^{W-D_p} (\alpha_n(x) - \alpha_p(x))\,dx \tag{A5}$$

Region III

$$\frac{dn(x)}{dx} = \alpha_n(x)n(x) \tag{A6}$$

$$n_t = n(W - D_p) \exp \int_{W-D_p}^{W} \alpha_n(x)\,dx. \tag{A7}$$

By eliminating $n(D_n)$ and $n(W - D_p)$ from equations (A3), (A5) and (A7) one obtains equation (14). It is obvious that one can obtain a corresponding relationship for M_p in the same manner. Equation (14) is a general expression which can be used for any electric field configuration. Note that when W becomes larger, the threshold energy correction becomes less important since band bending increases but the potential drops across both boundary dark spaces stay constant. This correction is also less important for graded junctions since the electric field inside the boundary dark spaces is small and thus the dark space contribution to the ionization integral is small.