JOURNAL OF MATERIALS SCIENCE 3 (1968) 89–98

Anomalous Diffusion Effects in Silicon (A Review)

A. F. W. WILLOUGHBY
Engineering Materials Laboratories, The University, Southampton, UK

This paper is concerned with anomalous diffusion effects in silicon, and particularly with the emitter-dip (or push-out) effect, which occurs in diffused transistors. Several possible mechanisms may be involved in anomalous diffusion and each is discussed, together with relevant experimental and theoretical work. Literature directly on the emitter-dip effect is next reviewed, and finally the author suggests fields of research which may clarify the mechanism or mechanisms involved.

1. Introduction

The importance of the silicon planar diffusion process in integrated circuit technology has emphasised the need for a review of diffusion studies in this material. This review is not intended to cover the whole field of diffusion in silicon; for this the reader is referred to general reviews by Reiss and Fuller [1], Boltaks [2] and Burger and Donovan [3]. The area reviewed will be that concerned with anomalous diffusion effects, and particularly with the emitter-dip effect (or push-out effect) which has hindered the development of very high frequency transistors. This effect will be described in detail in section 4, but it is clear that this, and similar effects, are associated with a change in diffusion characteristics of the base impurity brought about by the introduction of the emitter impurity. Mechanisms by which the presence of one impurity can effect the diffusion rate of another impurity, or of self-diffusion, are described in section 2, and experiments designed to investigate these mechanisms are described in sections 3.1 and 3.2. Section 4 describes experiments directly on the emitter-dip effect.

2. Possible Mechanisms of Anomalous Diffusion Effects

2.1. The Effect of Fermi Level on Diffusion Rates

Self-diffusion, and diffusion of substitutional impurities, may be affected by the Fermi level through the vacancy concentration, as suggested by Longini and Greene [4]. If a vacancy accepts an electron, its free energy of formation is reduced by the electron falling from the Fermi level into the vacancy acceptor level. Thus the equilibrium vacancy concentration, and hence the diffusion coefficient, will be greater in n-type than in p-type material. Valenta and Ramasastry [5] demonstrated this effect when they found that, in heavily doped germanium, the self diffusion coefficient was greater in As-doped material than in intrinsic material, which was also greater than that in Ga-doped crystals.

2.1.1. n-Type Material – Non-Degenerate Statistics

The following derivation, due to Millea [6], affords a method of estimating the magnitude of this effect in n-type material. Assuming that vacancies act as acceptors, the concentration of negatively charged vacancies $[V^-]$ is related to the total number of vacancies, $[V]$ by an expression of the form

$$[V^-] = \frac{[V]}{1 + \exp(E_V - F/kT)} \qquad (1)$$

where E_V is the vacancy acceptor level and F is the Fermi level. (A more detailed treatment, for which a model of the vacancy acceptor must be chosen, would involve a numerical factor multiplying the exponential term in equation 1; the value of this factor depends on the model chosen.)

In thermal equilibrium the concentration of neutral vacancies $[V^0]$ is independent of the Fermi level, and we may write

$$[V^0] = [V] - [V^-] \qquad (2)$$

Equations 1 and 2 give the relation

$$[V^0] = \frac{[V]}{1 + \exp(F - E_V/kT)} \qquad (3)$$

Equation 3 thus gives the concentration of neutral vacancies both in a doped and in an intrinsic sample, and it follows that

$$\frac{[V]_D}{[V]_I} = \frac{1 + \exp(F_D - E_V/kT)}{1 + \exp(F_I - E_V/kT)} \qquad (4)$$

where the subscripts D and I refer to the doped and intrinsic samples respectively. The conduction electron densities in the two samples may be expressed thus:

$$\frac{n_D}{n_I} = \exp\left(\frac{F_D - F_I}{kT}\right) \qquad (5)$$

and, combining equations 4 and 5

$$\frac{[V]_D}{[V]_I} = \frac{1 + (n_D/n_I)\exp(F_I - E_V/kT)}{1 + \exp(F_I - E_V/kT)} = \frac{D_D}{D_I} \qquad (6)$$

if the diffusion coefficient is directly proportional to the vacancy concentration. If the vacancy acceptor level lies near the valence band, $F_I - E_V \gg kT$, and

$$D_D/D_I = n_D/n_I \qquad (7)$$

If, however, the level lies within kT of the middle of the band gap,

$$\frac{D_D}{D_I} = \tfrac{1}{2}(1 + n_D/n_I) \qquad (8)$$

As the vacancy level approaches the conduction band the diffusion coefficient becomes insensitive to doping level.

Since n_D/n_I is always greater than, or equal to, unity, equation 7 represents the maximum possible change in diffusion coefficient due to this effect, and a value of D_D/D_I intermediate between this maximum value and unity will result if the vacancy level lies at some position in the forbidden gap away from the valence band. Measurement of D_D/D_I should, in principle therefore, help to locate the position of the vacancy acceptor level in the forbidden gap.

2.1.2. *p-Type Material – Non-Degenerate Statistics*

Millea's method of applying his analysis to p-type material contains errors, particularly his equation 4 for the electron concentration in p-type material which should read

$$n_p = \frac{2n_I^2}{N_a}\left[1 + (1 + 4n_I^2/N_a^2)^{\frac{1}{2}}\right]^{-1}$$

and it is suggested that the following treatment, involving hole concentrations, is more convenient.

In p-type material where the acceptor concentration N_a is of the same order of magnitude as the intrinsic concentration, n_I, we may write

$$p - n = N_a \qquad (9)$$

and

$$np = n_I^2 \qquad (10)$$

under non-degenerate conditions.

Solution of equations 9 and 10 leads to the following expression for the hole concentration

$$p = \tfrac{1}{2}\left[N_a + (N_a^2 + 4\,n_I^2)^{\frac{1}{2}}\right] \qquad (11)$$

Considering now the dependence of the hole concentration on the Fermi level, we may write

$$p = N_V \exp[-(\Delta E + F)/kT]$$

If the hole concentration in a doped sample is p_D, and in an intrinsic crystal is p_I, we have,

$$\frac{p_D}{p_I} = \exp(F_I - F_D/kT) \qquad (12)$$

Combining equations 12, and 4 (which still holds in p-type material) we have

$$\frac{[V]_D}{[V]_I} = \frac{1 + p_I/p_D \exp(F_I - E_V/kT)}{1 + \exp(F_I - E_V/kT)} = \frac{D_D}{D_I}$$

if the diffusion coefficient is directly proportional to the vacancy concentration.

If the vacancy acceptor level lies near the valence band, $F_I - E_V \gg kT$ and

$$\frac{D_D}{D_I} = \frac{p_I}{p_D} \qquad (13)$$

Or if the level lies within kT of the middle of the band gap

$$\frac{D_D}{D_I} = \tfrac{1}{2}(1 + p_I/p_D) \qquad (14)$$

Thus, equation 11 may be used to calculate p_D in order to predict D_D/D_I by equation 13 or 14.

2.1.3. *Exact Statistics*

For the above derivations, non-degenerate approximations have been used. These approximations are not justified under conditions of very high background doping levels, and it is suggested that the following treatment be used under near-degenerate conditions.

Initially the Fermi levels both in the intrinsic and in the heavily doped material must be calculated. In intrinsic material the Fermi level may be obtained from the charge neutrality condition $n = p$, where

$$n = 4\pi\,(2m_e\mathrm{k}T/h^2)^{\frac{3}{2}}\mathrm{f}_{\frac{1}{2}}\,[F/\mathrm{k}T] \tag{15}$$

and

$$p = 4\pi\,(2m_h\mathrm{k}T/h^2)^{\frac{3}{2}}\mathrm{f}_{\frac{1}{2}}\,[-(F + \varDelta E)/\mathrm{k}T] \tag{16}$$

in which

$$\mathrm{f_n}(\zeta) = \int_0^\infty \frac{x^n\mathrm{d}x}{\exp\,[(x-\zeta)+1]}$$

These equations may be solved numerically for F, using the Fermi-Dirac functions tabulated by McDougall and Stoner [7].

In heavily doped material the relation $np = n_\mathrm{I}^2$ does not hold, as in the non-degenerate case, and the Fermi level must be obtained from the appropriate charge neutrality condition, which is

$$n - p = N_\mathrm{D}^+ \tag{17}$$

where n and p are given by equations 15 and 16 respectively, and N_D^+, the concentration of *ionised* donors, is given by the following relation

$$N_\mathrm{D}^+ = \frac{N_\mathrm{D}}{1 + 2\exp\,[F - E_\mathrm{d}/\mathrm{k}T]} \tag{18}$$

where E_d is the donor energy level. The Fermi level may be obtained, again numerically, from equation 17.

Having obtained the Fermi level in intrinsic (F_I) and in doped (F_D) material, the ratio of vacancy concentrations in the two types of material may be calculated by equation 4, *viz*

$$\frac{[V]_\mathrm{D}}{[V]_\mathrm{I}} = \frac{1 + \exp\,(F_\mathrm{D} - E_\mathrm{V}/\mathrm{k}T)}{1 + \exp\,(F_\mathrm{I} - E_\mathrm{V}/\mathrm{k}T)}$$

The ratio of diffusion coefficients can then be estimated as before, assuming various positions for the vacancy acceptor level, E_V.

2.2. Effect of Built-in Electric Field

When there is a high concentration of the diffusing impurity, an effect known as the field effect (Zaromb [8], Kurtz and Yee [9]) may occur. If the diffusion of, say, an ionised donor impurity takes place, the charged donors and electrons are diffusing simultaneously. Since the donors are much less mobile than the electrons, the latter tend to advance ahead of the donors during diffusion. This tendency, however, is checked by the requirement for electro-neutrality, and a space charge field is formed which accelerates the slower particles and retards the fast ones. The field is most pronounced when the density of diffusing impurity atoms (N) approaches the density of intrinsic carriers at the diffusion temperature. The magnitude of the electric field was given by Lehovec and Slobodsky [10] as

$$E = \frac{\mathrm{k}T}{\mathrm{q}N}\frac{\mathrm{d}N}{\mathrm{d}x}\sqrt{\frac{(N/n_\mathrm{I})^2}{(N/n_\mathrm{I})^2 + 4}}$$

Smits [11] quotes the following expression derived by Hassion and Russo.

$$D = D_\mathrm{I}\left(1 + \frac{N}{\sqrt{(2n_\mathrm{I})^2 + N^2}}\right) \tag{19}$$

i.e. if $N \gg n_\mathrm{I}$, $D \approx 2\,D_\mathrm{I}$.

This effect can therefore account, at maximum, for a change in D of a factor of two. Lehovec and Slobodsky [10] extended this treatment by calculating modified diffusion profiles from equation 19. They deduced the correction to the apparent surface concentration obtained by extrapolating from the tail of the impurity distribution to the surface, using a complementary error function i.e. ignoring the electric field enhancing diffusion.

Zaromb [8], and Klein and Beale [12] have extended this theory to the case of simultaneous diffusion of oppositely charged impurities. In this case it is no longer possible to reduce the problem to a modification of diffusion constants, because the distribution of one impurity affects the diffusion of the other. The electric field is dominated by the impurity that is in the majority, and is in such a direction as to accelerate its rate of diffusion. The other impurity is oppositely charged and its diffusion is retarded.

2.3. Diffusion Induced Defects

The lattice parameter changes accompanying concentration gradients of a solute during diffusion cause stresses which can generate dislocations and other defects. These defects may well modify the diffusion profile of the solute responsible, or of other solutes, but the magnitude of this change is not certain. Prussin [13] has derived expressions for the density distribution of dislocations, and for the total number of dislocations in the cases of (a) constant surface concentration of solute, and (b) diffusion from a finite source. Prussin found, by an etch-pit technique, that diffusion under condition (a) resulted in a penetration of dislocations only about half-way to the p-n junction, whereas under condition (b) the dislocations penetrated almost up to the junction. The qualitative features of the theoretical dislocation distribution were confirmed by experiment. Dislocations produced by phosphorus diffusion have been studied in electron microscopy by Washburn, Thomas and Queisser [14], and mechanisms of dislocation production were suggested. Levine, Washburn and Thomas [15] compared dislocation distributions produced by boron diffusion with those produced by phosphorus diffusion, and observed that, for boron diffusion, the maximum density of dislocations was located at a depth corresponding to the steepest concentration gradient; the dislocation density near the p-n junction was very small. For phosphorus diffusion, the greatest density was near the surface as also observed by Joshi and Wilhelm [16], and Jaccodine [17]. Levine, Washburn and Thomas [18] examined the dislocation array introduced by the double-diffusion process used in the production of an n-p-n transistor wafer, the diffusion conditions being a surface concentration of $2.5 \times 10^{20}/\mathrm{cm}^3$ at 1200° C for the boron diffusion, and $3 \times 10^{20}/\mathrm{cm}^3$ at 1000° C for the phosphorus diffusion. Dislocations and precipitates were observed at the emitter surface, and at the emitter-base junction, but in no case had defects penetrated to the base-collector junction.

The extent to which lattice defects, such as those produced by diffusion, are likely to affect diffusion rates is not known with certainty at present. It is well established that diffusion rates are appreciably enhanced along small angle grain boundaries in silicon and germanium (Queisser, Hubner and Shockley [19], Karstensen [20]), but there has been much conflicting data concerning the effect of plastic deformation on diffusion rates in semiconductors. Widmer [21] found small (up to 38%) enhancements of self-diffusion in germanium specimens plastically deformed prior to diffusion; the enhancements were greatest in deformed specimens which were not annealed prior to the diffusion anneal. Heldt and Hobstetter [22], however, measured diffusivities in germanium samples which had been bent and annealed, and concluded that an array of parallel dislocations in densities of between 10^6 and $10^7/\mathrm{cm}^2$ has no effect on the diffusion of antimony or indium in germanium. Subsequently Calhoun and Heldt [23] investigated the effect of plastic straining during the diffusion anneal, and found that straining had no measurable effect on the diffusion of antimony in germanium. At present, therefore, it is not possible to draw any general conclusion as to the magnitude of the effect of plastic deformation on diffusion rates.

3. The Influence of Uniform Background Doping on Diffusion Rates

3.1. Self-Diffusion

The self-diffusion coefficient in silicon was measured directly for the first time recently (Peart [24], Ghostagore [25], Fairfield and Masters [26]). Values of D_0 and Q in intrinsic material are tabulated below.

	D_0 (cm²/sec)	Q (kcal/mole)
Peart [24]	2.0×10^3	110.0
Ghostagore [25]	1.2×10^3	109.0
Fairfield and Masters [26]	9.0×10^3	118.5

In a limited series of experiments, Ghostagore [25] found a higher diffusion coefficient in 8.0×10^{19} phosphorus doped n-type material than in 9.5×10^{19} boron doped p-type material. Fairfield and Masters state that the influence of heavy doping on diffusion coefficients in their samples was less than would be expected by the vacancy mechanism described in section 2.1, assuming a vacancy acceptor level near the middle of the gap, but the qualitative trends in both these investigations are consistent with a vacancy mechanism in which vacancies act as acceptor centres. Values of $D_\mathrm{D}/D_\mathrm{I}$ obtained by Ghostagore [25] are compared with values of $\frac{1}{2}(1 + n_\mathrm{D}/n_\mathrm{I})$ or $\frac{1}{2}(1 + p_\mathrm{I}/p_\mathrm{D})$ in table I. This comparison shows that equation 8, (or 14 where appropriate) rather overestimates the effect on diffusion coefficients. However, the results lie well within the limits discussed in section 2.1. (The calculation was repeated for near-degenerate conditions, as described in section 2.1.3, and predicted values of $D_\mathrm{D}/D_\mathrm{I}$ were within $\pm 10\%$ of the value given by the non-degenerate treatment.)

3.2. Diffusion of Impurities

The effect of background doping level has been studied for phosphorus (Tannenbaum [27], Millea [6], Thurston and Tsai [28], Mackintosh

TABLE I Effect of heavy doping on self-diffusion in silicon.

Author	Temperature (° K)	Background doping type	n_D	D(cm²/sec)	$\frac{1}{2}(1 + n_D/n_I)$	D_D/D_I
Ghostagore [25]	1451	N	8.6×10^{19}	7.3×10^{-14}	2.3	1.6
	1451	I	3.1×10^{13}	4.6×10^{-14}	—	—
			p_D		$\frac{1}{2}(1 + p_I/p_D)$	
	1451	P	1.0×10^{20}	3.3×10^{-14}	0.62	0.72

[29]), antimony (Millea [6], Thurston and Tsai [28]), tin (Millea [6]), boron (Williams [30], Thurston and Tsai [28]), and indium (Millea [6]).

3.2.1. Donor Impurities

On the theory described in section 2.1, movement of the Fermi level from valence band to conduction band should have the effect of increasing the diffusion coefficient. In the case of the donor impurities (table II) it can be seen that experimental results are not wholly in line with this prediction. For diffusion of phosphorus and antimony the results of Millea [6] and Tannenbaum [27] show the trends predicted by this theory, while those of Thurston and Tsai [28], and Mackintosh [29] show opposite effects. (The experiments of Tannenbaum are not strictly comparable here, as these trends were inferred from diffusant distributions which were inconsistent with a constant diffusion coefficient, rather than by varying background doping concentrations. Specimens were originally 2Ω-cm boron doped material.) However, the changes in diffusion constant observed by Thurston and Tsai, and Mackintosh, could not have been due to Fermi level effects since the doping level of their "doped" samples was appreciably lower than the intrinsic carrier concentration at the diffusion temperature. Furthermore, these changes cannot be explained by the effect of built-in electric field, since surface concentrations were kept constant in these experiments (phosphorus surface concentration was about 3×10^{20}/cm³ in Thurston and Tsai's experiments). These authors, [28] and [29], suggested that the effects were due to deviations from complementary error function distributions of diffusant;

TABLE II Effect of uniform doping on diffusion coefficient of donors.

Diffusing species	Temperature (° K)	Background doping type	n_D or p_D	D(cm²/sec)	$\frac{1}{2}(1 + n_D/n_I)$ or $\frac{1}{2}(1 + p_I/p_D)$	D_D/D_I	Author
P^{32}	1427	I	—	8.7×10^{-13}	—	—	Millea [6]
		N	5.1×10^{19}	1.2×10^{-12}	1.8	1.4	
		P	2.2×10^{20}	3.7×10^{-13}	0.54	0.4	
P	1323	I	2.0×10^{19}*	5.0×10^{-13}			Tannenbaum [27]
		P	6.0×10^{20}*	1.0×10^{-11}	15.5	20.0	
P	1523	I	1.0×10^{15}	4.5×10^{-12}	—	—	Thurston and Tsai [28]
		P	8.0×10^{18}	6.5×10^{-12}	0.941	1.44	
P	1473	I	5.0×10^{14}	7.5×10^{-12}	—	—	Mackintosh [29]
		P	1.0×10^{17}	1.0×10^{-11}	1.00	1.33	
Sb^{124}	1465	I	—	2.8×10^{-13}			Millea [6]
		N	7.1×10^{19}	7.1×10^{-13}	2.0	2.5	
		P	2.4×10^{20}	8.2×10^{-14}	0.55	0.29	
Sb	1473	I	1.0×10^{14}	3.0×10^{-13}	—	—	Thurston and Tsai [28]
		P	1.0×10^{18}	5.0×10^{-13}	0·992	1.66	
Sn^{113}	1488	I		1.2×10^{-13}			Millea [6]
		N	8.8×10^{19}	2.5×10^{-13}	2.0	2.1	
		P	8.4×10^{20}	1.7×10^{-13}	0.534	1.4	

*These values refer to *diffusant* concentration, not background doping concentration.

such deviations would not be detected in the p-n junction technique.

In table II, values of $\frac{1}{2}(1+n_D/n_I)$ or $\frac{1}{2}(1+p_I/p_D)$ are compared with D_D/D_I, showing fair agreement with equation 8 or 14 for the results of Millea and Tannenbaum. (In all cases the value calculated for near-degenerate conditions was within $\pm 10\%$ of that given by the non-degenerate treatment.) This indicates that phosphorus and antimony probably diffuse by a vacancy mechanism in which vacancies act as acceptors with a level near the centre of the band gap.

Data for diffusion of tin is also given in table II, and here the measurements of Millea are not in agreement with values predicted by equations 8 and 14. To explain this result, Millea proposed a simple interchange mechanism, similar to that proposed for indium as will be described below.

3.2.2. *Acceptor Impurities*

Diffusion data for the acceptor impurities are given in table III. In this case the doping levels in the samples of Thurston and Tsai [28] are comparable with the intrinsic concentration at the diffusion temperature, and one might expect Fermi level effects to be important. The values of D_D/D_I for boron as measured both by Thurston and Tsai [28], and Williams [30], agree fairly well with the vacancy model as proposed above for antimony and phosphorus and equation 8 describes the values reasonably well. The values for indium, however, show opposite trends to those of boron, since, while boron diffuses faster in n-type material, indium diffuses faster in p-type material. The behaviour of indium can be accounted for by an interchange mechanism proposed by Millea. In this mechanism a neutral indium atom moves into an interstitial site as In^+, leaving a negatively charged vacancy. A silicon atom then moves into the vacancy, and the indium ion moves into the place from which the silicon atom came. In this way the diffusion rate does not depend on the concentration of thermal vacancies, but rather on the concentration of *neutral* indium atoms, since a higher activation energy would be required if the indium was negatively charged. The diffusion coefficient, in this mechanism, will thus be larger in p-type than in n-type material. The predicted dependence is $D_D/D_I = n_I/n_D$ on this model, which is in good agreement with the results obtained by Millea. Another possible mechanism, interstitial diffusion of indium, does not agree so well with the experimental results.

At present there has been no consideration of the effect of the concentration of the diffusing impurity on the Fermi level, and hence on its own diffusion rate. In the experiments of Millea [6] no values of surface concentration are given, and thus it is not known whether this effect would be important. In the case of Tannenbaum's and Williams' investigations, however, one might expect the diffusion profiles to be modified by this effect. If, however, the surface concentration was kept constant throughout the experiments on specimens of different background doping, both this effect and the built-in electric field effect should be minimised.

4. Diffusion in Non-Uniformly Doped Material: the "Push-Out" Effect

L. E. Miller [31] first observed abnormal diffusion of base dopant impurity near an emitter region of a double-diffused n-p-n semiconductor device structure. This effect was observed as an enhanced penetration of the base region beneath a diffused emitter, the base dopant being gallium, and the emitter dopant phosphorus. A typical example of this phenomenon, revealed by

TABLE III Effect of uniform doping on diffusion coefficient of acceptors.

Diffusing species	Temperature (° K)	Background doping type	n_D or p_D	D(cm²/sec)	$\frac{1}{2}(1+n_D/n_I)$ or $\frac{1}{2}(1+p_I/p_D)$	D_D/D_I	Author
B	1523	I	2.0×10^{13}	5.6×10^{-12}	—	—	Williams [30]
		N	9.0×10^{19}	11.8×10^{-12}	2.0	2.1	
B	1508	I	1.0×10^{14}	5.0×10^{-12}	—	—	Thurston and Tsai [28]
		N	7.0×10^{19}	9.0×10^{-12}	1.68	1.85	
In^{114}	1519	I	—	2.4×10^{-12}			Millea [6]
		N	6.0×10^{19}	1.8×10^{-12}	1.5	0.75	
		P	2.1×10^{20}	1.7×10^{-11}	0.57	7.1	

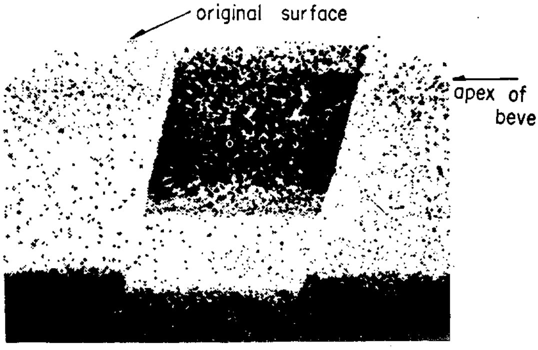

Figure 1a Photomicrograph of a bevelled and copper-plated section revealing collector junction structure in the region of restricted emitter. (By courtesy of the *Japanese Journal of Applied Physics*)

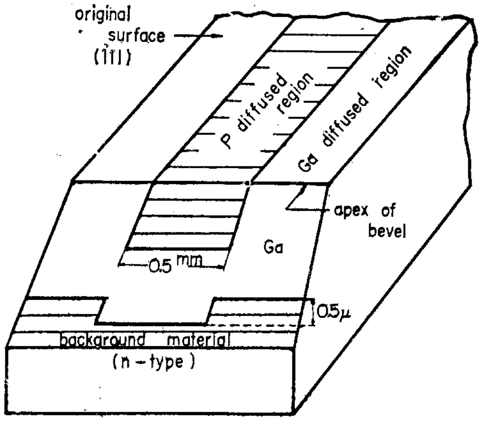

Figure 1b Schematic representation of fig. 1a.

staining, is shown in fig. 1 (Sato and Arata [32]). Miller suggested that this might be explained in terms of the change in Fermi level by the phosphorus diffusant which alters the diffusion coefficient of the gallium. Baruch *et al* [33] confirmed this effect and also suggested that an increase in vacancy concentration in the emitter was responsible, since they observed similar effects after irradiating with high-energy protons. They calculated that the change in Fermi level in the emitter would raise the vacancy concentration there by an order of magnitude. These excess vacancies would diffuse towards the collector-base junction during cooling, and increase the diffusion coefficient of the gallium base impurity.

Moore [34] suggested that strain from the high phosphorus concentration results in an increased vacancy concentration below the emitter, but he did not suggest whether plastic or elastic strain was responsible. Moore also mentioned that this effect was not observed in p-n-p structures where boron emitters are used.

Sato and Arata [32] investigated the distribution of dislocations near the junction formed by diffusion of phosphorus, and explained the push-out effect in terms of enhanced diffusion due to these dislocations. They found that the etch-pit density at the position where the collector-base junction would be, was about 10^7/cm^2. There is, however, considerable doubt as to whether such a dislocation array would have any measurable effect on the diffusion rates of substitutional impurities, as discussed in section 2.3, and, even where significant effects have been observed (Widmer [21]), these are much too small to explain the observed push-out unless the dislocation density was grossly underestimated by the etch-pit measurements of Sato and Arata.

Gereth, van Loon and Williams [35] made a comprehensive study of the emitter-dip effect. This effect was observed in n-p-n, but not in p-n-p structures, and was only observable when the phosphorus surface concentration was greater than the intrinsic carrier concentration at the diffusion temperature. Above this level, variation of phosphorus concentration had no effect on the magnitude of the dip. The size of the effect did not depend on the diffusion time used to form the emitter, but successive emitter diffusions (between which the specimen was removed from the furnace) caused successive increases in the emitter-dip. The reason appeared to be that the dip was created during the time the slice was being removed from the furnace rather than during the emitter diffusion, since it was found that the dip was enhanced by slow cooling and reduced by quenching. The base doping concentration was varied by using epitaxially deposited base layers of different doping concentrations, and it was found that a dip occurred only at boron concentrations greater than about 10^{18}/cm^3 (at 1000° C). Even if the epitaxial base layers of low boron doping were given, by diffusion, a high surface concentration of boron, the subsequent emitter diffusion did not cause a dip. Thus, it was concluded that this effect is not the result of rejection of boron by the region of high phosphorus concentration, but depends rather on the base concentration in the region of the base-collector junction. It was verified that the dip effect was not caused by the oxide mask, since no dip was observed if a slice having open windows was heated in the absence of phosphorus. The concentration profiles of base layers, before and after emitter diffusion, were measured using the technique of anodic sectioning (Tannenbaum [27]) and four-point probe resistivity measurements. These studies indicated that there is no pile-up of boron impurities in the base region directly in front of the emitter, and the distributions after emitter diffusion could be explained by an increased diffusion coefficient for boron near the collector junction. This increase must amount to a factor of 30 if the effect occurs during emitter diffusion, or, if it occurs during cooling (as indicated by these experiments) the increase must be at least a factor of 500. Gereth, van Loon and Williams did not suggest a model to explain all their results, but noted that their observations were all in qualitative agreement with Baruch's model. A theoretical objection raised to Baruch's model by Shockley is that the excess charged vacancies in the emitter will prefer to diffuse towards the surface rather than to the base, since the latter process becomes less favourable with decreasing temperature. However, if the effect occurs during emitter diffusion, it is clear that the excess vacancies would diffuse down their own concentration gradient, towards the base.

Nicholas [36] in a study of the emitter-dip effect obtained concentration profiles which were very similar to those of Gereth *et al* [35] and he also concluded that this effect is not due to a pile up of boron in front of the phosphorus, but rather to an increase in the diffusion coefficient of boron. Nicholas suggests that the emitter-dip effect is due to strain relief, which enhances vacancy concentration, presumably by dislocation interactions etc. This suggestion follows an experiment which showed that the presence of wet oxidation during diffusion caused a large increase in diffusion rate. Nicholas suggested that this increase was due to vacancies produced by strain relief (Queisser and van Loon [37]). Also, externally applied strain increased the diffusion coefficient under oxidising conditions; the diffusion coefficient was roughly proportional to the final dislocation density ($\sim 10^4/\mathrm{cm}^2$), being about a factor of 2 to 3 times the normal value, in the highly strained region. These experiments, however, do not lend themselves to a clear interpretation. The enhancements of diffusion coefficient are insufficient to explain the magnitude of the emitter-dip effect, and the presence of oxygen is a complicating factor.

Gereth and Schwuttke [38] present some rather unconvincing X-ray topograph evidence for precipitates in the phosphorus-diffused emitter region of an n-p-n structure. The authors suggest, by analogy with the electron microscope observations of Schmidt and Stickler [39] that these are SiP precipitates. It was suggested that these precipitates form during cooling, and excess neutral vacancies are generated at the sites of the precipitates. These vacancies then diffuse towards the base-collector junction where they enhance the base dopant diffusion constant and cause the emitter-dip effect.

Lawrence [40] observed the emitter-dip effect both in n^+-p-n and in p^+-p-n structures, a result which is not in agreement with the theory involving vacancy formation due to the Fermi level in the emitter. When n^+-p-n and p^+-p-n structures were diffused under similar conditions of concentration, it was found that the boron p^+ region had to diffuse to within 1 μm of the base penetration front for enhanced diffusion to occur, while the phosphorus n^+ region initiated enhanced base penetration when the junctions were separated by 2 μm. Lawrence associates this with the fact that the concentration of phosphorus was greater than that of boron, and hence phosphorus emitters should produce greater misfit stresses.

Lawrence observed dislocations in the base region of samples showing the emitter-dip effect, but no idea of overall dislocation density is given. The electron micrographs indicate that densities of 10^6 to $10^7/\mathrm{cm}^2$ were present in this region.

The effect of base concentration in p-n-p structures was investigated with boron-emitter concentrations consistently above $Q = 4 \times 10^{16}$ atoms/cm². Above a phosphorus base concentration of $Q = 4 \times 10^{15}/\mathrm{cm}^2$ the usual emitter-dip effect was observed, but below this concentration the base-collector junction exhibited a *retarded* penetration. No similar effect could be produced in n-p-n structures.

To test whether these effects were due to stress, a direct stress was applied via a silicon pin to p-n and n-p structures at 1200° C. In the case of boron-diffused structures, the p-n junction where the load was applied advanced away from the pin. In the case of phosphorus-diffused structures a retarded penetration (towards the pin) was observed. In both cases the abnormal diffusion was found to occur during the period when maximum plastic deformation occurred. On the basis of these experiments, Lawrence attributed the enhanced diffusion of base impurities to excess vacancies in the base region generated by dislocations moving away from the emitter. The retarded penetration of a base region was attributed to base impurities precipitating at the lattice disorder formed to relieve the emitter impurity stress.

Although the type of stress used in the direct stress experiment is expected to be rather different from that produced in a diffused emitter, this experiment demonstrated well that a direct stress can cause enhanced and retarded penetration. However, this theory does not explain all observations, particularly those of Gereth, van Loon and Williams [35] which indicated that the emitter-dip effect occurred during cooling of the sample.

5. Conclusions

A complete explanation of the emitter-dip, or push-out effect is still not available, for the theories put forward hitherto do not explain all the observations recorded. An important limitation to the understanding of this effect is that the possible mechanisms, described in section 2, have not been investigated fully enough themselves. Experimental evidence on the effect of heavy uniform doping on diffusion rates is very limited, and, as shown in section 3.2, even the qualitative effects are not understood in the case of certain elements. Furthermore, the range of elements studied is very limited and more fruitful experiments might be made on elements whose misfit in the silicon lattice was small, in order to minimise the effect of diffusion-induced defects. In the case of diffusion-induced defects also, the experimental evidence is very limited. This deficiency is emphasised by some of the more recent theories of the emitter-dip effect, based on defects produced by misfit stresses during the emitter diffusion. These theories rely on evidence such as that of Sato and Arata [32] that suggest that damage created by the emitter diffusion should extend down to the collector-base junction, but, as discussed in section 2.3, there is a considerable volume of evidence that diffusion-induced defects do not extend to this level. Even if plastic deformation does extend to the base-collector junction, the magnitude of the change in base diffusion coefficient produced cannot be predicted with any degree of certainty for, as discussed in section 2.3, there is no agreement on the magnitude of the effect of plastic deformation on diffusion rates. Furthermore the kind of stresses imposed by diffusion may have quite different effects from those imposed by macroscopic applied stresses. Finally, the possibility that precipitation plays an important part in the emitter-dip effect must not be ruled out, and it seems clear that another area of research which could throw light on this problem is that concerning the conditions under which precipitation occurs during diffusion.

Acknowledgements

The author is indebted to Professor R. L. Bell for help and advice concerning this review. This work was done under a CVD contract and is published by permission of the Ministry of Defence (Navy Department); financial support was also provided by an ICI Research Fellowship, and thanks are extended to each of these bodies. The author also acknowledges useful discussions with Dr J. B. Mullin and Mr D. Baker. The paper is published by permission of the Controller, HM Stationery Office.

References

1. H. REISS and C. S. FULLER, "Semiconductors", edited by N. B. Hannay (Reinhold, New York, 1959) p. 222.
2. B. I. BOLTAKS, "Diffusion in Semiconductors" (Infosearch, London, 1963).
3. R. M. BURGER and R. P. DONOVAN, "Fundamentals of Silicon Integrated Device Technology" Vol. 1 (Prentice Hall, New Jersey, 1967).
4. R. L. LONGINI and R. F. GREENE, *Phys. Rev.* **102** (1956) 992.
5. M. W. VALENTA and C. RAMASASTRY, *ibid* **106** (1957) 73.
6. M. F. MILLEA, *J. Phys. Chem. Solids* **27** (1966) 315.
7. J. MCDOUGALL and E. C. STONER, *Phil. Trans. A* **237** (1938) 67.
8. S. ZAROMB, *IBM J. Res. Dev.* **1** (1957) 57.
9. A. D. KURTZ and R. YEE, *J. Appl. Phys.* **31** (1960) 303.
10. K. LEHOVEC and A. SLOBODSKY, *Solid State Electronics* **3** (1961) 45.
11. F. M. SMITS, *Proc. IRE* **46** (1959) 1049.
12. T. KLEIN and J. R. A. BEALE, *Solid State Electronics* **9** (1966) 59.
13. S. PRUSSIN, *J. Appl. Phys.* **32** (1961) 1876.
14. J. WASHBURN, G. THOMAS, and H. J. QUEISSER, *ibid* **35** (1964) 1909.
15. E. LEVINE, J. WASHBURN, and G. THOMAS, *ibid* **38** (1967) 81.
16. M. L. JOSHI and F. WILHELM, *J. Electrochem. Soc.* **112** (1965) 185.

17. R. J. JACCODINE, *Appl. Phys. Letters* **4** (1964) 114.
18. E. LEVINE, J. WASHBURN, and G. THOMAS, *J. Appl. Phys.* **38** (1967) 87.
19. H. J. QUEISSER, K. HUBNER, and W. SHOCKLEY, *Phys. Rev.* **123** (1961) 1245.
20. F. KARSTENSEN, *J. Electronics and Control* **3** (1957) 305.
21. H. WIDMER, *Phys. Rev.* **125** (1962) 30.
22. L. A. HELDT and J. N. HOBSTETTER, *Acta. Met.* **11** (1963) 1165.
23. C. D. CALHOUN and L. A. HELDT, *ibid* **13** (1965) 932.
24. R. F. PEART, *Phys. Stat. Solidi* **15** (1966) K119.
25. R. N. GHOSTAGORE, *Phys. Rev. Letters* **16** (1966) 890.
26. J. M. FAIRFIELD and B. J. MASTERS, *Bull. Amer. Phys. Soc.* July 1966, and *Appl. Phys. Letters* **8** (1966) 280.
27. E. TANNENBAUM, *Solid State Electronics* **2** (1961) 123.
28. M. O. THURSTON and J. TSAI, Ohio State University Foundation Report No. 1233-49 (1962).
29. I. M. MACKINTOSH, *J. Electrochem. Soc.* **109** (1962) 392.
30. E. L. WILLIAMS, *ibid* **108** (1961) 795.
31. L. E. MILLER, "Properties of Elemental and Compound Semiconductors", edited by H. Gatos (Interscience, New York, 1960).
32. Y. SATO and H. ARATA, *J. Appl. Phys. Japan* **3** (1964) 511.
33. P. BARUCH, C. CONSTANTIN, J. C. PFISTER, and R. SAINTSPRIT, *Discuss. Faraday Soc.* **31** (1961) 76.
34. G. E. MOORE, "Microelectronics", edited by E. Keonjian (McGraw-Hill, New York, 1963) p. 262.
35. R. GERETH, P. G. G. VAN LOON, and W. WILLIAMS, *J. Electrochem. Soc.* **112** (1965) 323.
36. K. H. NICHOLAS, *Solid State Electronics* **9** (1966) 35.
37. H. J. QUEISSER and P. G. G. VAN LOON, *J. Appl. Phys.* **35** (1964) 3066.
38. R. GERETH and G. H. SCHWUTTKE, *Appl. Phys. Letters* **8** (1966) 55.
39. P. SCHMIDT and R. STICKLER, *J. Electrochem. Soc.* **111** (1964) 1188.
40. J. E. LAWRENCE, *J. Appl. Phys.* **37** (1966) 4106.

Letter

A Method for Determining the Thermal and Athermal Components of Flow Stress from Stress-Relaxation

It is now generally accepted that plastic deformation of metals can be thermally activated if strain rate is determined by the overcoming of short-range obstacles by dislocations [1, 2]. This behaviour can be satisfactorily accounted for by an Arrhenius type of equation

$$\dot{\gamma} = \dot{\gamma}_0 \exp\left[-\frac{H(\tau^*)}{kT}\right] \qquad (1)$$

where $\dot{\gamma}$ is the shear strain rate; $\dot{\gamma}_0$ contains geometrical factors and the frequency with which a dislocation pressed against an obstacle attempts to overcome it; H is the activation energy; k and T have their usual significance.

It was first suggested by Seeger [3] that the shear stress for yield or flow of a metal crystal, τ, can be considered to consist of two components; one, an athermal component τ_μ, depends on temperature only through the shear modulus μ, and the other, the thermal component τ^*, depends on temperature T and strain rate $\dot{\gamma}$;

$$\tau = \tau^*(T, \dot{\gamma}) + \tau_\mu \qquad (2)$$

The thermodynamic variables associated with thermal activation can be determined by evaluating τ^* at various temperatures and also by determining deformation partials such as $(\partial\tau^*/\partial T)\dot{\gamma}$ and $(\partial\tau^*/\partial \ln\dot{\gamma})_T$ [1, 2]. By evaluating these variables, it is then possible to speculate on the operating short-range hardening mechanism.

The starting point of any such analysis is the separation of τ^* from the total flow stress τ. This is usually accomplished by measuring τ at various temperatures up to a sufficiently high temperature T_0 above which all short-range obstacles are transparent to dislocations (i.e. $\tau^* = 0$). The total applied stress is then applied solely to overcoming the long-range stress-field; τ^* for other temperatures can then be determined by knowing the shear modulus variations with temperature

$$\tau^*_T = \tau_T - \frac{\tau_\mu(T_0).\mu_T}{\mu_{T0}} \approx \tau_T - \tau_{T0} \qquad (3)$$