

Nonequilibrium point defects and diffusion in silicon

S.M. Hu*

IBM Semiconductor Research and Development Center, East Fishkill Facility, Hopewell Junction, NY 12533 (USA)

Received 23 September 1993; accepted 3 January 1994

Abstract

Many surface and bulk processes generate or consume point defects in crystalline silicon. Some such processes that commonly occur in silicon device fabrications include impurity/dopant diffusion, thermal oxidation, thermal nitridation, silicidation, plasma treatment, ion implantation and oxygen precipitation. These processes can cause the concentrations of point defects to depart from their thermal equilibrium values. Nonequilibrium point defects can profoundly affect dopant diffusion. They may also engender extended defects, such as stacking faults and dislocations. Understanding of the causes and effects of nonequilibrium point defects is a prerequisite for the prediction and control of dopant diffusion in the fabrication of modern submicron devices. This report reviews technologically important processes that cause nonequilibrium point defects in silicon, and our current understanding of them.

1. Introduction

At sufficiently high temperatures, all crystalline solids contain thermally generated point defects. Two elementary native point defects are the vacancy and the self-interstitial. A vacancy is a lattice site with a missing atom. A self-interstitial is a host atom located anywhere in a crystal except at a lattice site, although for energetic reasons there are only a limited number of such off-lattice locations at which the interstitial atom may reside. Point defects affect many fundamental as well as technologically important phenomena in crystalline solids. Of their many roles, the best known is that of acting as vehicles of atomic transport processes, such as diffusion and certain activated processes of solid deformation. The equilibrium concentrations of these point defects are determined by their enthalpies and entropies of formation, and are thermodynamically defined functions of temperature, stress and, in semiconductors such as silicon, electron concentration.

When this author wrote a review on diffusion in silicon and germanium in 1973 [1], little was known about point defects in nonequilibrium. The peculiar shape of phosphorus diffusion profiles was characterized as anomalous. The intriguing phenomenon of surface orientation dependence of diffusion of boron and phosphorus was discussed in terms of speculations instead of analyses. Dislocations were a favorite culprit for these anomalous phenomena, including the so-called "emitter push" effect (Section 9.1). Then, in 1974, it was established [2] that excess silicon self-interstitials are generated during the thermal oxidation of silicon, and that they affect diffusion disparately for different substitutional impurities. This soon triggered a deluge of research activities, leading to the discovery of many different surface and bulk processes that give rise to nonequilibrium point defects. Many of these processes occur commonly in silicon device processing: thermal oxidation,

*Current address: 25 Bykenhulle Road, Hopewell Junction, NY 12533, USA.

nitridation, oxygen precipitation, silicidation, plasma treatment and ion implantation. And, most significantly, via a mechanism called “chemical pump” (see Section 9), the very diffusion process itself can cause point defects to become far from equilibrium.

When point defects exceed their thermal equilibrium concentrations they may condense into various types of extended defects, such as point defect clusters, stacking faults and dislocations, of which all are harmful to semiconductor devices. The occurrence of non-equilibrium point defects also affects impurity diffusion in very complicated ways, making it very difficult to predict resulting impurity profiles and device structures after a sequence of processes. With the smallness, tightness and complexity of modern silicon device structures, the prediction and control of impurity diffusion are of paramount importance. To achieve these goals requires a good understanding of the processes and mechanisms that give rise to nonequilibrium point defects. In this report, we shall review and discuss technologically important processes that cause nonequilibrium point defects in silicon and our current understanding of them. Some of these, as well as some not treated here, have been discussed extensively in a recent review by Fahey et al. [3]; this is highly recommended as a complement to the present paper.

2. Intrinsic point defects in silicon

2.1. Equilibrium intrinsic thermal point defects

In a finite crystal with a significant surface to volume ratio, vacancies and interstitials are generated independently of each other by the Schottky process. A vacancy is generated, in terms of net result, by moving a lattice atom in the bulk to the surface and attaching it to a kink of a surface step so as to conserve the surface area and kink density, and hence the surface free energy of the crystal, while the volume of the crystal has increased by one atomic volume. (The original picture of the Schottky process simply places an atom from the bulk on the crystal surface plane [4], thus involving an additional surface energy for its creation.) Similarly, the generation of self-interstitials by the Schottky process can be pictured as taking a surface atom at a kink of a surface ledge to the bulk of the silicon lattice. Let us consider the simple, neutral point defects. (Charged point defects are discussed elsewhere [5–8].) We can write the free energy change with the formation of N_V vacancies in a lattice of N_L lattice sites as

$$\Delta G_V = N_V(-\Delta S_V^f T + \Delta H_V^f) - kT \ln \frac{N_L!}{N_V!(N_L - N_V)!} \quad (1)$$

We can write a similar expression for ΔG_I^f . Since the vacancies and the self-interstitials can be generated independently of each other in the Schottky process, the minimum of the free energy sum $\Delta G_V + \Delta G_I$ is obtained by minimizing with respect to N_V and with respect to N_I independently. The two differentiations give the equilibrium concentrations of the vacancies and of the self-interstitials respectively as

$$C_V^* = C_L \exp\left(\frac{\Delta S_V^f}{k}\right) \exp\left(\frac{-\Delta H_V^f}{kT}\right) \quad (2)$$

and

$$C_I^* = C_L \exp\left(\frac{\Delta S_I^f}{k}\right) \exp\left(\frac{-\Delta H_I^f}{kT}\right) \quad (3)$$

On the other hand, when an infinite crystal (for which the surface is no longer relevant) is brought up to a finite temperature T , vacancies and interstitials can only be generated by a bulk process, the Frenkel process, in which equal numbers of vacancies and self-interstitials are simultaneously produced. This imposes a constraint of $C_V^* = C_I^*$, under which the minimization of free energy leads to

$$C_V^* = C_I^* = C_L \exp\left(\frac{\Delta S_V^f + \Delta S_I^f}{2k}\right) \exp\left(-\frac{\Delta H_V^f + \Delta H_I^f}{2kT}\right) \quad (4)$$

By use of Schwarz's inequality, it can be shown that the free energy reduction in the manner of Eq. (1) with vacancy and self-interstitial concentrations given by Eq. (4) is not as much as that with concentrations given by Eqs. (2) and (3). Mathematically, minimization under a constraint (here requiring $N_V = N_I$) always gives a larger result than that obtained by minimization without constraint.

In silicon, the thermal diffusivity ($\approx 1 \text{ cm}^2 \text{ s}^{-1}$) is several or more orders of magnitude larger than any atomic or defect diffusivity. When a typical silicon wafer is inserted into a furnace, the rate of the rise of temperature in the wafer is determined mainly by radiative heat transfer to the wafer, and hence by furnace temperature as well as the wafer thickness and the arrangement of wafers relative to each other and the wafer carrier. The temperature difference across the wafer thickness will typically be very small, $< 1 \text{ }^\circ\text{C}$, at all times due to the high thermal diffusivity [9]. This means that the point defect concentrations in the bulk are more closely given by Eq. (4), while in the surface region by Eqs. (2) and (3). But, from a practical viewpoint, the surface region, no more than a few μm deep, is where all the active devices are, and where all the diffusion processes of interest take place. It is also possible that point defects of the excess species may condense into extended defects which are energetically much less costly than point defects. Thermal equilibrium may then be achieved according to Eqs. (2) and (3).

It may be noted that, individually, both C_V^* and C_I^* are each different in an infinite crystal, where the Frenkel process dominates, from in a finite crystal, where the Schottky process dominates. However, their product $C_V^* C_I^*$ is the same in both crystals, and is given by

$$C_V^* C_I^* = C_L^2 \exp\left(\frac{\Delta S_V^f + \Delta S_I^f}{k}\right) \exp\left(-\frac{\Delta H_V^f + \Delta H_I^f}{kT}\right) \quad (5)$$

In a crystal of finite size, its surface plays a key role not just in determining the equilibrium, but also in affecting the nonequilibrium concentrations of vacancies and self-interstitials, which are independently generated and annihilated at the surface by the Schottky process.

2.2. Coexistence of vacancy and self-interstitial in silicon

Diffusion in silicon has been extensively investigated since the mid-1950s. Until 1968, it had always been assumed, implicitly or explicitly, that the point defect that mediates diffusion in silicon is the vacancy. A "regular" interstitial is conceptually one which wanders through lattice interstices all by itself without displacing host or substitutional lattice atoms. But diffusion by an "interstitialcy" mechanism was already suggested by Seitz in 1950 [10]. Seitz was discussing generic diffusion without reference to specific materials. Experimental diffusion data in metals tend to indicate a preponderance of the vacancy mechanism. In

1968, Seeger and Chik [11], argued that self-diffusion in silicon at high temperatures, as well as diffusion of Groups III and V dopants, most likely proceed via an interstitialcy mechanism rather than a vacancy mechanism. But no unequivocal experimental evidence was given in support of this speculation.

In 1974, Hu [2] proposed that both vacancies and self-interstitials coexist as two basic intrinsic point defects in silicon of approximately equal importance, and that diffusion of substitutional impurities in silicon proceeds via a dual vacancy-interstitialcy mechanism. The contribution of the interstitialcy mechanism as a fraction of the total diffusivity varies with impurity species. He arrived at this conclusion by bringing together the phenomena of enhanced diffusion of certain impurities and the formation of stacking faults in the thermal oxidation of silicon, and analyzing their common features and causal relations, such as the effects of surface orientation and oxidizing ambients. A second important proposition made at the same time is that thermal oxidation is a surface process that generates excess silicon interstitials. This proposition is essential for the experimental observation of the oxidation-enhanced diffusion (OED) of certain impurities in silicon, the oxidation retarded diffusion of others (ORD), and the formation of oxidation stacking faults (OSF).

Between 1969 and 1971, several groups of investigators [12–19] reported that, in oxidizing ambients, the diffusion rate varies with the crystallographic orientation of the silicon surface, in the order $\{100\} > \{110\} > \{111\}$. What was observed is actually the phenomenon of oxidation-enhanced diffusion (OED), an effect which varies with the surface orientation. The diffusion enhancement is greater in wet oxidation than in dry oxidation. A separate phenomenon had been reported several years earlier, that thermal oxidation of silicon often caused the formation of stacking faults (OSF, the oxidation stacking fault) [20–24]. The growth rate of OSF is also dependent on the crystallographic orientation of the surface being oxidized, in the order $\{100\} > \{110\} > \{111\}$. The growth rate is larger for oxidation in wet oxygen than in dry oxygen. These similarities between the two phenomena, OED and OSF, led Hu to suggest [2] that these two phenomena are closely related, and have a common origin, which he proposed to be the injection, and the consequent supersaturation, of self-interstitials by the thermal oxidation of silicon.

Connecting OED to OSF brought us to the crucial point on our way to understanding diffusion in silicon. This is because the nature of the stacking fault can be determined from image contrast in transmission electron microscopy (TEM) [25,26]. Through this means, the OSF in silicon has been determined to be extrinsic in nature, namely of the interstitial type [22,23].

There is another phenomenon of OSF growth that can allow only the logical conclusion that the OSF are interstitial disks. This phenomenon is associated with the precipitation of oxygen in silicon. When interstitially dissolved oxygen atoms agglomerate into an SiO_2 precipitate, a unit volume of silicon is converted to approximately 2.25 unit volumes of the precipitate, incurring an extremely large strain energy. One way to reduce the strain energy incurred by the localized volume expansion is to provide the needed cavity volume via the emission of self-interstitials and/or the absorption of lattice vacancies. This means there will be a supersaturation of self-interstitials, or/and an undersaturation of vacancies. The supersaturation of self-interstitials due to oxygen precipitation has been shown by Hu [27] and Rogers et al. [28,29] to feed the growth of pre-existing OSF.

The model of point defects and diffusion put forth by Hu contains the following important elements [2]:

- A. The vacancy and the self-interstitial coexist as major thermal point defects in silicon. The diffusion of Groups III and V atoms in silicon occurs via a dual mechanism mediated by both the vacancy and the self-interstitial. The ratio between the interstitialcy and the vacancy components varies according to the atomic species.
- B. Thermal oxidation of silicon injects excess silicon interstitials into the silicon substrate. The rate of injection is assumed to be directly proportional to the rate of thermal oxidation. The excess self-interstitials are assumed to be annihilated at the surface, via a regrowth process which is dependent on the surface orientation, and especially the density of surface kinks.
- C. Some of the self-interstitials in excess of the thermal equilibrium value would condense to form OSF, when nucleation sites are available.
- D. At the same time, the supersaturation of self-interstitials would enhance the diffusion of those substitutional dopants which exhibit an affinity with the self-interstitial. The difference in OED among various dopants then reflects the difference in the dopant's affinity with the self-interstitial. At that time, observations of OED were available only for boron, phosphorus and arsenic; but the difference in OED among these was already sufficient to establish the dual mechanism as the only logical explanation possible.

In the years that followed, a number of related phenomena have been discovered, all of which have greatly reinforced the concept of the coexistence of the vacancy and the self-interstitial as the two main intrinsic point defects. These include the observation, by Mizuo and Higuchi [30] and Antoniadis and Moskowitz [31], of the oxidation-retarded diffusion (ORD) of antimony, for which there is a very natural explanation: a supersaturation of self-interstitials causes an undersaturation of vacancies; if the antimony diffusion in silicon is mediated predominantly by the vacancy, it will be retarded during oxidation. There is also a physically intuitive interpretation for antimony's preference for the vacancy mechanism: the elastic interaction between a point defect and a substitutional atom is dominated by the mismatch of the sizes of the lattice and the substitutional atoms. A larger substitutional atom is attracted to a vacancy, while a smaller substitutional atom is attracted to a self-interstitial.

A complementary phenomenon was subsequently discovered [32–36], that thermal nitridation of silicon injects excess vacancies. That interpretation comes from the observations of enhanced/retarded diffusion of various dopants in thermal nitridation. The diffusion of phosphorus is retarded, while the diffusion of antimony is enhanced. This trend is just the opposite of that in thermal oxidation.

With the observation of the complementary effect of nitridation, the logic of a dual point defect system is now seamless: OSF are interstitial in nature as identified by TEM fringe contrast, and corroborated by their growth during oxygen precipitation in inert ambients. Therefore, the formation of OSF indicates that there is a supersaturation of self-interstitials and an undersaturation of vacancies during thermal oxidation. From the OED of boron and phosphorus and, to a lesser extent, arsenic, we conclude that these elements can diffuse via an interstitialcy mechanism. From the ORD of antimony, we must conclude that it diffuses via a vacancy mechanism. The fact that the antimony diffusion is enhanced by thermal nitridation must then indicate that nitridation generates excess vacancies. Finally, since the phosphorus diffusion becomes retarded in the presence of these excess vacancies, one can only conclude that, under normal conditions, it takes place predominantly through the vehicle of the self-interstitial, whose concentration is reduced by annihilation with the excess vacancies during the nitridation.

2.3. Point defects and self diffusion

Random movements of vacancies and/or self-interstitials give rise to self-diffusion. Self-diffusion can be studied simply by using an isotope of the lattice host atom. The diffusion proceeds in an ideal situation: the concentrations of the point defects are at thermal equilibrium and are not affected in any way, for example, electronically; there is no interaction between an isotope atom and a point defect. The following simple relationship holds:

$$D_s C_s = \phi_v D_v C_v + \phi_i D_i C_i \quad (6)$$

where D_s , D_v and D_i are the self-, vacancy and interstitial diffusivities, and C_s , C_v and C_i are the lattice site, vacancy and interstitial concentrations, respectively. ϕ_v and ϕ_i are the vacancy and the interstitial correlation factors, respectively. These correlation factors are specifically required for the diffusion of labeled lattice atoms, such as isotopes of the host atom in self-diffusion experiments. They are not required for self-diffusion that is inclusive of any host atoms, such as occurs in the Kirkendall effect, the growth and shrinkage of stacking faults, dynamic deformation recovery etc. (Impurity diffusion is also affected by correlated random walk, but the correlation factors are complicated by the impurity-defect interaction energy, and is usually lumped into the diffusivity of the impurity-defect complex. The subject has been discussed in detail by Hu [37].) A possible contribution of Pandey's concerted exchange mechanism [38] has not been included in Eq. (6). At the present time there is no experimental data in support of this model. In most materials, usually one species, either the self-interstitial, or more often the vacancy, is overwhelmingly dominant, and is the only species that needs to be considered for all practical purposes. It is one of nature's fortuities that in silicon, the vacancy and the self-interstitial are equally important, at least in the temperature range of practical interest. Two important properties of a point defect species are its concentration and diffusivity under intrinsic equilibrium condition at a given temperature:

$$C_x^* = C_s \exp\left(\frac{S_x^f}{k}\right) \exp\left(-\frac{H_x^f}{kT}\right) \quad (7)$$

and

$$D_x^* = \frac{r^2 \nu_0}{6} \exp\left(\frac{S_x^m}{k}\right) \exp\left(-\frac{H_x^m}{kT}\right) \quad (8)$$

where the subscript X represents either I or V, S_x^f and S_x^m are the entropies of formation and migration, respectively, and H_x^f and H_x^m are the enthalpies of formation and migration. ν_0 is the normal mode frequency, and r is the atomic jump distance. The asterisk denotes parameters under intrinsic (infinite dilution of any impurity) equilibrium conditions. It follows that

$$D_s^* = \frac{r^2 \nu_0}{6} \sum_x \phi_x \exp\left(\frac{S_x^f + S_x^m}{k}\right) \exp\left(-\frac{H_x^f + H_x^m}{kT}\right) \quad (9)$$

In solids, except for highly stressed cases, the enthalpy essentially equals the internal energy, and is referred to as simply as "energy", denoted by E . If the sum $E^f + E^m$ for the vacancy is different from that for the self-interstitial, the Arrhenius plot of the self-diffusivity may exhibit a break.

Self-diffusion in silicon has been investigated experimentally over the years since the mid-1960s. Because of the very low diffusivity, earlier experiments were carried out at temperatures ≥ 1100 °C. The activation energy of self-diffusion obtained from these experiments of varied methods falls in the rather narrow range of 4.8–5.1 eV [39–41]. Different activation energies were observed in later investigations that extended to lower temperatures. While Kalinowski and Seguin [42] reported an activation energy of 4.7 eV in the temperature range 855–1175 °C, Hirvonen and Antilla [43] and Demond et al. [44] reported activation of about 4.1–4.2 eV in the low-temperature range. Furthermore, Demond et al. [44] also reported a break in the Arrhenius plot at about 1100 °C, with an activation energy of 4.9 eV in the high-temperature regime. A similar break in the Arrhenius plot for germanium diffusivity in silicon was reported by Hettich et al. [45]. Because a germanium atom in silicon is rather similar to a silicon atom (although there is a 4% difference in atomic size, and a slight difference in electronegativity), it is commonly believed that germanium diffusion in silicon should be rather comparable to self-diffusion. The transition from the low-temperature to the high-temperature activation energy is commonly ascribed to a transition from vacancy-dominated to interstitialcy-dominated diffusion, a concept first introduced by Seeger and Chik [11,46]. This implies

$$E_V^f + E_V^m \sim 4 \text{ eV} \quad (10a)$$

$$E_I^f + E_I^m \sim 5 \text{ eV} \quad (10b)$$

This set of hypothesized relations has in many cases served as constraints for those seeking to separate E_V^f , E_V^m , E_I^f and E_I^m by model fitting of diffusion profiles. In order for the interstitialcy mechanism to operate at high temperatures, it is necessary that the sum of entropies for formation and migration be much larger for the self-interstitial than for the vacancy.

It is not possible from these experiments of self-diffusion to identify the point defect that represents the dominant term on the right-hand side of Eq. (6). Only experiments of self-diffusion carried out under controlled conditions of point defect supersaturations can help resolve this issue. While so far no such experiment of silicon self-diffusion has been carried out, an indication of nearly equal vacancy and interstitialcy contributions to silicon self-diffusion may be inferred from the investigation of Fahey et al. [47] on germanium diffusion in silicon under separate conditions of vacancy and interstitial supersaturations.

2.4. Properties of point defects from experimental investigations

There is a general acceptance of the activation energy of self-diffusion of between 4 and 5 eV, apparently because of the rather good agreements among results of very different experiments. However, there are very few direct ways to separate this activation energy into its component energies of defect formation and migration, and to identify these energies with specific point defect species. The reliabilities of these few methods are also debatable. In the past decade, many investigators have attempted to obtain these parameters indirectly, by model fitting of diffusion profiles of dopants, and more particularly transition elements like gold and platinum. One cannot really say that any set of fitting parameters, including point defect formation and migration, that is uniquely suitable for one particular model represents the true physical parameters. Another attempted approach to determine defect migration kinetics separately from its thermal formation is the study of the flow of injected point defects from one surface across a silicon slice of varying thickness, and measuring

its influence on the evolution of existing stacking faults or dopant profiles on the opposite surface.

2.4.1. Studies of point defects from irradiations

In the early 1960s, EPR studies of silicon point defects were carried out, especially by Watkins [48] and co-workers, at very low temperatures in silicon samples irradiated with MeV electrons. Thermally stimulated transient capacitance (TSCAP) and deep level transient capacitance spectroscopy (DLTS) methods were introduced in the early 1970s [49,50]. These techniques began to be used for the investigation of electron-irradiated silicon from the mid-1970s [51,52]. Experimental investigations of radiation created point defects cannot produce such information as the enthalpies and entropies of formation. But they can yield such information as the activation energies of migration, the geometrical structure, and the electronic states of the point defects. One question that has not yet been resolved is whether the point defects created by energetic bombardment at low temperatures are the same point defects thermally created at high temperatures. And, assuming they are the same point defects, are their properties measured at low temperatures and in highly excited nonequilibrium states the same as at thermal equilibrium at high temperatures?

It is expected that electron bombardment would create an equal number of vacancies and self-interstitials. It was somewhat of a mystery that, in low-temperature annealing of electron-irradiated silicon, only the vacancy and its various forms of complexes have been identified, while the self-interstitial and its various forms of complexes have not. In the case of aluminum-doped silicon, an explanation for the mystery came from the observation of aluminum interstitials. [48]. In aluminum-doped silicon irradiated with 1.5 MeV electrons at 4.2 K, Watkins found a spectrum, Si-G18, which he identified as Al^{++} . The production rate of these defects was high ($\approx 0.03 \text{ defect cm}^{-3}$, per electron cm^{-2}), and was similar to that of the isolated vacancies (V^+ and V^-) [48]. Watkins speculated that silicon interstitials produced by the irradiation had replaced substitutional aluminum atoms on lattice sites. The observation of boron interstitials has been interpreted in the same way [53]. All these processes were detected at 4.2 K when the self-interstitial must be mobile in order to produce these impurity interstitials. Such fast migration may be explained by the occurrence of the Bourgoin–Corbett mechanism of athermal migration of a highly excited system under irradiation conditions. In the Bourgoin–Corbett mechanism, an interstitial alternatingly changes its charge state by capturing and emitting a charge. As already shown by Weiser in 1962 [54] (see also Hu [1]), the difference in polarization energy due to a charged interstitial may cause the hexagonal site in the silicon lattice to be at a lower energy than the tetrahedral site, which is the equilibrium site for a neutral interstitial. Thus, a neutral interstitial sitting at a tetrahedral site, upon capturing an electron (or a hole), or two, finds its current equilibrium position turned into a saddle point, and can therefore move athermally to a new equilibrium position nearby. As its charge state changes again by re-emission of an electron (or hole) the cycle repeats.

Another possibility is the absence or instability of states having an odd number of electrons, making the interstitials invisible to EPR. Anderson [55] has proposed a model to explain why defects in certain amorphous semiconductors are not paramagnetic, and hence undetectable by EPR. His explanation is that for deep level defects within a semiconductor bandgap, the repulsive coulombic energy is more than offset by spin coupling and structural relaxation that may result from capturing (or expelling) a second electron at the defect. This property is now known as Anderson's negative-U. Thus, I^- (or I^{++})

may be a lower energy state than I^- (or I^+), and V^- (or V^{++}) may be a lower energy state than V^- (or V^+).

A question that has not been raised is: have any extended interstitial defects, such as extrinsic stacking faults or the climbing of dislocations, ever been observed in Watkins' experiments? It may be noted that, while researchers using electron paramagnetic resonance see only vacancies in irradiated silicon, electron microscopists tend to see only interstitial type stacking faults and dislocation loops, for example, in electron-irradiated nickel [56,57]. This is because each characterization technique has its own sensitivity limitations. The fast migration of the excess self-interstitials would very likely allow them to agglomerate into interstitial-type dislocation loops. It is not clear whether such loops have ever been observed (or examined) by Watkins or other researchers in their EPR studies.

The silicon vacancy, while it could be frozen at 4.2 K, is also quite mobile at cryogenic temperatures, with migration energies of ≈ 0.33 and ≈ 0.18 eV in the neutral and doubly negative charge states, respectively [48,58]. The migration energy of 0.33 eV was later revised to 0.32 eV [59]. A defect which anneals out with an activation energy of 0.45 eV has not been firmly established, but is thought to be that of the doubly positive vacancy [59]. The vacancy charge states show an Anderson negative-U property, and the levels are assigned as follows [60,61]:

Vacancy donor level at $E_v + 0.05$ eV for $V^0 \rightarrow V^+ + e^-$

Vacancy donor level at $E_v + 0.13$ eV for $V^+ \rightarrow V^{++} + e^-$

Vacancy acceptor level at $E_c - 0.57$ eV for $V^0 + e^- \rightarrow V^-$

Vacancy acceptor level at $E_c - 0.11$ eV for $V^- + e^- \rightarrow V^{--}$

Because of the negative-U properties of the vacancy charge states, the V species, mentioned earlier with a migration energy of 0.33 eV, is supposed to be unstable and does not exist normally. It can however be brought into existence in, for example, a photo-excited system, in which electrons from V^- can be pumped into the conduction band.

The charge state of a point defect is important because it affects not only the mobility of the defect but also its thermal equilibrium concentration, which, of course, also depends on the Fermi level, as shown by Shockley et al. [6,7]. For example, for defect X having two acceptor states, the concentrations of its singly negative and doubly negative states are

$$C_{X^-} = C_{X^0}^0 \exp\left(\frac{E_F - E_{X^-}}{kT}\right) \quad (11)$$

$$C_{X^{--}} = C_{X^0}^0 \exp\left(\frac{2E_F - E_{X^-} - E_{X^{--}}}{kT}\right) \quad (12)$$

Doubly negatively charged defects, assumed to be vacancies, have been found to be necessary for a satisfactory simulation of arsenic diffusion profiles by Chiu and Ghosh [62]. Similarly, Fair [63] and John and Law [64] have found that the diffusivity of phosphorus increases approximately with the square of phosphorus concentration, implying a dominantly doubly negative charge state of the point defect involved.

The equilibrium concentrations of Eqs. (11) and (12) are based simply on statistical thermodynamics; according to the tenet of energy balance, the formation energy of the ionized defect is equal to that of the neutral defect plus the difference between the Fermi level and the defect state energy. But some theoretically calculated results of the formation energies of variously charged defects (see Section 2.5) appear to contradict this tenet.

2.4.2. Vacancy formation energy from positron lifetime measurements

It has long been postulated that lattice vacancies, which represent localized volumes without positive nuclear charge, tend to trap positrons, thus delaying their annihilation with electrons. Thus, the measurement of positron lifetime affords an estimate of the concentration of vacancies present. Based on positron lifetime experiments conducted in the temperature range between 300 and 1523 K, Dannefaer et al. [65] reported an enthalpy of formation of 3.6 ± 0.2 eV, and an entropy of formation of 6 to $10k$. These values seem reasonable in view of other types of experimental results. However, many assumptions which are made in the analysis of such measurements may influence the reliability of results.

2.4.3. Point defect concentrations from thermal expansion measurements

At higher temperatures, more point defects are incorporated into a crystal and, depending on whether the dominant point defect is the vacancy or the self-interstitial, will cause the volume of the crystal to expand or contract by an amount that is in addition to the normal thermal expansion of the lattice. However, as is mentioned in Section 3.2, the incorporation of "thermal point defects", which are actually fast-diffusing metallic contaminants, will also expand the volume of a given crystal. There is a discrepancy between the thermal expansion coefficient from length measurements of Okaji [66] and the thermal expansion coefficient from lattice parameter measurements of Okada and Tokumaru [67]. Okada [68] assumed that the larger thermal expansion from length measurement is a direct result of a larger $C_V - C_I$, and gave a value of $1.8 \times 10^{16} \text{ cm}^{-3}$ at 1300 K for $C_V - C_I$. However, some important points of his results are open to question. First, from the same figure (his Fig. 1) [68] from which he drew the above conclusion, one would obtain $C_V - C_I$ of about the same value at 1073 K as at 1300 K, as if the formation enthalpy of the majority defect were essentially zero. The data and interpretation are even more problematic at lower temperatures: from 500 to 700 K, $C_V - C_I$ (again from his Fig. 1) would now become negative, but its magnitude, of the order of 10^{16} cm^{-3} , is still about the same as at 1300 K. It is true that a negative value may simply indicate that the self-interstitial is the predominant point defect at low temperatures. But we should also expect the concentrations of both the vacancies and the self-interstitials in silicon at 300–500 K to be immeasurably low, not the same magnitude as at 1300 K, according to Fig. 1 of Okada [68].

2.4.4. Diffusivity of the self-interstitial from membrane experiments

Several investigations have been made using a "membrane" method [69–75]. In this method, silicon "membranes" of different thickness, tens to hundreds of μm , are prepared. Interstitials are injected from the back side of a specimen by thermal oxidation, while the front side is protected from oxidation by a nitride or nitride–oxide composite film. The injected interstitials will flow through the thickness of the membrane to the front side and get annihilated there. By measuring the evolution of pre-existing stacking faults or dopant profiles on the front side, the diffusivity of the self-interstitial can be obtained through appropriate modeling. In this way, Griffin et al. [71] have obtained the lower bound diffusivity of the self-interstitial at various temperatures. It varies from $> 10^{-8} \text{ cm}^2 \text{ s}^{-1}$ at 1000 °C to $> 10^{-7} \text{ cm}^2 \text{ s}^{-1}$ at 1200 °C. The complications come from the front side interface reactions and the bulk trapping of the excess interstitials. The bulk recombination has been neglected in most analyses. As demonstrated by Griffin et al. [74], the enhanced diffusion on the front side differs between specimens from float zone and Czochralski silicon crystals, with the enhancement reduced in the Czochralski specimen. Since Czochralski silicon contains

a high concentration ($\approx 10^{18} \text{ cm}^{-3}$) of oxygen atoms, it may be concluded that oxygen could act as traps for the injected interstitials. However, Rogers and Massoud [75] reported no difference between float zone and Czochralski silicon substrate in bulk trapping of self-interstitials migrating across the thickness of silicon wafers. The bulk generation–recombination of Frenkel pairs has also been neglected in these studies. The results obtained from these studies vary widely.

2.4.5. Defect parameters from model fitting of Au and Pt diffusion

Gold diffusion in silicon has traditionally been analyzed in the Frank–Turnbull dissociative mechanism [76]. Two essential elements for this mechanism to operate are: (1) the impurity exists in a given host lattice both interstitially and substitutionally, with the substitutional-to-interstitial equilibrium ratio much greater than unity, and (2) the interstitial species diffuses much faster than the substitutional species, so that the substitutional species may be considered to be immobile. (This does not necessarily rule out the migration of substitutional gold by a vacancy mechanism, but implies only that it plays an insignificant part.) The diffusion of gold in total concentration, which is nearly that of the substitutional species, is then effected through the motion of the interstitial species, in combination with the following reaction:



Thus, the substitutional gold Au_s migrates by first dissociating into a vacancy and a mobile interstitial gold Au_i (the reverse of the above reaction), which then moves many elementary random walk steps before converting back to Au_s (the forward of the above reaction). In 1980, Gösele et al. [77] proposed an alternative mechanism through the following reaction:



In this mechanism, an interstitial gold atom takes up a substitutional site by “kicking out” a silicon lattice atom, Si_L . This mechanism is now often referred to as the kick-out mechanism. (We may note that “replacement” is perhaps a more appropriate term, since it also states unambiguously the fact that this is a replacement reaction — a silicon lattice atom is replaced by a gold atom. The term kick-out does not simultaneously suggest that the gold atom takes up the position left by the kicked-out silicon atom. Kick-out will more appropriately denote the ejection of a lattice atom by an energetic agent, e.g. high-energy electrons or ions, photons, or electron–hole reaction events.) If, due to the very large diffusivities of these two point defects, both the vacancy and the self-interstitial concentrations are maintained near their constant equilibrium values, then there will be no difference between these two mechanisms in terms of net results. This is because, mathematically, diffusion via either mechanism can be described by the same set of reaction–diffusion equations, as follows:

$$\frac{\partial C_i}{\partial t} = D_i \frac{\partial^2 C_i}{\partial x^2} - k^f C_i + k^r C_s \quad (15)$$

$$\frac{\partial C_s}{\partial t} = k^f C_i - k^r C_s \quad (16)$$

where the subscripts i and s denote the interstitial and the substitutional species, respectively. C_i^* has been lumped into the reaction constant k^r for the kick-out mechanism, and C_v^* has been lumped into the reaction constant k^f for the dissociative mechanism. It should be noted, however, that if C_i^* and C_v^* are not constant by being different functions of C_s

and/or C_i , the two mechanisms can no longer be expressed by an identical set of equations as above. But in this case, the difference would also distinguish the normal vacancy and interstitialcy mechanisms of diffusion of a purely substitutional impurity.

If the diffusion of these point defects (vacancies from, or interstitials to, the surface of the specimen) cannot keep up with the reactions of Eqs. (13) or (14), a concentration profile of point defects is produced, which is different for the vacancy than for the self-interstitial. As shown by Gösele et al. [77] in an approximate analysis, the gold concentration profile will consequently be different for the two mechanisms. They contended that the gold profiles from the kick-out model exhibit a U-shaped distribution, in agreement with experimental findings.

Aside from the good fitting of the kick-out model to experimental gold diffusion profiles as reported by Gösele et al. [77], there seems to be no compelling physical reason why the dissociative mechanism should not take place, in view of the experimentally established coexistence of the vacancy and the self-interstitial in silicon. It seems more reasonable to expect that the dissociative mechanism would take place at least in the surface region, where vacancies can be supplied from the surface, although not necessarily at an infinite rate, while in the interior, where vacancies may be severely depleted, the kick-out provides a feasible alternative.

The kick-out model has since been utilized by a number of investigators [78–84] to obtain the silicon self-diffusivity as well as the diffusivity of the self-interstitial. In their analysis, Zimmermann et al. [82,83] have considered both the dissociative and the replacement mechanisms, as well as the bulk recombination between vacancies and self-interstitials. Their results for the temperature range 700–950 °C are summarized below:

$$C_I^* = 1.94 \times 10^{27} \exp(-3.835 \text{ eV}/kT) \text{ cm}^{-3} \quad (17a)$$

$$D_I = 2.58 \times 10^{-2} \exp(-0.965 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1} \quad (17b)$$

$$C_V^* = 1.83 \times 10^{19} \exp(-1.162 \text{ eV}/kT) \text{ cm}^{-3} \quad (17c)$$

$$D_V = 1.09 \times 10^3 \exp(-2.838 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1} \quad (17d)$$

The vacancy formation enthalpy given by Eqs. (17) seems to be unreasonably small. Furthermore, according to Eq. (2), a pre-exponential factor of 1.83×10^{19} for C_V^* would imply a negative formation entropy, something that is not possible. Assuming that only the kick-out mechanism operates, and ignoring vacancy–self-interstitial generation–recombination, Boit et al. [81] have obtained from rapid optical annealing of gold diffusion in silicon the following results:

$$D_I = 1.03 \times 10^6 \exp(-3.22 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1} \quad (18a)$$

$$C_I^* = 3.11 \times 10^{19} \exp(-1.58 \text{ eV}/kT) \text{ cm}^{-3} \quad (18b)$$

Another set of experimental data extracted from gold diffusion, reported by Bronner and Plummer [85], is

$$D_I = 600 \exp(-2.44 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1} \quad (19a)$$

$$C_I^* = 5 \times 10^{22} \exp(-2.36 \text{ eV}/kT) \text{ cm}^{-3} \quad (19b)$$

The last set of parameters has been used in a number of numerical simulations of diffusion profiles [86–88]. One can see that the activation energies of formation and migration of the self-interstitial obtained this way scatter widely among different model fitting results,

and in general disagree with theoretical calculations to be discussed later. The only agreement among them is their common use of the constraint that the sum of the formation energy and the migration energy is about 5 eV.

2.5. Properties of point defects from theoretical calculations

Theoretical calculations of the energies of formation and migration of the vacancy and the self-interstitial in silicon were made in the 1960s by use of empirical pairwise interatomic potentials [89–91]. In the last several years, several new empirical interatomic potentials have been introduced for attempting better descriptions of covalent materials, with some ad hoc terms to account for environments beyond pairwise interactions. Such empirical interatomic potentials have been employed in a number of recent calculations of the energetics of the vacancy and the self-interstitial in silicon [92–95]. Reasonable results, though not agreeing with each other, have been reported in all these investigations. But the faith in such calculations would seem unwarranted, in view of the lack of theoretical basis in such empirical potentials. All scientists know that empirical formulae cannot be stretched too far. While some such empirical potentials are acceptable for calculating energies resulting from small changes from equilibrium lattice parameters, they cannot be expected to produce reliable predictions for drastic changes involving rearrangements of valence electrons in the formation of point defects. In particular, empirical interatomic potentials cannot provide means for calculating the energetics of point defects in their different charge states, which can be quite complex and unexpected (e.g. the occurrence of the “negative-U” system mentioned in Section 2.4.1.).

In the early 1970s, it was quite popular to calculate the energetics of point defects by use of semi-empirical extended Hückel theory. The reliability of the EHT method has been criticized by Pantelides et al. [96].

Since the late 1970s, the energetics of the silicon vacancy and the self-interstitial have been obtained with first principles calculations of total energy of a crystal containing a defect. In these calculations, the total energies of a perfect crystal and of a crystal containing a point defect at different lattice locations are calculated, and the appropriate differences then give the formation energy and migration energies along different paths. The Schrödinger equation is solved for a crystal containing a defect for a self-consistent pseudopotential (including the exchange-correlation energy). The self-consistency in the pseudopotential is achieved by iterative computations, using the density-functional theory in the local density approximation (LDA) for calculating the exchange-correlation potential [97]. These calculations are accomplished via one of two schemes: the use of a self-consistent Green's function method for a finite defect cluster [98–106], and the method of supercell. The supercell replicates by translations, and each typically contains 16 or 32 nearest atoms surrounding a defect [107–111]. Up to 64 atoms are included in a supercell in a recent calculation [112] for the migration of an impurity atom from which a point defect has to move away to at least the third coordination site in order to return later to that impurity atom at another nearest neighboring site that is different from the originating site [37]. This minimum necessary cell size may still not guarantee avoidance of the occurrence of a percolation phenomenon [113] in which the point defect migration is short-circuited through closely spaced impurity atoms. The limitations on the reliability of these calculations are at present handicapped by the computationally practical size of the supercell or cluster, due to the limitations of present-day computer power. Elastic relaxation and, particularly,

electrostatic potentials are long-ranged, occurring over many atomic shells. The effect is particularly evident in Weiser's calculations of polarization energy due to an interstitial ion [54]. This presents a serious problem for such calculations. A typical supercell scheme would force these long-range effects to terminate at the third or fourth coordination site, with unknown consequential errors.

According to the calculations of Bar-Yam and Joannopoulos [107], the equilibrium sites for I^0 and I^{++} are the hexagonal and tetrahedral sites, respectively. This is just the opposite to what one would expect from the effect of ionic polarization according to the Weiser theory, as discussed earlier. The migration energies they calculated, with lattice relaxation, are 1.0 and 1.4 eV for I^0 and I^{++} , respectively. They later revised their results [109] to 1.2 eV for the self-interstitial in either charge state. They further differentiated the equilibrium geometrical configurations between p-type and n-type silicon, and reported that [109] in n-type silicon the tetrahedral site is the equilibrium site for I^{++} and the hexagonal site is the saddle point — i.e. the reverse of the p-type situation. Again, the hexagonal-tetrahedral configuration is the reverse for I^0 between p-type and n-type silicon. This finding is interesting, but does not seem to have an easy physical explanation. What differentiates n-type from p-type silicon is only the presence of nonlocalized conduction electrons. Why should nonlocalized electrons so strongly affect the potential energy at different sites? One would intuitively suppose that the doping type would merely shift the formation energy of a charged defect uniformly in real space, by an amount corresponding to the Fermi level. In n-type silicon, the concentration of I^{++} would be low according to Fermi statistics, as shown by Shockley et al. [6,7]. Bar-Yam and Joannopoulos did not discuss this problem in their paper [109].

Car et al. [103] have also calculated the silicon self-interstitial energetics for various geometrical configurations and charge states. They also found that the tetrahedral site is the equilibrium site for I^{++} , with the hexagonal site being its saddle point. This tetrahedral-hexagonal geometrical configuration is reversed for I^0 , which is again in disagreement with the polarization energy concept of Weiser (which favors the hexagonal site as the equilibrium site for an ionized interstitial). Another peculiarity in the total energy vs. geometrical location plot for I^0 , from the results of Car et al. [103], is the existence of an energy cusp at the tetrahedral site. In the results of Car et al., the effect of the Fermi level is simply to shift the formation energy uniformly in real space by an amount equal to the Fermi energy (or twice that amount for doubly charged defects), up or down according to the sign of the charge. This is physically expected. They calculated the total energies for I^{++} , I^+ and I^0 . Their results suggest that the Bourgoin-Corbett mechanism may occur along the TB path: an I^{++} at the tetrahedral site (T) captures an electron and become I^+ , which then finds the neighboring bond center site (B) to be at a lower energy and moves there without thermal assistance. Another possible path for the Bourgoin-Corbett mechanism was suggested to be the TBTH path. The BS path (bond center to split interstitial) of a thermal migration, proposed earlier by Watkins et al. [114], was found not to be possible. They found that I^+ is not a stable charge state. In other words, I^0 , I^+ and I^{++} form an Anderson negative-U system.

Bar-Yam and Joannopoulos [107–109], Car et al. [103,105] and Kelly and Car [106] did not report any negatively charged self-interstitials. This appears to be in conflict with overwhelming experimental evidence that phosphorus diffuses in silicon via a dominantly interstitialcy mechanism [3,115–117], and that phosphorus diffusion is enhanced approximately in proportion to electron concentration to the second power (for example, see [3,63,64]),

indicating that the self-interstitial may have a dominant double negative charge state. (The concentration of neutral defects is independent of the Fermi level [6,7]).

The values of E_V^f , E_V^m , E_I^f and E_I^m vary somewhat from different calculations. A summary of the results from various theoretical calculations is given in Table 1.

The pressure dependence of defect energetics in silicon appropriate for self-diffusion, including the concerted exchange mechanism of Pandey [38], has been calculated by Antonelli and Bernholc [110]. The pressure dependence of the energetics of point defects in mediating impurity diffusion in silicon has been calculated by Sugino and Oshiyama [112]. They found that the activation energies for the diffusion of phosphorus, arsenic and antimony in silicon decrease with pressure for the vacancy mechanism, but increase with pressure for the interstitialcy mechanism. Comparing their calculated results with the experimental results of Nygren et al. [118], who found the rate of diffusion of arsenic in silicon to increase with pressure, Sugino and Oshiyama concluded [112] that arsenic diffuses in silicon via a dominantly vacancy mechanism. They seemed to have neglected the fact that what Nygren et al. measured is the pressure effect on the *activation enthalpy*, rather than the *activation energy*, of diffusion. Sugino and Oshiyama [112] reported a decrease of activation energy of diffusion by 0.6 eV at 60 kbar. The change in activation enthalpy would then be equal to -0.6 eV, plus a term of $P \Delta V$ of about 0.68 eV (ΔV being approximately given by the atomic volume), or +0.08 eV. A similar correction, but with an opposite sign, is needed for the activation enthalpy of the diffusion via an interstitialcy mechanism. In view of the level of accuracy of such supercell calculations, a definitive assessment of the diffusion mechanism would seem rather unwarranted.

While the LDA-based total energy calculations have yielded remarkably good results, in so far as their compatibility with experimental data of self-diffusion is concerned, some of the limitations and reliability of such calculations should be not be overlooked. Kelly and Car [106] have noted that for the calculation of the formation energy of silicon, a relaxation of the next-nearest neighbors would give such a huge change in value that a reliable first principles calculation with a sufficiently large cluster is impractical for present-day computers. A similar problem occurs in the supercell method. A practical limit on the size of the supercell leads to very large dispersion in the gap state. Furthmüller and Fähnle [119] have noted that, for chalcogen defects in silicon, a significant dispersion occurs even for a supercell size of 128 atoms. Furthermore, the density-functional theory is a theory

Table 1
Theoretical formation and migration energies (eV) of intrinsic point defects in silicon

Defect	E_I^f or E_V^f	E_I^m or E_V^m	References
$I^{++}(T)$	4.7 ± 0.5	0.4 ± 0.5	[104]
	4.0	0.2 ± 0.2	[108]
	3.6		[110]
	4.4	1.2	[106]
$I^0(T)$	4.3	1.6 ± 0.5	[104]
		1.2 ± 0.3	[108]
	4.3	0.3	[110]
$I_0(B)$	5.0		[110]
V^0	4.4		[110]
	4.4 (unrelaxed)		[106]

for the ground state. In the local density approximation, it has proven unreliable for energies of excited states, such as the conduction band and conduction band derived states. This caveat has indeed been noted by some authors of those calculations, of whom some would refrain from giving values for such energy states.

3. Transformation of point defects in cooling

3.1. *Swirl defect manifestation of intrinsic point defects*

As a new part of a silicon crystal is freshly grown from a melt, it contains equilibrium concentrations of vacancies and self-interstitials, as well as their complexes. As that part of the crystal moves away from the melt and becomes colder, the majority of these grown-in point defects become excess over their new equilibrium concentrations at a lower temperature, and are subject to population reduction. In a crystal that is substantially dislocation-free, the only ways open to their depopulation are vacancy–interstitial annihilation and the parallel process of agglomeration among like species themselves into extended defects. Silicon wafers produced under a variety of conditions exhibit a kind of *microdefect* that is commonly referred to as a *swirl defect*. Such defects can readily be examined visually by a number of techniques, such as preferential etching, copper decoration and transmission electron microscopy (TEM). These microdefects are usually distributed throughout a crystal in a striated pattern, due to the nonuniformity in the temperature field in the crystal growth environment in which a crystal rotates during growth, or due to an oscillatory dynamics of crystal growth. As the solid–melt interface is usually concave, the distribution of these microdefects on a flat-cut surface of a wafer exhibits a swirl pattern from preferential etching. A stylized schematic distribution of swirl defects is shown in Fig. 1. In reality, however, the striation is spiraling along the growth axis because during growth the crystal is being pulled axially while rotating through the furnace environment, and the statistical fluctuations of defect agglomeration cause the distribution pattern on an axial cross-section to appear as swirls. For this reason, these microdefects are also referred to as *swirl defects*.

Plaskett [120] first reported the occurrence of what were thought to be clusters of vacancies in the inner region of a silicon crystal. The width of the cluster-free outer ring of a wafer cut from the crystal is about 1.5 mm, which was thought to be indicative of the vacancy diffusion length in the temperature transient during the growth. Abe [121] first reported the distribution of shallow etch pits of microdefects in swirl patterns. Swirl defects were first studied in detail by de Kock [122], especially regarding the growth conditions under which such microdefects would appear and disappear. He noted two kinds of such microdefects, which he termed type A and type B. These were conjectured to be clusters of point defects or clusters of complexes of impurities and point defects. Type A defects are, relatively speaking, quite large and can be readily examined by TEM. Subsequently, Föll and Kolbesen [123] identified the A defects as interstitial type dislocation loops. From this observation, they proposed that the dominant intrinsic point defect in silicon is the self-interstitial. The A defects would form from the agglomeration of excess self-interstitials that had been left after vacancy–interstitial annihilation. The B defects could not be characterized by TEM; but they can be revealed in their spatial distribution by means of selective etching or metallic decoration. Hu suggested an alternative model of swirl defect formation [124]: the excess vacancies and self-interstitials do not annihilate completely, and

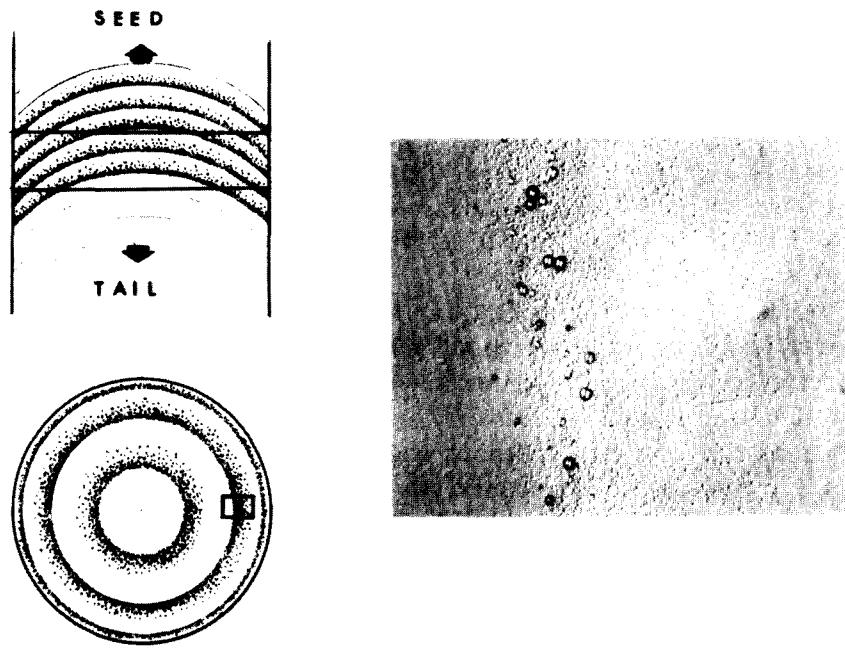


Fig. 1. Schematic illustration (left) showing the distribution of swirl defects in a crystal: left top, longitudinal section; left bottom, transverse section. The photomicrograph (size 400 μm by 500 μm) of swirl-defect etch pits on the right corresponds to the framed rectangular area at the left bottom (taken from Hu [124]).

significant fractions of the excess populations of both species agglomerate, separately and in parallel, into the A and B defects. The larger A clusters are agglomerates of self-interstitials in the form of dislocation loops (interstitial-type disks), while the smaller B clusters, as yet unidentified, are agglomerates of vacancies in the form of small globules. There is a simple reason for the preference of interstitial-type extended defects to take disk shape: it is the geometry of minimum strain energy for a mismatched inclusion in a matrix, as shown by Nabarro [125,126]. An interstitial type inclusion represents an insertion of extra matter into a matrix, and consequently produces a normal compression against the host matrix. A spherical cavity, by contrast, produces no, or very little, strain in the host matrix. It is for this reason that it cannot be detected by TEM (aside from a small change in matter density). For vacancy clusters, the spherical cavity is therefore the geometry of minimum energy, in both the strain energy and the surface energy components, and is the preferred form. A plausible explanation for the separate agglomeration of excess vacancies and self-interstitials among like species is that there may exist an energy barrier to vacancy-interstitial recombination. The rationale of the existence of a recombination barrier is that the equilibrium configurations of the vacancy and the interstitial involve relaxation of atomic structure surrounding each defect, and a reordering of the unrelaxed structures for both the vacancy and the self-interstitial is required preceding their recombination [124]. From a study of the transients of retarded antimony diffusion under thermal oxidation, a recombination energy barrier of ≈ 1.4 eV has been estimated [31]. Alternatively, an entropy barrier for vacancy-interstitial recombination has also been proposed [127]. The idea derives from the suggestion [11] that, at high temperatures, the silicon self-interstitial is an extended structure of a disordered region; and so, too, may be the vacancy. An extended point defect is one of the explanations for the very large pre-exponential factor experimentally found for the silicon self-diffusion. It is argued [127] that the vacancy-interstitial recombination

event is preceded by the reordering of the structures of both point defects that results in a negative entropy of about $-11.5k$. The vacancy-interstitial recombination will take place, regardless of which or both barriers are present; but the rate of recombination may not be fast enough to prevent self-agglomeration of vacancies and self-interstitials in parallel. De Kock and van de Wijgert [128] have endorsed such a model for the formation of the A and B defects. An alternative model [129–131] is that the self-interstitial is the predominant point defect in silicon at high temperatures, and that B defects are interstitial agglomerates, which would collapse into A defects.

Two more types of microdefects, the C defect [132] and the D defect [133] have later been reported. The C defect is seen only sometimes, and remains somewhat of a mystery. The D defect, like the B defect, is very small and in general invisible to TEM. A distinguishing feature of the D defects is that they are distributed uniformly, rather than in a striated pattern in the manner of A and B defects. Rosknoer suggested that they result from agglomeration of vacancies, a view that has received support from many subsequent investigators. But a contrary view has been expressed by the Ioffe group [134], who reported that strain contrasts, obtained with high-resolution electron microscopy, of all microdefects, including the D defect, are of the interstitial type. Their model of microdefect formation is as follows: in regions of vacancy excess, vacancies will complex with interstitial oxygen atoms (not self-interstitials) to form D defects. In regions of interstitial supersaturation, self-interstitials, oxygen and carbon atoms agglomerate to form B defects which, with further addition of self-interstitials, become A defects. At present, a definitive conclusion about these microdefects, with the exception of the A defect, remains unavailable.

Under the experimental conditions of Roksnoer [131], who grew their crystals by the pedestal method (somewhat similar to the floating zone method), D defects will not form until the crystal growth rate exceeds 5.4 mm min^{-1} (see Fig. 2). In contrast, A and B defects may start to form at a crystal growth rate of $\geq 0.2 \text{ mm min}^{-1}$, but will not form when the growth rate exceeds $4.5\text{--}5.0 \text{ mm min}^{-1}$. The threshold growth rates for the appearance and disappearance of various types of microdefects depend strongly on the diameter of the crystals and the method of crystal growth. For example, it has been reported [135] that in Czochralski crystals D defects start to appear at a growth rate of $\geq 0.8 \text{ mm min}^{-1}$. It should be noted that there are many differences between float zone and Czochralski

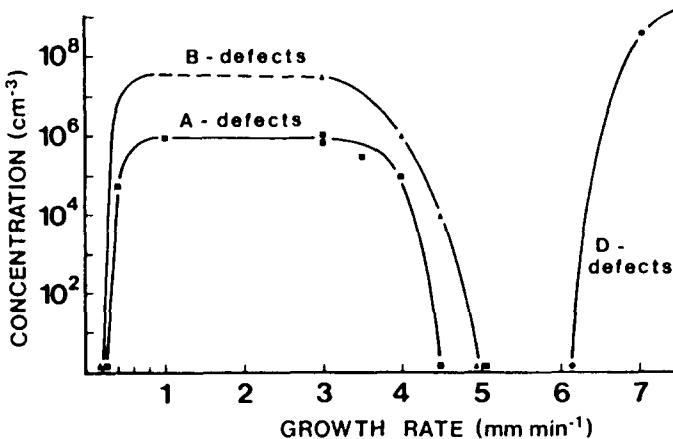


Fig. 2. Formation of A, B and D microdefects as a function of crystal growth rate. Crystal diameter 23 mm, atmosphere pure argon, carbon concentration $1.0\text{--}1.5 \times 10^{16} \text{ cm}^{-3}$ (taken from Roksnoer and van den Boom [133]).

silicon crystals, aside from the higher oxygen content in the latter. The temperature distribution during the growth is also quite different. While it would seem that the threshold growth rate may be indicative of the diffusion-limited annihilation velocity of a particular point defect species at the melt-crystal interface, model analysis is predicated on the appropriateness of assumptions. Voronkov [136] has given estimates for C_v , D_v , C_i and D_i from just such an analysis. But Voronkov's formulation of the swirl defect formation is an oversimplification, considering the crystal velocity from the interface and the concentration gradient as the only factors driving a point defect flux. Among the more important factors ignored is the thermal diffusion in a rapidly varying temperature field along the crystal axis. Also, the effect of impurity atoms on the formation of swirl defects, as has been noted by many researchers, was also not taken into account in Voronkov's formulation.

At the present time, there is still no satisfactory model for the formation of the various types of swirl defects; but they are undoubtedly manifestation of intrinsic point defects. There have been relatively few investigations of swirl defects in Czochralski silicon crystals. Most results [128,132,135,137,138] seem to indicate similar behaviors in the formation of microdefects in both kinds of crystals. Hence, at present, it is difficult to assess the role of oxygen in the formation of microdefects.

3.2. Thermal point defects in quenched silicon

Perhaps due to the influence from metal research, extensive earlier efforts have been made attempting to quench point defects in silicon from high temperature, hoping to study their annealing characteristics and electronic levels. The quenched-in defects are frequently referred to as "thermal defects", and the subject of research has a rather interesting and protracted history. Since the first observation, reported by Gallagher [139] in 1955, various investigators have reported that, when a silicon substrate is heated at some high temperatures (generally ≥ 800 °C) for some length of time and then quenched, a significant concentration of deep-level donors is introduced into the substrate. These deep-level centers have since been called "thermal defects". Early speculations on the identity of these centers range from silicon self-interstitials [140] to vacancy clusters [141,142]. In the investigation of Boltaks and Budarin [142], the concentration of the thermal point defects was derived from measurements of the change in specimen density and lattice constant after quenching. Both Elstner and Kamprath [141] and Boltaks and Budarin [142] reported the energy of formation of the "vacancy" to be 2.5–2.8 eV. Quenching experiments for the investigation of native point defects in silicon have proven to be difficult and unreliable. Unlike in most metals, the concentrations of both the self-interstitial and the vacancy are very low. The change in the length of a specimen is related to the change in the volume of the specimen by $\Delta l/l = \Delta V/3V$. For an increase of the vacancy concentration by $1 \times 10^{16} \text{ cm}^{-3}$, the length of a silicon specimen increases by less than 0.7×10^{-7} , a change too small to be measured accurately. Furthermore, silicon samples can be easily contaminated by fast diffusing impurities such as copper, nickel and iron, which can diffuse through the silica furnace tube at high temperatures. While the opinions of various investigators have diverged, their experimental results exhibit a high degree of agreement. The donor level observed in various investigations was in a narrow neighborhood around $E_v + 0.40$ eV. Various investigations since 1977 [143–150], however, have led to the general consensus that the "thermal defects" are after all not native defects or their clusters, but are contaminants from the furnace, passing through silica tubes and getting into the silicon substrates. It may be remarked that in

1956 Collins and Carlson [151] already expressed their suspicion that there could be a connection between the thermal defects observed by Gallagher and the iron donor in silicon that they studied, because of their many similarities.

The main conclusion is that, so far, attempts to quench-in point defects have not been successful. This perhaps should have been expected, in view of our current understanding of the extremely high diffusivities of the point defects.

4. Thermal oxidation, enhanced diffusion and stacking fault formation

4.1. Mechanism of interstitial injection in silicon oxidation

While extensive investigations have been made on the oxidation enhanced diffusion and the formation of oxidation stacking faults, and while it has been firmly established that thermal oxidation of silicon injects self-interstitials, the precise mechanism of how this occurs has remained unclear. It was proposed by Hu [2] that incomplete oxidation of the silicon surface immediately beneath the oxide leads to some surface silicon atoms being ejected as interstitials; the incomplete oxidation is a consequence that only a fraction of monolayer of silicon surface atoms is needed to form a monolayer of SiO_2 . Ejection of surface silicon atoms as interstitials is a way of relieving the stress in SiO_2 that would have otherwise developed to an immense intensity because of the fact that a unit volume of silicon would turn into 2.25 unit volumes of SiO_2 upon oxidation.

Another way of relieving stress in thermally grown SiO_2 is for the SiO_2 to grow only vertically via a kink oxidation mechanism proposed by Mott [152,153]. In this model, an oxidant molecule attacks the silicon atom at the kink site, and the extra volume of the oxide formed is accommodated perpendicularly to the interface. The oxidation of the atom at the kink also causes the kink to move to the next atomic site, and the process repeats (see Fig. 3). The kink oxidation model had earlier been used to successfully explain a number of phenomena in metal oxidation [154–156]. The operation of the kink model of silicon oxidation would not only avoid stress development in the oxide, but also obviate the generation of self-interstitials. Furthermore, the kink oxidation model would predict a much higher linear oxidation rate constant for a (100) vicinal plane than for a (100) plane, contrary to experimental observations.

The kink model of oxidation was later taken up by Tiller [157]. Tiller's paper on oxidation mechanism was later cited by Hirth and Tiller [158] in a single-sentence comment on Hu's kink regrowth model, expressing their opinion that kinks can only serve as sources, not sinks, of self-interstitials. This comment was later repeated by Dunham and Plummer [159]. This opinion is easily shown to be incorrect. Since ledge kinks are equilibrium

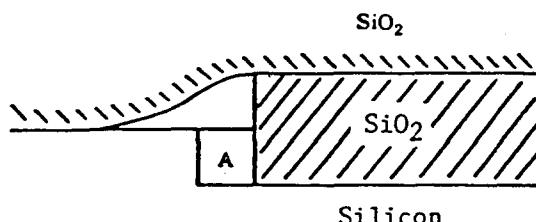


Fig. 3. A kink model of thermal oxidation of silicon. An oxidant molecule attacks the silicon atom A at the kink. That atom grows into SiO_2 vertically and produces no lateral strain (taken from Mott [152]).

structures, thermodynamics requires that they serve as sources as well as sinks, depending on whether the interstitials are in supersaturation or undersaturation. In fact, using STM (scanning tunneling microscopy) for the investigation of the initial stages of Si(100) oxidation, Cahill and Avouris [160] have found that silicon interstitials ejected from oxidation have regrown into islands. Regrowth on kinks is well known and well founded, and operates ubiquitously in such processes as epitaxial growth.

Viscoelastic flow is another way of relieving stress in thermally grown SiO_2 . SiO_2 has an extremely high viscosity, characterized by an activation energy of flow of 5 to 7 eV, depending on its hydroxyl content [161]. A significant rate of viscous flow can occur only at temperatures $\geq 1000^\circ\text{C}$.

An idea embraced by a number of investigators is that compressive stress in the thermal SiO_2 may produce in the silicon substrate a tensile stress, which in turn causes the formation of extrinsic stacking fault and the injection of silicon self-interstitials. Another way of viewing it is that the interstitial injection is related to the viscous flow of the oxide. A model on this line was put forth by Tan and Gösele [162]. There is no doubt that, at high temperature, viscous flow will occur to some extent. But the idea of relating stress and viscous flow to interstitial injection contradicts many known facts. For example, it is known that compressive stress in the oxide during oxidation is greatly reduced in steam oxidation [163,164], vis-à-vis dry oxidation. In fact, EerNisse [165] has shown that, at temperatures $\geq 975^\circ\text{C}$, oxides during oxidation in wet oxygen are completely free of stress. Yet OSF have been found to be larger and more numerous, and the related OED to be greater in steam oxidation [2]. It is also known that stress in SiO_2 decreases with increasing oxidation temperatures, and becomes essentially zero at $\geq 1100^\circ\text{C}$, yet OSF are found to increase with temperature. Both of these two facts show that the trends of stress intensity and of OSF size actually run in opposite directions. Furthermore, it has been shown that, in backside-oxidized silicon [75], pre-existing stacking faults would grow when silicon surfaces were covered with Si_3N_4 , whereas they would shrink when silicon surfaces were covered with SiO_2 . It is well known that Si_3N_4 surface films exhibit a very high tensile stress of $\approx 1 \text{ GPa}$ [166], thus placing the silicon substrate under compression.

4.2. Experimental observations of OSF and OED

Some earlier investigators variously reported the kinetics of OSF growth to follow a linear growth law [21], a parabolic law [167], or a 4/5th power law [24]. More careful investigations carried out later found the growth kinetics to be more complex, but may be approximated by a power law of t^m , where m ranges from 0.66 to 0.85 but is generally around 3/4 [168–170]. The activation energy of growth is about 2.3 eV, but has also been reported up to 2.6 eV. There is also a retrogrowth phenomenon at high temperatures, around 1240°C or so, depending on the surface orientation of the silicon substrate [168,169,171]. An example is shown in Fig. 4. In inert ambients, the shrinkage of pre-existing OSF has an activation energy of around 4.6–5.2 eV [75,172–176]. As already discussed in Section 2.3, the Arrhenius plot of silicon self-diffusivity shows a break at about 1100°C , with an activation energy 4.7–4.9 eV in the high-temperature regime and 4.1–4.2 eV in the low-temperature regime [43,44]. If the kinetics of shrinkage is diffusion-limited, one should expect the activation energy of shrinkage in inert ambients to be identical to that of self-diffusion. If there is an additional reaction barrier to OSF growth shrinkage, the activation energy would be higher. Considering the lower activation energy observed in silicon isotope

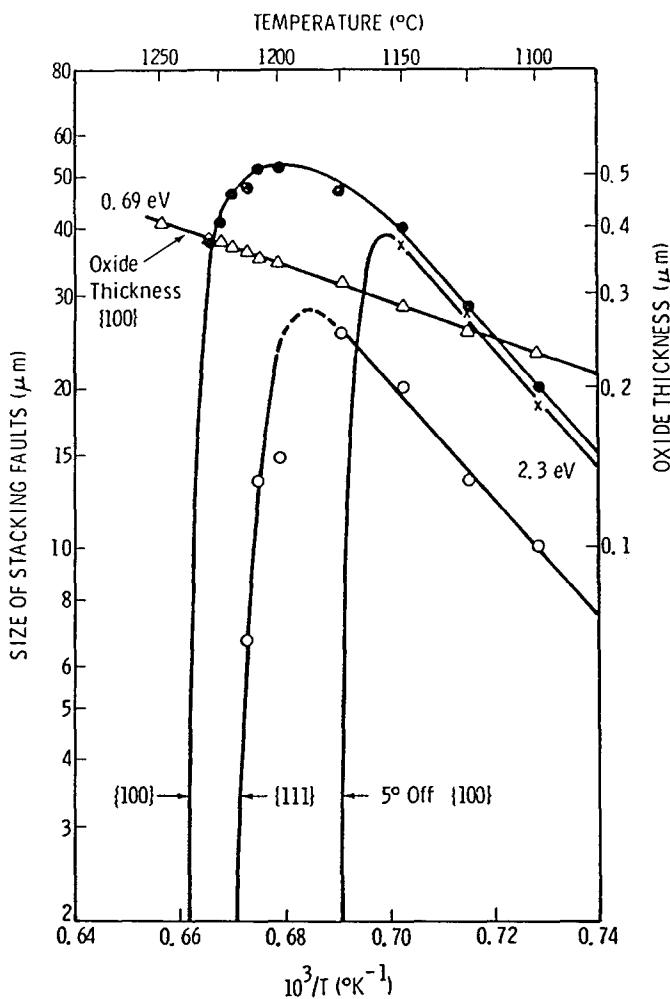


Fig. 4. Growth vs. temperature of oxidation stacking faults on {100}, {111}, and vicinal {100} (5° toward {110}) surfaces in dry oxygen for 3 h (taken from Hu [168]).

diffusion, and allowing for the fact that the experiments on stacking fault shrinkage are generally carried out at temperatures somewhat lower than for experiments on isotope self-diffusion, we can only say that there may be a reaction barrier which is not very large, perhaps ≈ 0.4 eV.

4.3. Kinetics of interstitial injection in silicon oxidation

4.3.1. The model of Hu [2,177-179]

It seems natural to assume that the rate of generation of self-interstitials is linearly proportional to the rate of thermal oxidation, as suggested by Hu [2,177]. The main problem is to derive the time dependence of the supersaturation of self-interstitials from this premise. The kinetics of interstitial supersaturation should then be corroborated by OED and the kinetics of the growth of OSF. In 1974, when the theory of OED and OSF was first formulated, very reliable and accurate experimental kinetic data were not available. Based on mainly qualitative observations, Hu [2] built his model on the following two assumptions:

1. Thermal oxidation injects self-interstitials at a rate that is directly proportional to the rate of thermal oxidation.
2. The excess self-interstitials flood the bulk, and at the same time annihilate at the Si–SiO₂ interface via a surface regrowth process. The rate of interface annihilation is directly proportional to the concentration of excess interstitials and the concentration of surface sinks considered to be ledge kinks.

The postulation of ledge kinks as sinks for excess self-interstitials also serves to explain the effect of surface orientation on the kinetics [2]. The silicon ejection and regrowth model is physically well grounded and, as already mentioned, is consistent with the STM observations of the initial stages of Si(100) oxidation by Cahill and Avouris [160]. The proposition of annihilation of excess interstitials at the Si–SiO₂ interface is also important in the simulation of two-dimensional distribution of excess point defects, first carried out by Hamasaki [180]. It is also important in explaining a phenomenon related to 2D distribution of excess point defects: the stripe width dependence of oxidation and oxynitridation-enhanced diffusion [181,182].

The buildup of excess self-interstitials is determined by the balance between generation and annihilation. The kinetics of interstitial supersaturation is then to be obtained from the solution of the following continuity equation [2]:

$$\frac{\partial C_I}{\partial t} = D_I \frac{\partial^2 C_I}{\partial x^2} - R_{\text{ann}}^b, \quad x > 0 \quad (20)$$

with the boundary condition

$$-D_I \frac{\partial C_I}{\partial x} + k_{\text{ann}}^s (C_I - C_I^*) = At^{-1/2}, \quad x = 0 \quad (21)$$

In the above equations, R_{ann}^b is the rate of bulk annihilation of the excess interstitials and k_{ann}^s is the rate constant of surface annihilation of the excess interstitials. The last term in Eq. (21) represents the rate of interstitial injection in direct proportion to the rate of thermal oxidation, which varies as $t^{-1/2}$ (in the parabolic regime of oxide growth), A being a proportionality constant.

In his original model, he opted for mathematical simplicity by making another assumption that the diffusion of self-interstitials is so fast, and there is so negligibly little bulk annihilation, that the concentration of self-interstitials is approximately uniform across the substrate thickness. This simplifying assumption reduces the problem to solving the following ordinary differential equation [2]:

$$\frac{dC_I}{dt} = At^{-1/2} - Bh(C_I - C_I^*) \quad (22)$$

with $B = \pi \rho a_0 D_I / h$.

In the above, A is a lumped parameter of oxidation kinetics and interstitial injection, and is time-independent, ρ is the density of interface annihilation sites, a_0 is the radius of the annihilation center, D_I is the diffusivity of the self-interstitial, and h is the thickness of the substrate. The solution of the above equation is

$$C_I - C_I^* = (2A/B^{1/2})F((Bt)^{1/2}) \quad (23)$$

where $F((Bt)^{1/2})$ is Dawson's integral, defined as

$$F(\eta) = \exp(-\eta^2) \int_0^\eta \exp(\zeta^2) d\zeta \quad (24)$$

The properties of Dawson's integral are such that for $\eta \gtrsim 0.4$, $F(\eta) \approx \eta$; it attains a maximum value of 0.541 at $\eta = 0.92$ and takes an asymptotic form of $F(\eta) \approx 1/2\eta$ for $\eta \gtrsim 6$. The time constant of annihilation, $1/B$, is of the order of 10 s, much smaller than the oxidation time. Hence the asymptotic form of Dawson's integral is appropriate for the case under consideration. In this regime, the model predicts the excess self-interstitials to vary as $t^{-1/2}$. The rate of growth of stacking faults in the presence of a supersaturation of self-interstitials may be written as

$$\frac{dr}{dt} = \pi a_0^2 D_I (C_I - C_I^*) \quad (25)$$

where, for simplicity, the atomic volume has been set approximately equal to a_0^3 . Using the asymptotic form of $C_I - C_I^*$ from Eq. (23), Eq. (25) is integrated to give

$$r(t) = (2\pi a_0^2 D_I A/B) t^{1/2} \quad (26)$$

The above equation predicts a parabolic growth law of OSF. Shortly afterward, from a more extensive set of careful experiments carried out in 1975, Hu [168] found that the kinetics of OSF follows more closely a $t^{0.75}$ power law, rather than a parabolic law. As already discussed earlier, subsequent investigations by many researchers generally corroborate this power law; but, dependent on growth parameters, the power law may vary between $t^{0.66}$ to $t^{0.85}$, as mentioned earlier. Furthermore, many investigators have subsequently reported that the interstitial supersaturation appears to vary as the square root of the rate of thermal oxidation [183–185].

It is therefore necessary to modify the original model in order to bring about an agreement with the experimental findings. To arrive at a $t^{0.75}$ power law for stacking fault growth, the interstitial supersaturation should change with time as $t^{-0.25}$ (See Eq. (25)). The idea that the rate of interstitial injection is proportional to the rate of oxidation seems so reasonable as to be irreproachable. One part of the model that may be open to question is the assumption that the interface is the only sink of excess interstitials. Many possible sinks for the excess self-interstitials readily come to mind. For example, the excess interstitials may also be annihilated in the bulk and/or in the oxide. Another part of the model that is obviously questionable is the grossly simplifying assumption of a uniform concentration of excess interstitials. Alternative models have since been explored by various investigators. We will discuss a few of them.

4.3.2. Bimolecular process of surface annihilation [177]

If the concentration of the excess self-interstitials is sufficiently high, a bimolecular annihilation process may prevail. Again with the simplifying assumption of uniform concentration of excess point defects, one may write in place of Eq. (22)

$$\frac{dC_I}{dt} = At^{-1/2} - k_2 C_I^2 \quad (27)$$

where k_2 is the rate constant of bimolecular annihilation. The asymptotic solution is given by [177]:

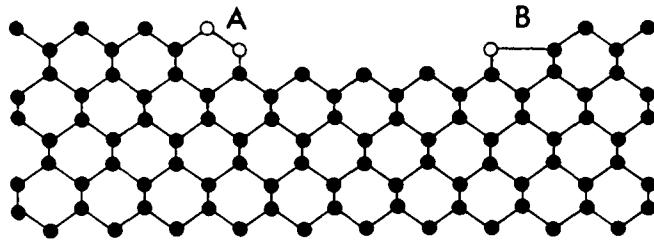


Fig. 5. Schematic illustration of step structures A and B for diatomic and monatomic mechanisms of surface annihilation of self-interstitials via step propagation (taken from Hu [177]).

$$C_I = (A/k_2)^{1/2} t^{-1/4} \quad (28)$$

The rationale for the feasibility of bimolecular annihilation at high concentration is illustrated in Fig. 5, which depicts a likely structure of steps on a silicon (100) surface. As seen, the simultaneous capture of two silicon atoms (open circles attached to the step on the left) is required in order to preserve the original step structure. The capture of a single silicon atom (open circle on the right) results in a new structure B, which may later revert to structure A by capturing another silicon atom. Since A is the equilibrium structure on the original surface, it is the minimum energy configuration. The bimolecular mechanism of step propagation may prevail if

$$C_I/C_S > \exp[(H_A - H_B)/kT] \quad (29)$$

where H_A and H_B are the enthalpies of formation of step structures A and B, respectively.

4.3.3. Excess interstitials diffusing into silicon [178]

Analytically solving Eq. (20) without the bulk annihilation term, and with the boundary condition of Eq. (21), Hu [178] obtained for a semi-infinite silicon:

$$C_I - C_I^* = A(\pi/D_I)^{1/2} e^{(\tau^{1/2} + \lambda)^2 - \lambda^2} \operatorname{erfc}(\tau^{1/2} + \lambda) \quad (30)$$

where

$$\lambda = x/2(D_I t)^{1/2} \quad (31)$$

and

$$\tau = tk_{\text{ann}}^s / D_I \quad (32)$$

As $C_I(x, t)$ varies with depth we will be interested only in its value in the region pertinent to OSF and OED. Since the diffusion length of important dopants is so much shorter than the diffusion length of self-interstitials, for OED we will be interested in the region very close to the surface, or $x=0$, where

$$C_I - C_I^* = A(\pi/D_I)^{1/2} e^\tau \operatorname{erfc}(\tau^{1/2}) \quad (33)$$

For OSF we will be interested in the average of C_I over the region equal to the size of OSF. Because the size of OSF is itself dependent on $C_I - C_I^*$, this average leads to an integral equation which cannot be solved analytically. Instead, we average C_I over the interstitial diffusion length, with the result

$$\langle C_I - C_I^* \rangle_\lambda = \frac{A}{2} \left(\frac{\pi}{D_I \tau} \right)^{1/2} [e^{(\tau^{1/2} + 1)^2 - 1} \operatorname{erfc}((\tau^{1/2} + 1) - e^\tau \operatorname{erfc}(\tau^{1/2}) + \operatorname{erf}(1))] \quad (34)$$

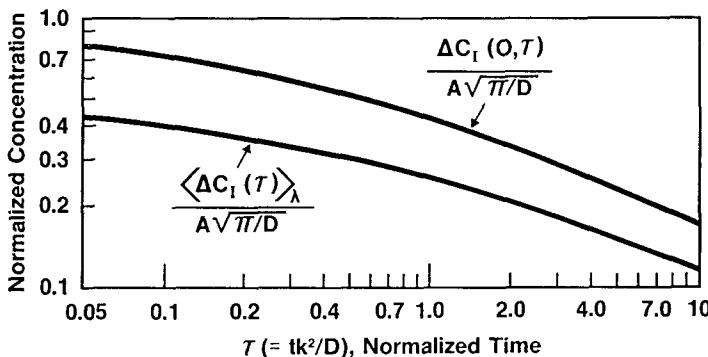


Fig. 6. Normalized concentrations of excess interstitials at the surface and averaged over the interstitial diffusion length, respectively, are plotted vs. normalized time. The analytical solution allows the excess interstitials injected by thermal oxidation to diffuse into a silicon half-space (taken from Hu [178]).

In the time range $\tau=0.1$ to 1, either expression can be approximated by a power law with a power of -0.25 to -0.21 . Fig. 6. shows both $C_I(0, \tau)$ and $\langle C_I - C_I^* \rangle_\lambda$. This time dependence of C_I will fit the experimentally observed time dependence of OSF and OED. For greater oxidation time, however, n of the t^{-n} power law approximation grows larger.

4.3.4. Excess interstitials diffusing and annihilating in silicon

Hu [179] has also considered the case with bulk recombination, as given in Eq. (20), with the bulk annihilation given by

$$R_{\text{ann}}^b = k_{\text{ann}}^b (C_I - C_I^*) \quad (35)$$

When the oxidation is in the short time period low-temperature regime, the oxide growth is linear and the interstitial injection rate becomes a constant. Hu has obtained and discussed an analytical solution for this case. His result will not be reproduced here, and the interested reader is referred to his paper [179].

4.3.5. Excess interstitials diffusing and annihilating in SiO_2 [159,176,184,186–189]

Lin et al. [176] presented a kinetic model of supersaturation of self-interstitials by allowing the excess self-interstitials to diffuse into the oxide and get oxidized along the way. This model has gone through various modifications by Dunham and associates [159,184,186,187], and Taniguchi and associates. The model contains the following essential elements:

1. Excess interstitials are generated by thermal oxidation at a rate that is directly proportional to the oxidation rate.
2. The majority of these generated interstitials flow into the oxide and get oxidized on their way, while a negligible amount flows into the silicon substrate. The balance between the generation and the leakage through the oxide determines the concentration of excess interstitials at the $\text{Si}-\text{SiO}_2$ interface.
3. The reaction-diffusion of silicon interstitials in the oxide is assumed to be in a steady state. The oxygen concentration within the reaction zone is assumed to be uniform and given by that at the $\text{Si}-\text{SiO}_2$ interface, $C_{\text{O}_2}(0)$.
4. The concentration on the silicon side is determined by a segregation coefficient for the interstitials between the silicon and the SiO_2 .

The steady state reaction-diffusion in the oxide is given by

$$D_I^{\text{ox}} \frac{d^2 C_I^{\text{ox}}}{dy^2} - k_{\text{ox}} C_{\text{O}_2}(0) C_I^{\text{ox}} = 0 \quad (36)$$

where parameters with the superscript ox pertain to the oxide phase. (Parameters without the superscript pertain to the silicon phase). k_{ox} is a reaction rate constant, and y is the axis in the direction of the normal to the silicon substrate (i.e. directed from the interface into the oxide). The linear differential equation has the simple solution of

$$C_I^{\text{ox}}(y) = C_I^{\text{ox}}(0) \exp(-y/L) \quad (37)$$

where L is the self-interstitial diffusion length in the oxide, and is given by

$$L = \left(\frac{D_I^{\text{ox}}}{k_{\text{ox}} C_{\text{O}_2}} \right)^{1/2} \quad (38)$$

The flux of interstitials into the oxide is given by

$$J_I^{\text{ox}}(0) = -D_I^{\text{ox}} \left(\frac{dC_I^{\text{ox}}}{dy} \right)_{y=0} = \frac{D_I^{\text{ox}} C_I^{\text{ox}}(0)}{L} \quad (39)$$

The oxidant concentration at the interface is related to the oxidation rate according to, for example, Deal and Grove [190], as

$$\frac{1}{\Omega_{\text{SiO}_2}} \frac{dw_{\text{ox}}}{dt} = k_s C_{\text{O}_2}(0) \quad (40)$$

where Ω_{SiO_2} is the molecular volume of SiO_2 , dw_{ox}/dt is the rate of oxide growth, and k_s is the reaction rate constant. The rate of generation of excess self-interstitials is assumed to be directly proportional to the oxidation rate as

$$G = \theta \frac{dw_{\text{ox}}}{dt} \quad (41)$$

Balance the rate of interstitial generation with the flux of interstitials leaking into the oxide:

$$J_I^{\text{ox}}(0) = G \quad (42)$$

Putting Eqs. (38)–(41) into Eq. (42), one obtains

$$C_I^{\text{ox}}(0) = \theta \left(\frac{k_s \Omega_{\text{SiO}_2}}{k_{\text{ox}} D_I^{\text{ox}}} \frac{dw_{\text{ox}}}{dt} \right)^{-1/2} \quad (43)$$

Finally, the interstitial supersaturation in the silicon is simply obtained by

$$C_I = m C_I^{\text{ox}} \quad (44)$$

where m is the segregation coefficient, assumed to be constant. Since in typical situations the thermal oxidation kinetics lies dominantly in the parabolic regime where $dw_{\text{ox}}/dt \propto t^{-1/2}$, Eq. (43) will yield a $t^{-0.25}$ power law of interstitial supersaturation, in agreement with experimental observations. Lin et al. [176] and Dunham and Plummer [159,187] suggested that the segregation coefficient may be dependent on surface orientation, so that it would explain the observed dependence of OED and ORD on surface orientation. But thermodynamics requires that the equilibrium segregation between two phases be independent of their interface. Were it otherwise, one could construct a perpetual motion machine by

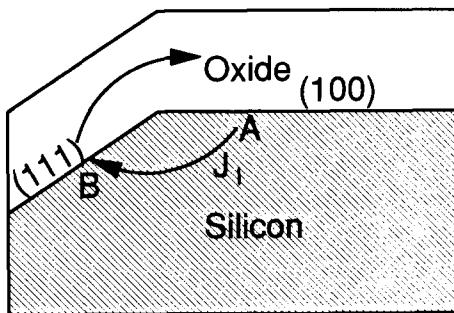


Fig. 7. A perpetual motion device based on the supposed anisotropic interface segregation coefficient m . If $m_{(100)} > m_{(111)}$, then in an initially quiescent (in a neutral ambient) structure as shown, the concentration of self-interstitials in silicon at point A would be higher than that at point B, causing an interstitial flux J_I , as indicated by the lower arrow, and tending to reduce the concentration difference between A and B. The increased concentration at B in turn causes an increase of the interstitial concentration in the SiO_2 across the interface from B, thus establishing a concentration gradient in the oxide so as to drive an interstitial flux in the oxide, as indicated by the upper arrow.

joining two pieces of different materials at two ends at two different crystalline angles, as illustrated in Fig. 7. It should be clear that interface annihilation is the only process that can and should manifest an orientation effect in the kinetics of interstitial supersaturation. Furthermore, the orientation dependence of phosphorus and boron diffusivity is observed only in oxidizing ambients, but has never been observed in inert ambients (with point defects in thermal equilibrium) [12–19]. In inert ambients, there will still be self-interstitials, albeit at a lower concentration than in oxidizing ambients. Self-interstitials must still diffuse into the oxide, except that they will not get oxidized. Since the oxide is amorphous, the concentration there should be independent of the orientation of the silicon substrate. An orientation-dependent segregation coefficient between SiO_2 and silicon would have caused an orientation dependence of diffusivity in any ambients.

For the boundary condition in their latest model, Taniguchi et al. [188,189] have added the interface annihilation as well as the diffusion of excess interstitials into the silicon substrate. The interface annihilation included in their model is the same as the one included in all of the variations of Hu's model. In all cases, however, the orientation effect has not been made quantitative, although an interpretation of it was given early by Hu [2].

Finally, it may be noted that the recent study of Murrell et al. [191] shows that the amount of silicon diffusing into the oxide during silicon oxidation is very small; they placed an upper limit of ≈ 0.2 of a monolayer on the distance silicon diffuses into the oxide. Thus the formation of new oxide occurs essentially at the $\text{Si}-\text{SiO}_2$ interface.

4.4. Effect of chloro-compounds on excess interstitials

Many researchers [169,192–195] have reported that addition of HCl, chloroethylene or other chlorine-containing compounds to the oxidizing ambient significantly reduces interstitial supersaturation caused by oxidation. Fig. 8 shows an example of the effect of adding a varying percentage of HCl in the oxidizing ambient. So far, there has been no theoretical model of this phenomenon.

5. Vacancy generation in silicide formation

Silicides are often employed in silicon integrated circuits for electric contacts between metal interconnection lines and devices in the silicon substrate. Typically, after formation

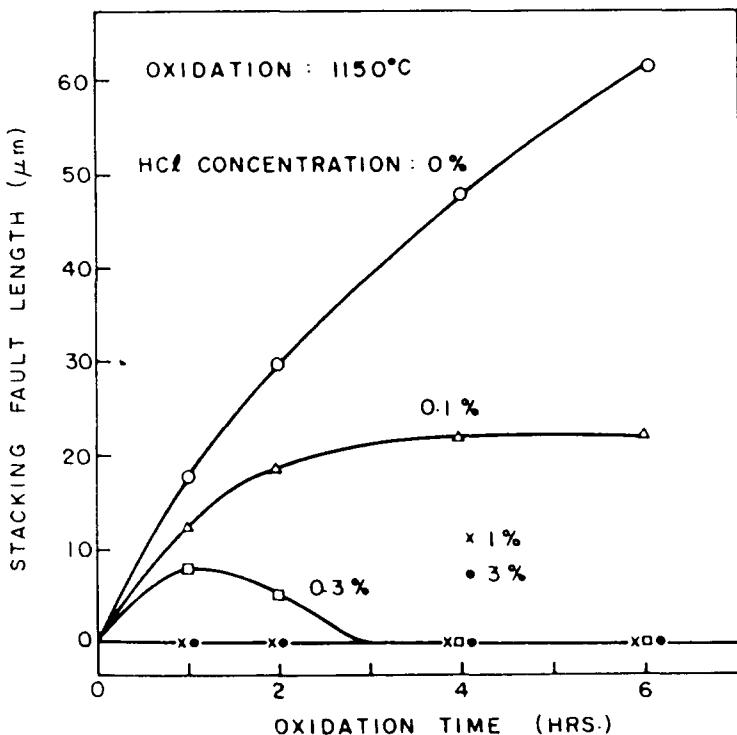


Fig. 8. The effect of HCl on the growth of oxidation stacking faults at 1150 °C. 0%, 0.1%, 0.3%, 1% and 3% of HCl are respectively added to the dry oxygen oxidizing ambient (taken from Shiraki [192]).

of silicides the IC will not be exposed to temperatures higher than ≈ 400 °C. But at even lower temperatures, the formation of silicides such as PtSi and NiSi has been reported to cause significant enhanced movement of arsenic from the surface into the substrate [196,197]. This has been explained by a “snow-plow” effect, by which arsenic atoms are rejected, or “plowed”, from the region of silicon being converted into silicide. However, this explanation is not quite satisfactory, at least by itself. Snow-plowing can only explain the pile-up of impurities at one side of an advancing interface, but does not explain how an enhanced movement occurs as manifested in a broadened impurity profile, or a lengthened profile tail. It remains unclear how the arsenic atoms at the interface can move a significant distance into the substrate at such low temperatures. A plausible mechanism is that arsenic atoms are ejected by the advancing interface into interstitial sites and become highly mobile until taking up substitutional sites again. It was also reported [196] that the observed enhanced movement of dopants occurs only during the formation of near-noble metal silicides, but not refractory silicides such as TiSi and TaSi₂.

More recently, Wittmer et al. [198,199] reported enhanced movement of antimony in buried layers in silicon substrates under surface silicidation (PdSi₂) at very low temperatures (200 °C). The snow-plow effect would no longer matter in the buried layers. Furthermore, the enhanced movement was reported to be asymmetric, and the profile broadened only toward the surface, as if driven by a unidirectional force. It is very difficult to conceive of a force that would bring about this effect. Recently, Pichler and Orlowski [200] suggested that this anomalous diffusion can be consistently explained by postulating the existence of an inhomogeneous stress field as an external driving force in conjunction with stress relaxation mechanisms due to the silicon self-interstitial diffusion. They actually reported a simulation

using this model and showed good agreement with the experimental results of Wittmer et al. It now appears that this effect is actually spurious, and is caused by the roughening of the sample surface during sputtering in the SIMS (secondary ion mass spectrometry) analysis of the impurity profiles. Ronsheim and Tejwani [201], in an attempt to repeat the experiment of Wittmer et al., have found that the broadening in the measured dopant profile of the silicided sample can be eliminated by changing the penetration or mixing depth of the sputtering beam in the SIMS analysis. This method of changing the sputtering conditions to minimize depth resolution degradation in dopant profiling is an established technique [202,203]. Furthermore, they were able to obtain an asymmetrically broadened SIMS profile similar to that obtained by Wittmer et al. [198] by using a higher energy (5.5 keV vs. 3.0 keV) O_2^+ sputtering beam at a different incident angle (40° vs. 47°), a condition that led to surface roughening.

Fig. 9 illustrates how two different surface pitting topographies would cause two different artificial broadenings of impurity profiles by SIMS. The depth of a point on an impurity profile measured by SIMS is actually the depth from the floor of a pit, rather than from the original surface. The impurity concentration profile $C(x)$ becomes $C(x-h)$, where h is the depth of the pit. By convolution with a random distribution of depth, pits having straight walls would broaden an impurity profile symmetrically. If the depth distribution of the pits is statistically nonrandom, or geometrically skewed as in pyramidal pits, then the SIMS-measured profile would broaden asymmetrically. Since a section in the bottom part of a pyramidal pit has a smaller surface area than a section on the upper part, a smaller areal fraction of an impurity profile will be shifted toward the surface, resulting in an asymmetrical broadening of the impurity profile as seen by Wittmer et al. [198,199].

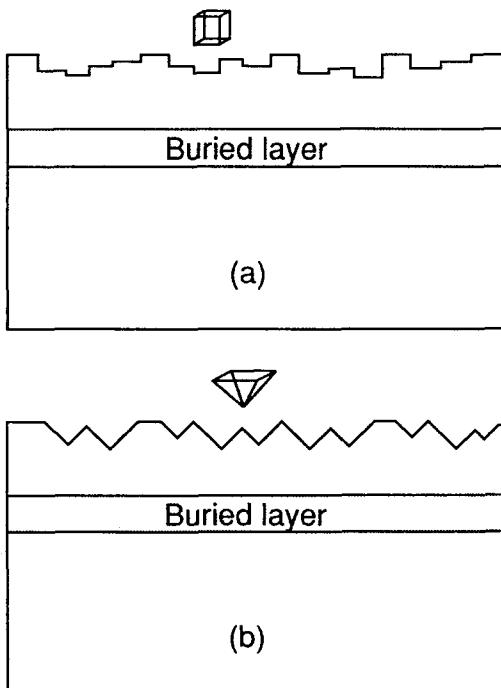


Fig. 9. Schematic of two types of roughened surface geometry. Structure (a) with straight-walled pits and random depth distribution would lead to symmetrical broadening in SIMS profiles. Structure (b) has pyramidal pits such that deeper recesses are geometrically weighted less, giving SIMS profiles that are asymmetrically broadened toward the surface.

At higher temperatures, instead of or in addition to the snow-plow effect, the silicidation process could conceivably generate excess point defects, either vacancies or self-interstitials. To test this idea, Hu [204] carried out experiments using buried-layer monitors that are several μm below the silicide–silicon interface, avoiding the snow-plow effect. Buried-layer monitors also allowed the distinction between point defect supersaturation/undersaturation from the Kirkendall effect, or the effect of vacancy/interstitialcy wind. Boron and antimony are used as the monitor diffusants in the buried layer. Boron diffusion in silicon had been shown to have a large interstitialcy component, while antimony diffusion has an almost entirely vacancy mechanism. $\text{TaSi}_{1.8}$ films were deposited by sputtering onto the silicon substrate through patterned windows in the silicon nitride film, as illustrated in Fig. 10. The samples were annealed in nitrogen at 950 °C for 15 h, beveled and then analyzed by spreading resistance profiling; subsequently they were copper-stained. The silicidation is expected to convert TaSi_x , where $x < 2$, to TaSi_2 . The result is that both boron and antimony buried layers are broadened, both toward the surface and into the bulk, as illustrated in Fig. 10. The result can only be explained by a vacancy supersaturation created by silicidation. It may also be noted that the result contradicts the supposition of Wittmer and Tu [196] that refractory metal silicides, and more particularly silicides in which silicon is the predominant diffusing species, would not cause enhanced diffusion.

The mechanism of generation of vacancies by a silicidation process may be as follows: silicidation occurs by removing individual silicon atoms from the silicon substrate at the silicon–silicide interface, and transporting them via a vacancy mechanism across the silicide film to the silicide–metal interface, where the reaction between silicon and metal takes place. As individual silicon atoms leave the substrate, vacancies are created and diffuse into the bulk of the substrate, causing a local supersaturation (Fig. 11.) This is in contrast to thermal oxidation, where oxygen (or another oxidant such as the water molecule) is

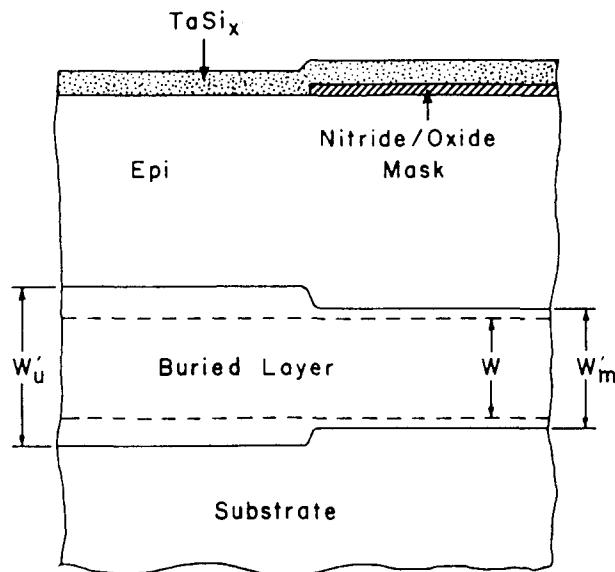


Fig. 10. Structures used in the study of point defect generation by silicidation. The conductivity-type of both the epitaxial layer and the substrate are in all cases opposite to the buried layer (B or Sb), so as to allow junction delineation by staining. W and W' are the junction widths before and after silicidation, respectively, and the subscripts m and u denote areas masked and unmasked by the nitride pattern. Enhanced junction movements are observed for both antimony and boron buried layers (taken from Hu [204]).

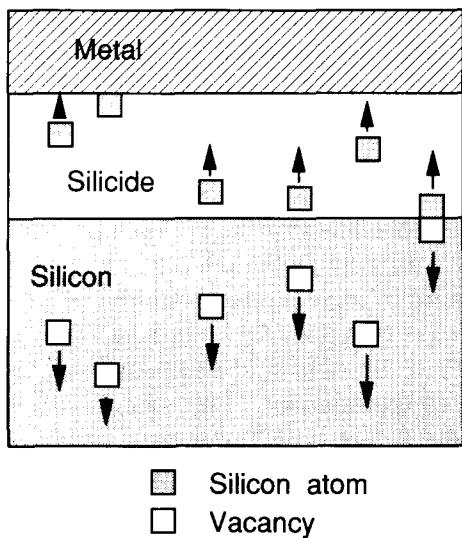


Fig. 11. Schematic illustration of how silicidation generates vacancies in the silicon. On the right-most side of the figure, a surface silicon atom has moved a step into the silicide, leaving a vacancy behind. That vacancy will subsequently move into the bulk of the silicon following the example of the other vacancies shown to its left. The silicidation is assumed to occur by a mechanism in which silicon atoms migrate across the silicide to react with the metal. The opposite mechanism of metal atoms migrating across the silicide to react with the silicon will not produce vacancies, but may produce silicon interstitials. A kink silicidation mechanism will produce neither vacancies nor interstitials.

transported from the oxide surface to the oxide–silicon interface, and oxidizes most of the silicon atoms there. But some silicon atoms are left unoxidized and are ejected as interstitials which then flood the bulk of the substrate, causing a local interstitial supersaturation.

Hu [204] also studied the effect of an intervening layer of 150 nm polycrystalline silicon film between the silicide and the silicon substrate, and found that the enhanced part of dopant diffusion disappeared in presence of the intervening layer. The explanation is simply that the grain boundaries of polycrystalline silicon are effective sinks for excess vacancies. (Polycrystalline silicon films are also effective sinks for excess self-interstitials, as will be discussed in Section 7.)

In a related study, Wen et al. [205] found that the residual extended defects due to end-of-range ion implantation damage can be totally eliminated by titanium silicidation. The end-of-range extended defects are usually obtained after annealing–recrystallization of an amorphized layer, and are interstitial in nature. Pre-amorphization is rather commonly used in order to reduce channeling of shallow implant of light ions such as boron. Earlier, Ajmera and Rozgonyi [206] found that shallow end-of-range dislocation loops can be eliminated by rapid thermal annealing at 1050 °C in a reducing ambient, provided the amorphous–crystalline interface is less than 70 nm deep. Deeper end-of-range defects cannot be removed by annealing under these conditions. In the experiment of Wen et al. [205], the end-of-range defects were produced by pre-amorphization by Ge^+ implantation at 85 keV with a dose of $1 \times 10^{15} \text{ cm}^{-2}$, and subsequent recrystallization at 550 °C following boron implantation. It was further followed by 10 s rapid thermal annealing at 1050 °C. The silicidation started with the deposition of a 30 nm Ti film. It was then followed by various rapid thermal annealing at different temperatures. It was found that defects were observed at densities of the order of 10^5 cm^{-1} for one-step silicidation at 600 °C and at 800 °C. No defects were observed for one-step silicidation at 1000 °C. Wen et al. explained

the results of the annihilation of the interstitial end-of-range defects as due to excess vacancies generated by silicidation.

Quite similarly, Maex et al. found that the formation of TiSi_2 and CoSi_2 led to the reduction of interstitial-type defects introduced by ion implantation [207]. An interesting aspect of the result of Maex et al. is that a significant reduction of the defects was obtained for silicidation at 700 °C, in contrast to the result of Wen et al. Ohdomari et al. [208] reported that the formation of Pd_2Si also led to the reduction of defects introduced by As implantation.

More recently, Honeycutt and Rozgonyi [209] investigated enhanced diffusion of buried Sb layers under the formation of TiSi_2 and CoSi_2 at 700 °C, one of the conditions in the experiments of Maex et al. [207]. They observed not only an enhanced diffusion from which they estimated an astonishing vacancy supersaturation of 10^7 , but also an asymmetrically broadened Sb concentration profile similar to those of Wittmer et al. [198]. As did Wittmer et al., they used SIMS to analyze the Sb profiles. It is hence likely that their results are spurious, as their measured profiles may have been smeared toward the substrate surface by a sputtering-induced surface roughening, as already discussed.

6. Oxygen precipitation and nonequilibrium point defects

6.1. Importance of oxygen and its precipitation in silicon

Silicon crystals grown by the Czochralski method from silica crucibles typically contain $\approx 10^{18}$ atoms cm^{-3} of dissolved oxygen. Interest in the subject of oxygen in silicon first appeared in the mid-1950s, especially related to the topic of thermal donors. But the interest in the subject, which had not been connected to practical matters, had almost disappeared by the end of the 1960s. Then, around 1970, it was discovered that making integrated circuits on silicon wafers from the seed-end portion of a Czochralski-grown crystal tended to give a higher yield than on those from the tail-end portion. This brought forth two urgent problems: (a) should we wastefully discard the tail portion of a silicon crystal, and (b) what caused this to happen? Without an understanding of what was happening, one could not hope to be in control of fabrication yield. Among many clues sought to explain this finding was the known fact that Czochralski silicon crystals contain oxygen, and that the oxygen concentration is highest in the seed-end portion of a crystal, decreasing towards the tail end. The interest in the subject of oxygen in silicon was thus rekindled and it rapidly intensified. A number of possible effects due to oxygen were soon postulated, and were experimentally demonstrated in the subsequent several years. One important effect is the improvement of the high-temperature mechanical strength of silicon [210–214], which helps to reduce the chances of dislocation generation during thermal processing.

Another important effect is that precipitates of oxygen, if confined to well below the surface region, where devices are located, can serve as internal getters of metallic impurities [124,215–218]. In fact, the tendency of metallic impurities, such as copper, to gather in regions rich in oxygen or its precipitates had already been observed in the early 1960s [219,220], but had not been specifically exploited until the mid-1970s.

The kinetics of the precipitation of oxygen is usually determined experimentally by following thermal processing with the evolution of the infrared absorption band at 1107 cm^{-1} , characteristic of interstitially dissolved oxygen in silicon. Oxygen precipitates also exhibit characteristic infrared absorption bands; but the absorbances of these bands are

substantially weaker. Furthermore, according to Hu's theory [221], which was later confirmed experimentally [222,223], the absorption bands vary according to the shape of the oxygen precipitates, due to a polarization effect. These absorption bands, experimental as well as calculated [221], are not very sharp, and may be difficult to measure very accurately.

6.2. Interstitial emission from oxygen precipitation

In 1977, Hu [211] reported as an incidental note an accidental discovery that oxygen precipitation is retarded when silicon specimens are annealed in oxidizing ambients relative to neutral ambients. He subsequently studied this phenomenon in more detail [224]. It had by that time already been established [2] that thermal oxidation of silicon causes excess self-interstitials to be injected into the silicon substrate. Taking this together with the well-known fact that the agglomeration of interstitial oxygen atoms into SiO_2 precipitates produces local strain, Hu proposed [224] that, to reduce strain energy, oxygen precipitation is accompanied by emission of silicon self-interstitials, and that a supersaturation of self-interstitials due to thermal oxidation tends to suppress the interstitial emission, thus retarding the oxygen precipitation. He described the oxygen precipitation process by the following equation:



where the subscript I indicates an interstitial configuration. An excess of silicon self-interstitials, such as from thermal oxidation, causes the above reaction to shift to the left, consequently retarding the precipitation. Various experimental observations have since confirmed the precipitation retardation effect of oxidizing ambients [225–233].

The finding that oxygen precipitation actually causes the generation of self-interstitials was also reported in the same year by Hu [27] in a study in which pre-existing extrinsic stacking faults, which normally would shrink when annealed in neutral ambients, were observed to grow when oxygen precipitation was on course. Since then, the oxygen precipitation-induced emission of self-interstitials has abundantly been confirmed, both through the observations of the growth of extrinsic stacking faults [27–29,233], and through the enhanced diffusion of phosphorus and the retarded diffusion of antimony [234]. Taken together, these two phenomena have established the validity of the precipitation model as described by Eq. (45).

The above reaction equation for oxygen precipitation is written because we do not know the precise reaction scheme, and hence an unknown coefficient x is assigned to the molecularity of oxygen atoms participating in the reaction. However, from experimental observations, the precipitation process appears to be effectively unimolecular, and the rate of precipitation is directly proportional to the interstitial oxygen concentration, $C_{\text{O}i}$, rather than proportional to $C_{\text{O}i}^2$. Here, SiO_2 , a separate phase, is merely a sink, or a source, of fixed strength. Hence, we may rewrite the precipitation process as



y in the above reaction equation has not been established experimentally, and can be expected to vary from 0 to 1.25, depending on many factors, especially the annealing temperature. From the a priori assumption that one unit volume of silicon is turned into 2.25 unit volumes of stoichiometric SiO_2 upon the precipitation of oxygen, it may be tempting to assign a value of $y=1.25$ in order to eliminate the strain completely. This is to say that,

for each oxygen atom precipitated, there will be 0.62 silicon self-interstitials emitted. There are problems with this assumption, however. In the first place, there is no experimental evidence or theoretical ground to suggest that self-interstitials will be emitted, at any temperatures, to the extent required for a complete relief of strain, if at all. We will discuss this topic in more detail subsequently. In the second place, the oxygen precipitate in silicon does not have an exact stoichiometry of SiO_2 . The exact stoichiometry has been a question for a long time. Recently, Borghesi et al. [235] have determined the stoichiometry of oxygen precipitate platelets in silicon to be $\text{SiO}_{1.8}$. They came to this conclusion from their calculations of the infrared absorption band by adopting Hu's model [221] of polarization of plate-shaped precipitates within the framework of the effective medium theory.

6.3. Precipitate morphology and interstitial emission

A salient feature of the morphology of oxygen precipitates is its dependence on the annealing temperature. Approximately three temperature regimes may be distinguished. Precipitates take the shape of ribbons [236–238] up to about 800 °C. From about 800 °C to about 1050 °C platelet precipitates dominate and often are accompanied by prismatic dislocation loops [237,239–246]. Above about 1050 °C, precipitates predominantly take the shape of regular or truncated octahedra [237,245–250]. However, there is considerable overlap of these regimes. Platelet precipitates have been observed at annealing temperatures as low as 600 °C [238] and as high as 1150 °C [246]. The platelet and ribbon precipitates have the {100} habit plane, while the octahedral precipitate has {111} bounding faces. The precipitates being amorphous, their habit planes are determined by the elastic and the interfacial anisotropies of the host matrix silicon. Where the strain energy matters (when there is little or no interstitial emission), the precipitate platelet adopts the habit plane {100} because the platelet is under a normal compressive stress and the ⟨100⟩ direction has the minimum Young's modulus of elasticity in silicon [251,252] (as, with rare exceptions, is generally the case in cubic crystals [253]). Where the strain energy no longer matters (when there is almost complete elimination of strain by sufficient emission of self-interstitials) and the interface energy becomes the next on the energy minimization agenda, the precipitate chooses the {111} polyhedral face because it has the minimum interface energy with an amorphous phase of SiO_2 . (We really do not know the energy of the Si– SiO_2 interface; but it seems reasonable to expect it to follow the trend of the surface energy anisotropy of silicon, which exhibits a minimum on {111} [254].)

At low temperatures, however, a dense crystalline polymorph of SiO_2 , coesite, has been reported [236–238]. The coesite polymorph can exist only under extremely high pressures (≥ 25 kbar) [255]. Its occurrence is testimony both to the extremely high local stress created by the formation of the precipitate, and to the negligible emission of self-interstitials at about 650 °C. It is also clear that at that low temperature, punching of prismatic dislocation loops is also rather ineffective.

The shape of a precipitate is determined by a number of contributing factors to the total free energy of precipitation. The strain energy is largest for the spherical shape, and smallest for the plate shape [125,126]. The minimum strain energy is attained in a disk or ribbon shape with the smallest possible thickness (say, of the order of the atomic dimension). But then the interface energy becomes largest. Hu [256] has derived a least-energy relationship, as given below, of the size-to-thickness aspect ratio as a function of the size of platelet:

$$\frac{r}{\lambda} = \frac{(r/h\delta)^2}{\ln(r/h\delta) + \ln(8/e^2)} \quad (47)$$

where r and h are respectively the radius and the thickness of the disk, and δ is the linear misfit between the silicon matrix and the SiO_2 precipitate. $\lambda = 8\pi\gamma(1-\nu)/3\mu$, where γ is the interface energy, ν is Poisson's ratio and μ is the shear modulus. Hu's derivation is based on the assumption of a fixed elastic misfit, amounting to the assumption of no emission of self-interstitials. He found that his theoretical results agree reasonably well (about a factor of 2) with the experimental data of Patrick et al. [243]. This could mean, though not so definitively, that in fact there is not enough emission of self-interstitials during the formation of disk-shaped oxygen precipitates. Otherwise it would be difficult to explain why the precipitates would adopt a disk shape, a shape of extravagant interface energy. We thus emphasize that one cannot decide a priori that the number of self-interstitials emitted per one oxygen atom precipitated is the ideal number 0.62, or for that matter, any other number. It would seem obvious that while free energy consideration may favor interstitial emission, the rate of emission is controlled by kinetic factors.

The fact that not enough interstitials are emitted is also evident from the strain contrast and prismatic dislocation loops surrounding disk-shaped precipitates. A good example of punching of prismatic dislocation loops is shown in the TEM picture in Fig. 12.

At higher temperatures, say ≥ 1050 °C, the process of self-interstitial emission becomes quite efficient in creating free volumes needed for substantially strain-free precipitates. Thus, at 1200 °C, with strain practically eliminated, oxygen precipitates are observed to take up the shape of least interface energy, the truncated octahedron [237,245–250], but not the shape of least surface area, the sphere. A TEM example of Ponce et al. [249] is shown in Fig. 13. The transmission microscopic lattice images of such precipitates reveal no detectable lattice strain [237,249]. The strain-free state means that, assuming an exact stoichiometry of SiO_2 , there are 1.25 self-interstitials emitted per silicon atom oxidized, or

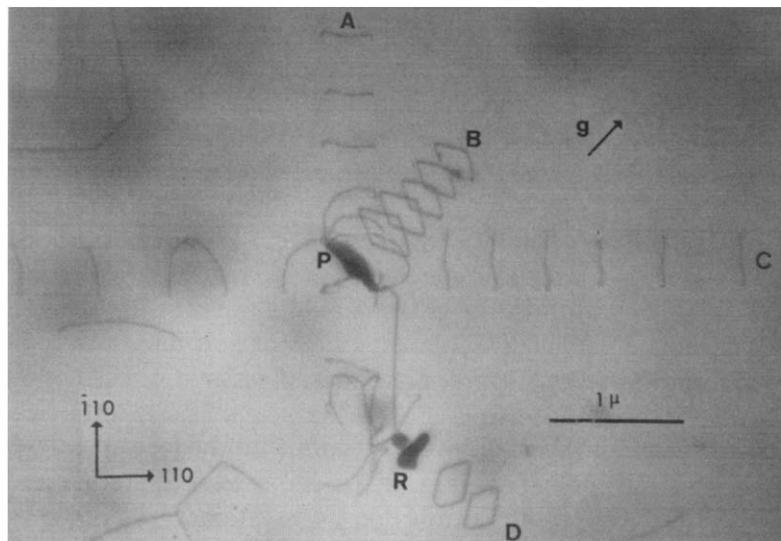


Fig. 12. Prismatic punching of dislocation loops by SiO_2 precipitate platelets in a silicon matrix. The P platelet has a (010) habit plane and the R platelet a (100) habit plane. Prismatic loops group A: $(a/2)\bar{1}\bar{1}0$; group B: $(a/2)[01\bar{1}](0\bar{1}\bar{1})$; etc. (taken from Tice and Tan [240]).

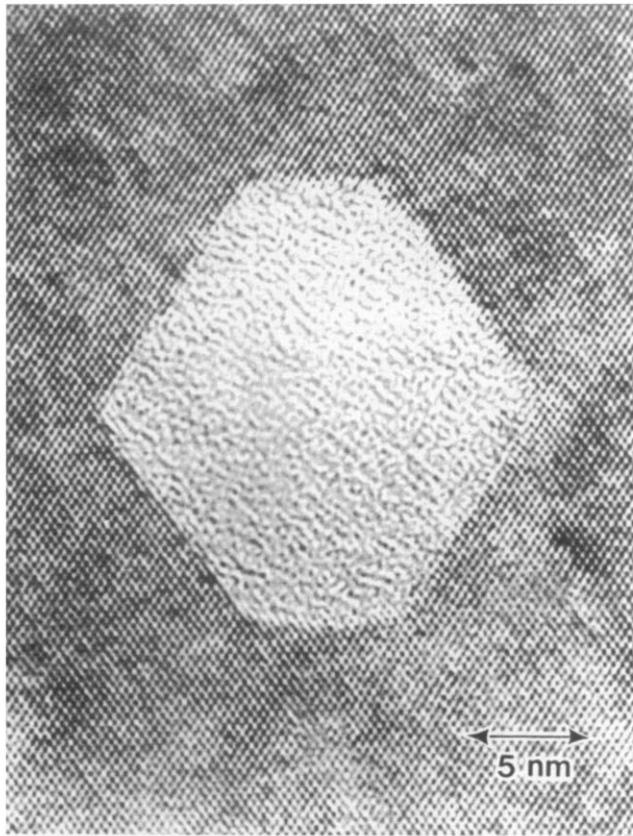


Fig. 13. A high-resolution lattice image of a polyhedral SiO_2 precipitate produced in silicon after 1175 °C, 64 h anneal. Notice that the precipitate is amorphous and the silicon lattice is completely free of strain (taken from Ponce et al. [249]).

per two oxygen atoms precipitated. We may assume this ratio for oxygen precipitation at high temperatures. We will use this value in a subsequent section to calculate the concentration of self-interstitials during oxygen precipitation at 1200 °C annealing temperature. A small deviation from this value will not significantly affect our estimation of the excess self-interstitial concentration during oxygen precipitation.

6.4. Critical radius of precipitate nucleus

In order to applying Hu's idea (Eq. (45)) to the nucleation of oxygen precipitates quantitatively, Gösele and Tan [257] assume $x=0.5$ and $y=1$ for the coefficients in Eq. (45). They obtained for a spherical shaped precipitate the critical radius of precipitate nucleus of

$$r_{\text{crit}} = \frac{\gamma \nu_{\text{SiO}_2}}{kT \ln \left[\left(\frac{C_O}{C_O^*} \right) \left(\frac{C_I^*}{C_I} \right)^{1/2} \right]} \quad (48)$$

where ν_{SiO_2} is the molecular volume of SiO_2 , and C_O is the oxygen concentration. As we have already discussed, it is difficult to assess the correct value of the ratio of the number of self-interstitials emitted to the number of oxygen atoms precipitated, a ratio that is an important element in their expression.

6.5. Shrinkage and growth of stacking faults in oxygen precipitation

The kinetics of oxygen precipitation can be followed by monitoring the change of interstitial oxygen concentration, measured by the infrared absorption band at 1107 cm^{-1} [27,224]. Fig. 14 shows an example of the change of the interstitial oxygen concentration with annealing time as measured by the 1107 cm^{-1} absorption band [27]. The corresponding kinetics of the shrinkage and growth of pre-existing OSF is given by curve No. 1 of Fig. 15. There is some hint of a brief incubation period; but this cannot be ascertained because of considerable data scatter. An incubation period, however, is not a prerequisite for the initial OSF shrinkage. It takes time for the concentration of excess self-interstitials to build up. Until a critical concentration, C' , is reached, the stacking faults will shrink under the driving force of the faulting energy γ . C' , the concentration of self-interstitials in equilibrium with a stacking fault, is given by [177]

$$C' = C^* \exp \left[\frac{\Omega}{bkT} \left(\gamma + \frac{Gb^2}{2\pi r(1-\nu)} \left(\ln \frac{8r}{b} - 1 \right) \right) \right] \quad (49)$$

In the above expression, within the square brackets the first term, γ , is the faulting energy per unit area of the stacking fault, and the second term represents the strain energy due to the stacking fault. G is the shear modulus, ν Poisson's ratio, and Ω the atomic volume of silicon. b is the Burgers vector of the stacking fault, about 3.14 \AA . For sufficiently large stacking faults, the strain energy term becomes negligible compared with the faulting energy,

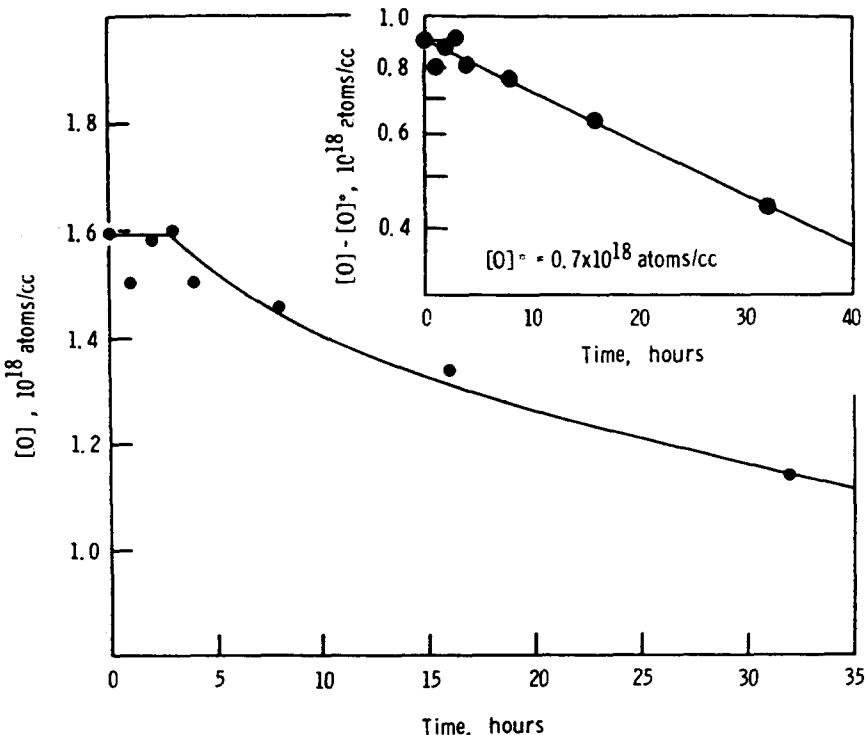


Fig. 14. The kinetics of oxygen precipitation measured by infrared spectrometry of the decrease of interstitial oxygen concentration remaining in solution. This evolution corresponds to the shrinkage-growth of stacking faults of curve 1 in Fig. 15 (taken from Hu [27]).

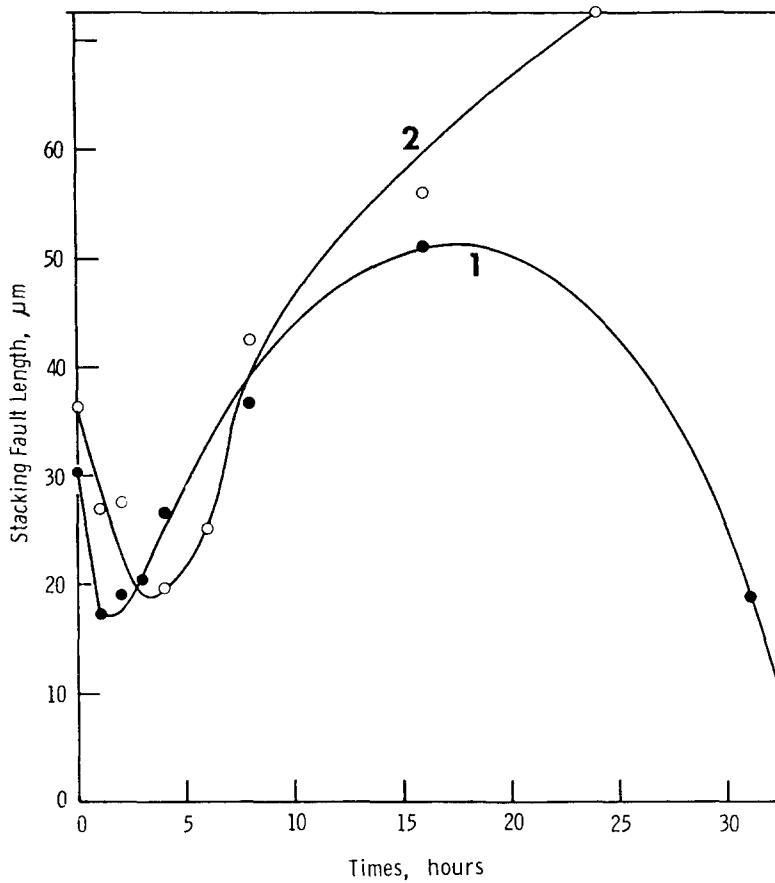


Fig. 15. The shrinkage and growth of pre-existing stacking faults in Czochralski silicon wafers containing oxygen during annealing in argon at 1200 °C. The two curves are from two sets of samples, attesting to the heterogeneity in wafers (taken from Hu [27]).

and Eq. (49) reduces to

$$C' = C^* \exp\left(\frac{\Omega\gamma}{bkT}\right) \text{ for } \gamma \gg \frac{Gb^2}{2\pi\gamma(1-\nu)} \quad (50)$$

For the extrinsic stacking fault in silicon, γ is about 70 ergs cm⁻² [258], or about 0.026 eV per atom. Thus, the above expression should be good for $r \geq 1 \mu\text{m}$.

The rate of stacking fault growth as a function of self-interstitial supersaturation may be formulated as diffusion-controlled, or reaction-controlled, or both [177]. The fact that the growth rate of multiple, closely spaced OSF is not significantly smaller than that of an isolated OSF is suggestive of a reaction-limited process. But if the reaction barrier ΔH is sufficiently small, say $\geq 4kT$, then the rate of OSF growth shrinkage may be expressed approximately as [177]

$$\begin{aligned} \frac{dr}{dt} &= \frac{\pi\Omega r_c^2}{bd_a^2} e^{-\Delta H/kT} D_I C_I^* \left[\left(\frac{C_I}{C_I^*} - 1 \right) - (e^{\Omega\gamma/bkT} - 1) \right] \\ &\approx k_{SF}^I \left[\left(\frac{C_I}{C_I^*} - 1 \right) - \frac{\Omega\gamma}{bkT} \right] \end{aligned} \quad (51)$$

where r_c is the core radius of the torus representing the sink surface surrounding the rim of the stacking fault, approximately equal to the Burgers vector b , and d_a is the interatomic distance. In the above equation, the lumped constant k_{SF}^I represents the entire quantity between the equals sign and the first left square bracket. We have also taken a linear approximation for the faulting energy term, justifiably because $\Omega\gamma/bkT \approx 0.25 \ll 1$ under typical conditions. If the kinetics of stacking fault growth is diffusion-limited rather than reaction-limited, then the quantity outside the brackets is no longer a constant, but contains a factor of $\ln(8r/r_c)$ [177,259,260], and hence cannot be represented by the constant k_{SF} . Experimental observations [172–175,195] have shown that, in inert ambients, pre-existing OSF shrink linearly with time. This may be taken as yet another indication that the kinetics of stacking fault shrink growth is reaction-limited rather than diffusion-limited.

If the stacking fault kinetics is diffusion-limited, the contribution of vacancies should also be considered. Fahey et al. have estimated the interstitialcy component of silicon self-diffusion to be 0.3–0.4 at 1050 °C. [47]. In the presence of a significant self-interstitial supersaturation, the vacancy concentration is expected to decrease below C_v^* , so that the contribution of vacancy flow to the kinetics of OSF growth may be ignored in comparison to the self-interstitial flow.

The formulation can be simply extended to include the consideration of both the interstitial and the vacancy contribution, as was done by Gösele et al. Their formulation is based on the assumption of diffusion-limited kinetics [259,260]:

$$\frac{dr}{dt} = -\alpha_{eff} b^2 \left[(D_I C_I^* + D_V C_V^*) \frac{\Omega\gamma}{bkT} - D_I C_I^* \left(\frac{C_I}{C_I^*} - 1 \right) + D_V C_V^* \left(\frac{C_V}{C_V^*} - 1 \right) \right] \quad (52)$$

where α_{eff} is a geometrical factor for the diffusion-limited case, and is given by $2\pi/\ln(8r/r_c)$, which is a weak function of r . (Note that the expression given in Eq. (52) has been modified from the one originally given by [259,260], which is dimensionally inconsistent. The dimension of the faulting energy γ is also different as used here and in [259,260].) This formulation is especially appropriate in the presence of vacancy supersaturation. Under the usual condition of stacking fault growth, in which the concentration of self-interstitials is supersaturated while the concentration of vacancies is undersaturated, self-diffusion of silicon is mediated predominantly by the self-interstitial, and the additional vacancy term in the above formulation becomes unnecessary.

C_I and C_V in either Eq. (51) or Eq. (52) are, in general, functions of time, depending on the kinetics of surface and bulk processes of point defect generation taking place. When no such process is occurring, C_I and C_V are at their equilibrium values, and the extended defects would shrink at a constant rate, at least before the size of extended defects falls below ≈ 100 nm.

6.6. Kinetics of nonequilibrium interstitials during oxygen precipitation

We now analyze the kinetics of oxygen precipitation for the data of Fig. 14. Since the exponential fitting is seen to be quite good, and is in accord with most of the previous investigations [261,262], as well as Ham's theory of diffusion-limited precipitation kinetics [263], we will assume the decay of the supersaturation of the interstitial oxygen to be given by

$$C_O - C_O^* = A e^{-t/\tau_{gen}} \quad (53)$$

The rate of self-interstitial generation, from Eq. (46), is then given by

$$r_{gen} = -\frac{y}{2} \frac{dC_O}{dt} = \frac{Ay}{2\tau_{gen}} e^{-t/\tau_{gen}} \quad (54)$$

The interstitial oxygen concentration is of the order of 10^{18} cm^{-3} , and the self-interstitials generated by the precipitation process are approximately 60% of that, well over two orders of magnitude larger than the concentration of vacancies in stock in the bulk. It is clear, then, that the annihilation of the excess interstitials must rely on their diffusion to the surface, perhaps also aided by vacancies diffusing from the surface. Furthermore, it has been experimentally observed [27] that the magnitude of interstitial supersaturation, as indicated by the growth or shrinkage of pre-existing stacking faults, is strongly dependent on the substrate surface covering, whether it is Si_3N_4 , SiO_2 or bare silicon. This phenomenon would suggest that the surface reaction controls the rate of annihilation of the excess self-interstitials. We may therefore express the rate of interstitial annihilation as a linear process

$$r_{ann} = \frac{C_I - C_I^*}{\tau_{ann}} \quad (55)$$

where τ_{ann} is the time constant of surface annihilation. We can now write the rate of change of the self-interstitial concentration as

$$\frac{d(C_I - C_I^*)}{dt} = r_{gen} - r_{ann} = \frac{Ay}{2\tau_{gen}} e^{-t/\tau_{gen}} - \frac{C_I - C_I^*}{\tau_{ann}} \quad (56)$$

for which the solution is

$$C_I - C_I^* = \frac{Ay\tau_{ann}}{2(\tau_{gen} - \tau_{ann})} (e^{-t/\tau_{gen}} - e^{-t/\tau_{ann}}) \quad (57)$$

C_I peaks at

$$t = \frac{\tau_{gen}\tau_{ann}}{\tau_{gen} - \tau_{ann}} \ln \frac{\tau_{gen}}{\tau_{ann}} \quad (58)$$

corresponding to an inflection point on the SF shrinkage growth curve. With $y=1.25$, we obtain from Fig. 14 $A=0.9 \times 10^{18} \text{ atoms cm}^{-3}$, and $\tau_{gen}=44.3 \text{ h}$. Then the non-trivial solution to Eq. (58) gives $\tau_{ann}=0.87 \text{ h}$ at the time corresponding to the point of inflection. With these values, the evolution of $C_I - C_I^*$ can be obtained from Eq. (57), and is shown in Fig. 16. At $dr/dt=0$

$$C_I = C_I^* \exp\left(\frac{\Omega\gamma}{bkT}\right) \quad (59)$$

from which we obtain the equilibrium concentration of self-interstitials:

$$C_I^* = \frac{bkT}{\Omega\gamma} (C_I - C_I^*)_{dr/dt=0} = 3.0 \times 10^{16} \text{ cm}^{-3} \text{ at } 1200 \text{ }^\circ\text{C} \quad (60)$$

This value is about two orders of magnitude larger than estimates made from the "kick-out" mechanism of diffusion of transition elements in silicon [79,81,86].

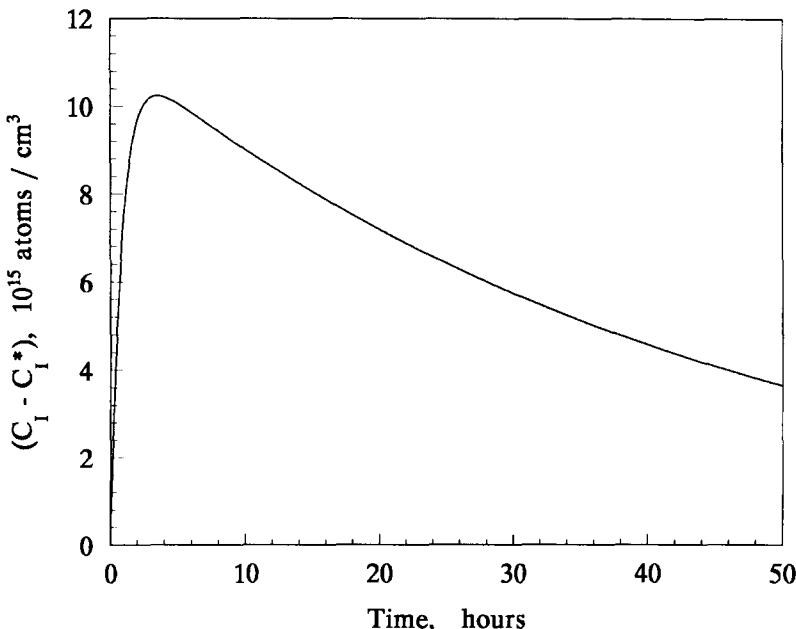


Fig. 16. The evolution of excess self-interstitials during the precipitation of oxygen, corresponding to the kinetics of oxygen precipitation in Fig. 14 (taken from Hu [27]).

7. Effect of surface films on nonequilibrium point defects

It was found [211,224] that oxygen precipitation is retarded in silicon substrates covered by amorphous films of silicon nitride compared with those covered by thermal SiO_2 or with bare surfaces. It was also found that pre-existing oxidation stacking faults would grow rather rapidly and continuing for a long time in the course of oxygen precipitation [27]. Hu explained his findings with the model that the $\text{Si}-\text{Si}_3\text{N}_4$ interface is a poor sink for point defects. The concentration of the excess self-interstitials generated by oxygen precipitation builds up quickly in a nitride-covered substrate. The supersaturation of self-interstitials in turn retards the precipitation process itself, and at the same time feeds the growth of extrinsic stacking faults. This fits well within the framework of his model of OED [2], in which the excess self-interstitials generated by thermal oxidation would be annihilated at the $\text{Si}-\text{SiO}_2$ interface by a regrowth process. The $\text{Si}-\text{Si}_3\text{N}_4$ interface makes such a regrowth process much more difficult. An alternative model to the surface regrowth process of a surface sink is that excess self-interstitials can diffuse through an SiO_2 film much more readily than through an Si_3N_4 film. The two main problems with this alternative are that it cannot explain the effect of surface orientation on OED and OSF and that the effect of a bare surface requires a different explanation.

There are surface coverings that are better sinks for excess point defects than the bare surface or the thermal SiO_2 . It has been reported [264,265] that polycrystalline silicon films on silicon substrates increase the rate of oxygen precipitation. This is easily understood because a polycrystalline silicon film has an area of grain boundaries many times more than a smooth silicon surface. In fact, in a patent by Hu [266], it has been shown that backside polycrystalline silicon films prevent the formation of oxidation stacking faults, and improve minority carrier lifetime through its excellent gettering property. The gettering

efficiency of polycrystalline silicon films has been confirmed by many subsequent investigations [267,268].

8. Vacancy concentration in the presence of interstitial supersaturation

8.1. Commonly assumed relationship between C_V and C_I

Due to vacancy–interstitial annihilation, a supersaturation of, say, self-interstitials from, say, surface injection from thermal oxidation will lead to a depletion of vacancies. In modeling a dual mechanism of diffusion, it is necessary to know the concentrations of both the vacancy and the self-interstitial. It has been suggested [30,269,270] that when the concentration of self-interstitials is in supersaturation, the concentration of vacancies would observe the relationship

$$C_V C_I = C_{V^*} C_{I^*} \quad (61)$$

under “dynamic equilibrium”. This assumed relationship was first used by Mizuo and Higuchi [30] to analyze their observation of oxidation-retarded diffusion of antimony. It has since been used by many others for similar analyses of vacancy concentration decrease in the presence of interstitial injection. This assumption makes things much easier by removing the necessity to solve complicated coupled nonlinear equations of diffusion–reaction. For example, the expression of the effective diffusivity [271]

$$\frac{D_A}{D_{A^*}} = (1-f) \frac{C_V}{C_{V^*}} + f \frac{C_I}{C_{I^*}} \quad (62)$$

may be written simply as [30,270,272]

$$\frac{D_A}{D_{A^*}} = (1-f) \frac{C_I^*}{C_I} + f \frac{C_I}{C_I^*} \quad (63)$$

Gösele et al. [127] have estimated, for diffusion-limited recombination, a time of <0.05 s to reach dynamic equilibrium, by which they meant the relationship of $C_V C_I = C_{V^*} C_{I^*}$. But the real issue is not the time required to reach dynamic equilibrium; it is the validity of that relationship under a continuous injection of point defects. The recombination time constant would be an appropriate issue when dealing with the case of injection of an instantaneous dose of point defects. In the following sections, we will analyze the non-equilibrium concentrations of point defects in the presence of bulk and surface generations. The case of bulk generation is easy to understand, and is analogous to the case of the generation of electron–hole pairs by, for example, photoexcitation, where the dynamic equilibrium relationship between n and p , as is well known, is $np > n_i p_i$. The case of surface injection is much more complicated. It requires the solution of coupled nonlinear diffusion–reaction equations for the point defects. But it should be apparent that, within the “diffusion length” from the point of defect injection, the “dynamic equilibrium” relationship is not $C_V C_I = C_{V^*} C_{I^*}$, but $C_V C_I > C_{V^*} C_{I^*}$. The issue of practical importance is the magnitude of this length compared to the depth of the region in which dopant diffusion takes place.

8.2. C_V – C_I relationship in the presence of a nonthermal bulk source

Despite the dominance of the Schottky processes in finite crystals in the determination of equilibrium thermal point defects at the lowest possible free energy, the bulk Frenkel

process of generation and recombination of vacancy-interstitial pairs is occurring all the time in the following mass action:



where B denotes the host lattice atom. (We will denote an impurity atom as A, a vacancy as V, a self-interstitial as I and a point defect as X, which may be V or I where a distinction is not required.) The law of mass action gives at equilibrium

$$C_V C_I = \frac{k_g C_B}{k_r} = \text{constant} \quad (65)$$

because the concentration of host lattice atoms, C_B , is practically a constant (C_A , C_V and C_I are all negligible compared to C_B , so that $C_B \approx C_S$, the concentration of lattice sites). Furthermore, this constant is equal to $C_V^* C_I^*$, as already discussed above.

When in addition to thermal excitation there are other sources such as X-ray, electron or ion irradiations, the balance of recombination and generation becomes given by

$$k_r C_V C_I = k_g C_B + G_b \quad (66)$$

where G_b is the bulk generation rate of Frenkel pairs due to nonthermal sources. It follows that

$$C_V C_I = C_V^* C_I^* + \frac{G_b}{k_r} \quad (67)$$

We see that in the condition of the so-called "dynamic equilibrium", the relationship is really $C_V C_I > C_V^* C_I^*$.

For ion implantation, G_b will not be uniform, so that diffusion in addition to excess vacancies and self-interstitials will take place. Due to a likely disparity between their diffusivities, the concentration profiles of the vacancy and the self-interstitial will not be the same. If the lifetime of the vacancy and the interstitial is small compared to the irradiation time, a steady state may be achieved within the irradiation time span. In this case, we need only to solve the following coupled equations:

$$D_V \frac{\partial^2 C_V}{\partial x^2} - k_r (C_V C_I - C_V^* C_I^*) + G_b(x) = 0 \quad (68)$$

$$D_I \frac{\partial^2 C_I}{\partial x^2} - k_r (C_V C_I - C_V^* C_I^*) + G_b(x) = 0 \quad (69)$$

Typically in ion implantation or in proton irradiation, $G_b(x)$ is a rather broad distribution. Then, within the entire regime of this distribution, we must have $C_V C_I > C_V^* C_I^*$. This is illustrated in Fig. 17(a). If $G_b(x)$ is a delta function, then it is equivalent to a generating "internal surface". The result is schematically shown in Fig. 17(b). It is clear that, in this case, within the diffusion length of the point defects, we must have $C_V C_I > C_V^* C_I^*$.

8.3. Vacancy concentration under surface injection of interstitials

In a finite crystal, a prominent source of point defects comes from surface reaction, from which point defects of a particular species are injected into the crystal. The situation here is more complicated, but the product $C_V C_I$ under "dynamic equilibrium" will not be equal to $C_V^* C_I^*$ as claimed by various authors. This is a very important case that occurs

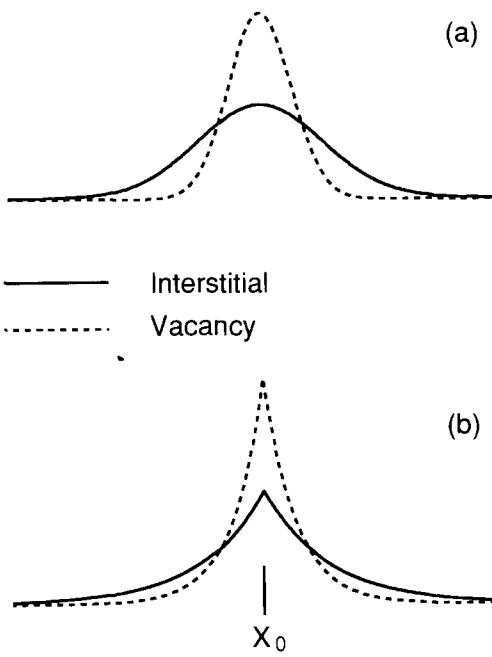


Fig. 17. Schematic illustration of the concentration profiles of vacancies and self-interstitials generated by an internal source of Frenkel pairs such as ion implantation. (a) shows the case in which the source distribution is broader than the diffusion length of the generated excess point defects. Within the whole range of source distribution, $C_V C_I > C_V^* C_I^*$. (b) shows the case in which the source has a delta function distribution located at x_0 . Here, within the diffusion length of excess point defects, $C_V C_I > C_V^* C_I^*$. It is quite possible that the diffusion length may be larger than the width of a real distributed source. Such steady state distributions are in fact in "dynamic equilibrium".

commonly in silicon device fabrications. A simple insight can be gained by simply considering that, in the surface region within the diffusion length of the injected point defects, the situation is similar to that of bulk generation source.

Consider the condition of the injection of self-interstitials from the surface. Consider the surface capable of acting as an unlimited source of vacancies. It is easy to see that a smaller sum of $\Delta G_V + \Delta G_I$ is obtained by keeping $C_V = C_V^*$ than by making $C_V C_I = C_V^* C_I^*$. That is, near the surface region, the free energy driving force will always attempt to drive the vacancy concentration to the value of C_V^* , regardless of how far C_I departs from C_I^* . But the kinetics of surface generation-annihilation of vacancies, and the kinetics of diffusion of these vacancies to the bulk, will determine at what distance from the surface and how close C_V will approach C_V^* , instead of $C_V^* C_I^*/C_I$. Bulk annihilation between vacancies and self-interstitials will cause a vacancy depletion, which will incur a vacancy flux from the surface. To supply this vacancy flux requires a finite value of $C_V^* - C_V$ as a driving force for any finite rate of vacancy generation at the surface. Clearly, then, one must expect $C_V < C_V^*$ even at the surface, unless the kinetic coefficient of surface generation-annihilation is very large compared with that of the bulk. To determine rigorously how far C_V will depart from C_V^* requires the solution of coupled continuity equations of vacancy and interstitial diffusion with bulk reaction. An exact solution for such a system of continuity equations has been obtained by Hu for the steady state condition [273].

8.3.1. Steady state analysis

For simplicity of analysis, consider the case of a constant interstitial injection rate at the surface, G_S . In this case, a steady state is approached at sufficiently long time in a

substrate of finite thickness h . It may be noted that a steady state can still be attained at sufficiently large time, even when G_s is a function of time, but of order t^m , where $|m| < 1/2$. The continuity equations for the self-interstitial and the vacancy are given by

$$\frac{\partial C_I}{\partial t} = D_I \frac{\partial^2 C_I}{\partial x^2} - k_r(C_V C_I - C_V^* C_I^*) \quad (70)$$

$$\frac{\partial C_V}{\partial t} = D_V \frac{\partial^2 C_V}{\partial x^2} - k_r(C_V C_I - C_V^* C_I^*) \quad (71)$$

where k_r is the rate constant of vacancy-interstitial recombination. In the steady state, $\partial C_I / \partial t = 0$ and $\partial C_V / \partial t = 0$, and the above equations reduce to the following set of simultaneous second-order nonlinear ordinary differential equations:

$$D_I \frac{\partial^2 C_I}{\partial x^2} - k_r(C_V C_I - C_V^* C_I^*) = 0 \quad (72)$$

$$D_V \frac{\partial^2 C_V}{\partial x^2} - k_r(C_V C_I - C_V^* C_I^*) = 0 \quad (73)$$

The boundary conditions are

$$-D_I \frac{dC_I}{dx} + k_I(C_I - C_I^*) = G_s \quad \text{at } x=0 \quad (74)$$

$$-D_V \frac{dC_V}{dx} + k_V(C_V - C_V^*) = 0 \quad \text{at } x=0 \quad (75)$$

$$\frac{dC_I}{dx} = 0 \quad \text{at } x=h/2 \quad (76)$$

$$\frac{dC_V}{dx} = 0 \quad \text{at } x=h/2 \quad (77)$$

k_I and k_V are the surface annihilation rate constants for the self-interstitial and the vacancy, respectively. Subtracting Eq. (73) from Eq. (72), one obtains

$$D_I \frac{\partial^2 C_I}{\partial x^2} = D_V \frac{\partial^2 C_V}{\partial x^2} \quad (78)$$

The general solution under the given boundary conditions leads to the following relationship between C_V and C_I [273]:

$$C_V = (D_I/D_V)C_I + B \quad (79)$$

where B is a constant given by

$$B = \frac{k_I}{k_V} [C_I(0) - C_I^*] - \frac{G_s}{k_V} + C_V^* - \frac{D_I}{D_V} C_I(0) \quad (80)$$

Thus, in the steady state (as $t \rightarrow \infty$), the C_V profile rises and falls with the C_I profile according to Eq. (79). It remains to find $C_I(x)$. With the relationship of Eq. (79), Eq. (78) can be reduced to a second-order nonlinear differential equation for $C_I(x)$:

$$\frac{d^2 C_I}{dx^2} = \frac{k_r}{D_V} C_I^2 + \frac{k_r B}{D_I} C_I - \frac{k_r}{D_I} C_V^* C_I^* \quad (81)$$

which can be solved analytically [273]. The analytical solution is somewhat complicated, and the reader is referred to Ref. [273] for detail. Fig. 18 schematically illustrates the C_I and C_V distributions in the steady state under a constant surface injection. It is clear from the figure that the relationship $C_V C_I = C_V^* C_I^*$ is never observed, even at the mid point of the substrate ($x=h/2$). The departure from the relation becomes much bigger in the surface region.

To put it simply, since there is only a finite initial supply of vacancies in the bulk, a continuous injection of self-interstitials from the surface necessarily demands a vacancy supply from the surface. (The transient case is different in that the deep bulk can also supply vacancies to meet consumption.) This means both $C_I(\text{surface}) > C_I(\text{bulk})$ and $C_V(\text{surface}) > C_V(\text{bulk})$. Allowing that $C_I C_V(\text{bulk}) \approx C_I^* C_V^*$, it is simple logic that $C_I C_V(\text{surface}) > C_I^* C_V^*$. This is essentially what is depicted in Fig. 18. Fig. 18 is obtained based on the assumption that, while the rate of interstitial injection is given, the rate of vacancy injection is governed by both surface reaction and diffusion, i.e. the boundary condition of Eq. (75). If it is diffusion-controlled, then we would have $C_V(0) = C_V^*$, because of the minimum free energy requirement.

The observation of ORD requires that there be some finite vacancy undersaturation, not zero. That this is indeed the case is indicative of surface reaction limited kinetics, which is consistent with the observation of the surface annihilation limited kinetics of excess self-interstitials, as manifested in the dependence of the kinetics on surface films [27].

It is readily seen from Fig. 18 that ORD should increase with depth, in contrast to OED, which decreases with depth.

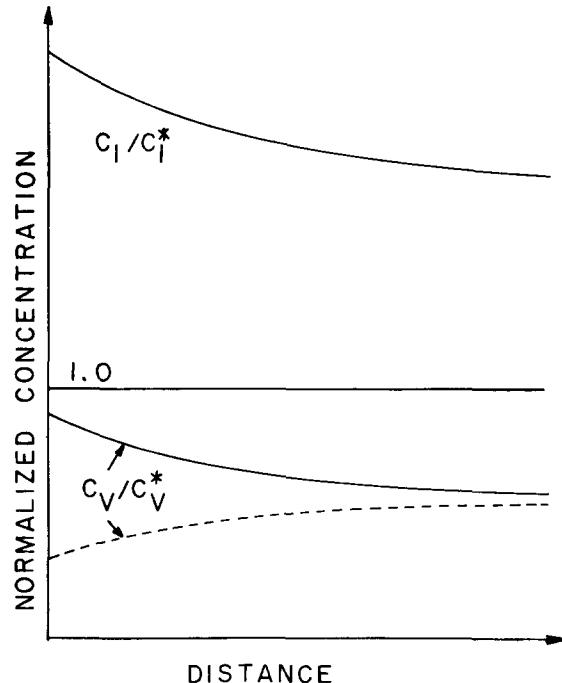


Fig. 18. Schematic illustration showing the relationship between the steady state concentrations of vacancies and interstitials under a constant injection of self-interstitials that would obtain from the solution of coupled reaction-diffusion equations (solid curves), and the vacancy concentration that would obtain under the assumption $C_V C_I = C_V^* C_I^*$ (taken from Hu [273]).

8.3.2. Transient analysis

Yeager and Dutton [274] and Fahey et al. [3] have carried out numerical solutions of Eqs. (70) and (71) to investigate the transients of vacancies and self-interstitials in the presence of a constant injection rate, G_S , for the boundary condition of Eqs. (74) and (75), using plausible values of kinetic parameters. Their results are largely similar. The result of Fahey et al. is shown in Fig. 19. Comparing Figs. 18 and 19, one sees that the transient case is much more complex. Beyond that, one sees from Fig. 19 that, for distance $\leq 5 \mu\text{m}$ from the surface, C_V increases toward the surface, as does C_I , and consequently there is a large departure from the relationship $C_V C_I = C_V^* C_I^*$. The assumption of this relationship is hence invalid for modeling dopant diffusion which takes place within a distance of $\approx 5 \mu\text{m}$ from the substrate surface. The rate constant of bulk annihilation between vacancies and self-interstitials was taken to be diffusion-limited, both in the studies of Yeager and Dutton and of Fahey et al. A reaction barrier will increase this distance.

As in Hu's steady state solution, the transient analysis would also predict ORD to increase with depth in the surface region within the diffusion length of vacancies. So far, however, there have been no suitable experimental data to confirm this prediction.

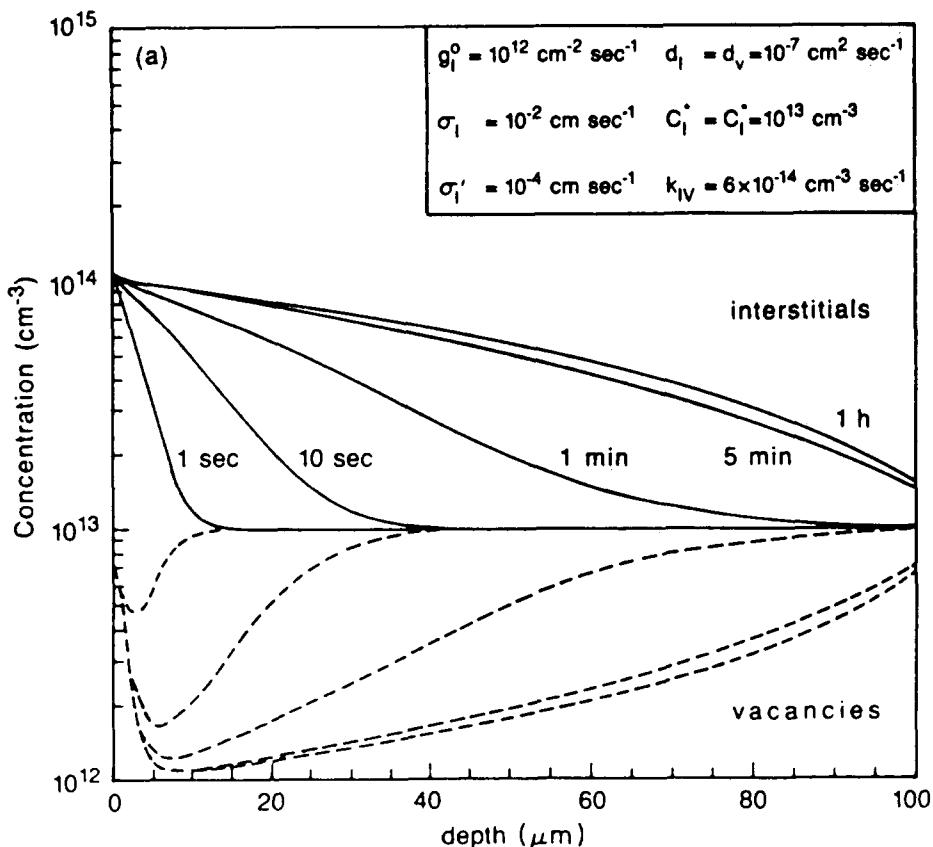


Fig. 19. Transient concentrations of self-interstitials and vacancies under a constant injection, as obtained from a numerical solution of the coupled point defect reaction-diffusion equations, assuming diffusion-limited recombination between vacancies and self-interstitials. Notice that the diffusion length of the excess point defects is about $5 \mu\text{m}$, within the depth of which from the surface we have the relationship $C_V C_I \gg C_V^* C_I^*$. The possible existence of a reaction barrier would increase the diffusion length (taken from Hu [3]).

A conclusion to be drawn from these results is that, regardless of whether in steady state or in transient conditions, the dynamic equilibrium within the point defect diffusion length is really $C_v C_i > C_v^* C_i^*$. For a lower-bound estimate of the diffusion length, we consider the vacancy–interstitial annihilation to be diffusion-limited, with the rate constant given by [275,276]:

$$k_{\text{ann}} = 4\pi r_0 D_X C_{\text{ann}} \quad (82)$$

where r_0 is the capture radius of the annihilation center, about 2.34 Å (the sum of the radii of two reacting particles). The annihilation center is the conjugate defect of X, for example, a self-interstitial for a vacancy. With the diffusion and annihilation rates sharing the same kinetic parameter, namely the defect diffusivity, the diffusion distance of the excess vacancies and self-interstitials is governed only by the concentration of the conjugate defect. Strictly speaking, a diffusion length defined in the sense of exponential decay is inappropriate for this case, because both C_v and C_i decay rapidly, so that the diffusion reaction differential equation is not linear. With this caveat, we would give a rough estimate of the diffusion length to be $L = (D_X/k_{\text{ann}})^{1/2} = 1/2(\pi r_0 C_{\text{ann}})^{1/2}$. Using a value of $C_{\text{ann}} = 10^{13} \text{ cm}^{-3}$ (rough estimate for typical diffusion temperatures 1000–1100 °C), one gets $L = 9 \mu\text{m}$. The diffusion length will be larger if the annihilation is reaction-limited. In the hypothetical case where $C_v^* = C_i^*$ and $D_v = D_i$, under interstitial injection the diffusion length for the self-interstitials will be substantially larger than for the vacancies, because of the lower concentration of vacancies (the reciprocal square root relationship).

It may be noted that, from proton-enhanced diffusion, Masters [277] has measured a vacancy diffusion length of 3.5 μm at 900 °C. The diffusion length of point defects in particle-irradiated silicon, such as in Masters' case, is expected to be substantially shorter than that in oxidized silicon, because of the very high concentrations of conjugate defects produced in the former. Finally we should point out an interesting feature in Masters' result, that the diffusion length at 900 °C is larger than at 600 °C. As a rule of thumb, the square of the diffusion length of a decaying species is proportional to its diffusivity, which increases with temperatures; but it is also inversely proportional to the recombination rate constant, which also increases with temperature, at least equally strongly in the case of diffusion-limited reaction. A possible interpretation is that the diffusivity that affects the diffusion length of a species pertains to that species, say the vacancy, with a larger migration energy, whereas the diffusivity that determines the diffusion-limited reaction rate pertains to the faster diffusion conjugate defect, say the interstitial, with a smaller migration energy. The disparity in the diffusivity between the vacancy and the self-interstitial leads to a disparity in the nonequilibrium concentration profiles of the two species of point defects, as schematically illustrated in Fig. 17. The slower diffusing species will thus have a smaller diffusion length. In contrast, for the faster diffusing species, its diffusivity determines both its migration away from the location of its creation, as well as the diffusion-limited reaction rate. Its diffusion length should then be independent of temperature in the simplified model.

9. Pumping of point defects by diffusion

9.1. Experimental observation of point defect supersaturation by chemical pumping

Phosphorus diffusion provides a prime example of chemical pumping of point defects – and more specifically self-interstitials in this case. Three phenomena are clear manifestations

of the pumping of self-interstitials by the diffusion of phosphorus into silicon, leading to a large supersaturation of self-interstitials in the interior. The forerunner of these phenomena is the so-called “emitter push effect”, also called the “emitter dip”. This phenomenon was first reported by Miller [278], and was subsequently investigated and/or discussed by many others [279–284]. The effect refers to the enhanced movement of the base–collector junction directly below a phosphorus emitter, relative to the base–collector junction lateral to the emitter (see Fig. 20). The base dopant in this phenomenon is usually boron or gallium. Miller [278] suggested that it might be due to an enhanced diffusion of base acceptor dopants, boron or gallium, caused by the increased equilibrium vacancy concentration as a result of an increased Fermi level in the phosphorus-doped emitter region. But an elevated Fermi level increases only the concentration of negatively charged vacancies (and other negatively charged point defects). Because of electrostatic repulsion, negatively charged point defects do not effectively mediate the diffusion of acceptor dopants. On the contrary, the concentration of positively charged vacancies (and other positively charged point defects) will decrease with an increased Fermi level, while the concentration of neutral point defects remains unaffected. Therefore the diffusion of acceptor dopants should be retarded rather than enhanced in heavily doped n-type regions. Indeed, Makris and Masters [285] have found retarded gallium diffusion in silicon doped heavily and homogeneously with arsenic. Moore [279] suggested that the concentration of vacancies in the emitter region may be increased by elastic strain from the extremely high phosphorus concentration. (We note that the charge state of the point defects does not matter in interactions with elastic strain, and concentrations of all charge states of a defect species should be affected in direct proportion. What matters is the nature of the point defect. The equilibrium concentration of vacancies should increase while the equilibrium concentration of self-interstitials should decrease in regions with heavy phosphorus doping, which causes the silicon lattice in a confined geometry to be tensioned. But this explanation, too, is contradicted by the observation of Makris and Masters. Furthermore, the emitter push effect occurs even when no dislocation can be found in a phosphorus emitter [286], thus negating the role of dislocations in generating the excess point defects.

An example of the peculiarity of a phosphorus diffusion profile is shown in Fig. 21 [287]. An established method of analyzing the concentration dependence of diffusivity is the so-called Matano analysis [288]. Yoshida et al. have performed such an analysis on many phosphorus profiles [286] to obtain the phosphorus diffusivity as a function of the local and the surface concentrations of phosphorus. Their results are shown in Fig. 22. One sees that, for phosphorus profiles with a low surface concentration, the diffusivity as a function of local concentration is quite similar to that discussed by Hu and Schmidt [289],

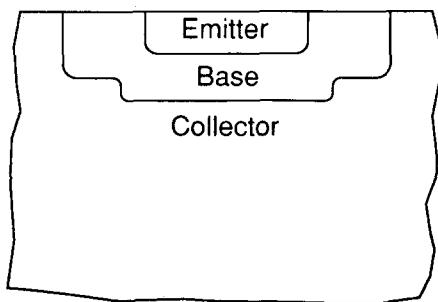


Fig. 20. The emitter push effect: the base–collector junction in a transistor is pushed out under a phosphorus emitter.

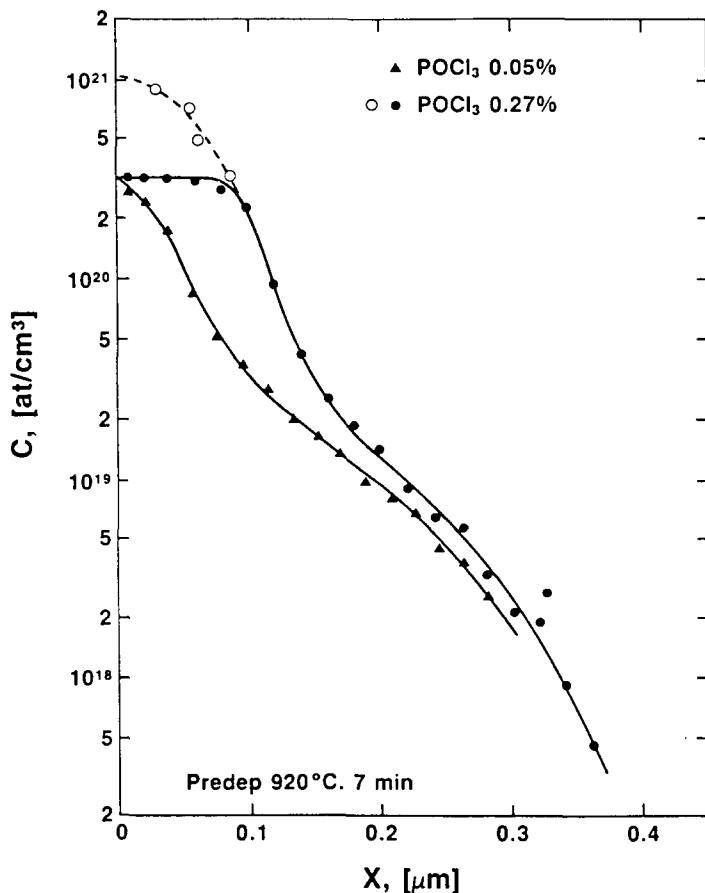


Fig. 21. Phosphorus concentration profiles from POCl_3 diffusion with two different source concentrations. Notice the profile tails (taken from Solmi et al. [287]).

based on the concept of the Fermi level dependence of equilibrium charged point defects. For profiles with high surface concentrations, the situation is quite complicated, and the diffusivity obtained with the Matano method at low local concentration is much greater, by hundreds of times. Apparently, a nonlocal effect is manifested here. It should be noted that the nonlocal effect on diffusivity is not limited to phosphorus diffusion. Similar Matano analyses of boron and arsenic diffusion profiles and similar nonlocal effect of high surface concentrations has been found. The only difference is that this effect is much smaller for arsenic diffusion (see Fig. 23 [290]). We should mention that the application of Matano analysis is incorrect when the diffusivity is dependent on anything other than a local impurity concentration (for example, a function of t or x) [288]. However, in the case of diffusivity enhancement due to point defect supersaturation from a "chemical pump" (see Section 9.2), a quasi-steady state analysis may yield a diffusivity that is dependent only on the local impurity concentration, as well as the surface impurity concentration, which is, however, a constant throughout the diffusion process of the case in question (see Eq. (105)).

The best explanation now available for the emitter push effect and the occurrence of long profile tails is the chemical pump of point defects by phosphorus diffusion, building up an extremely high supersaturation of these point defects in the interior of the substrate substantially beyond the diffusion front. The occurrence of long-ranging nonequilibrium point defects was demonstrated in an elegant experiment of Lee and Willoughby [291].

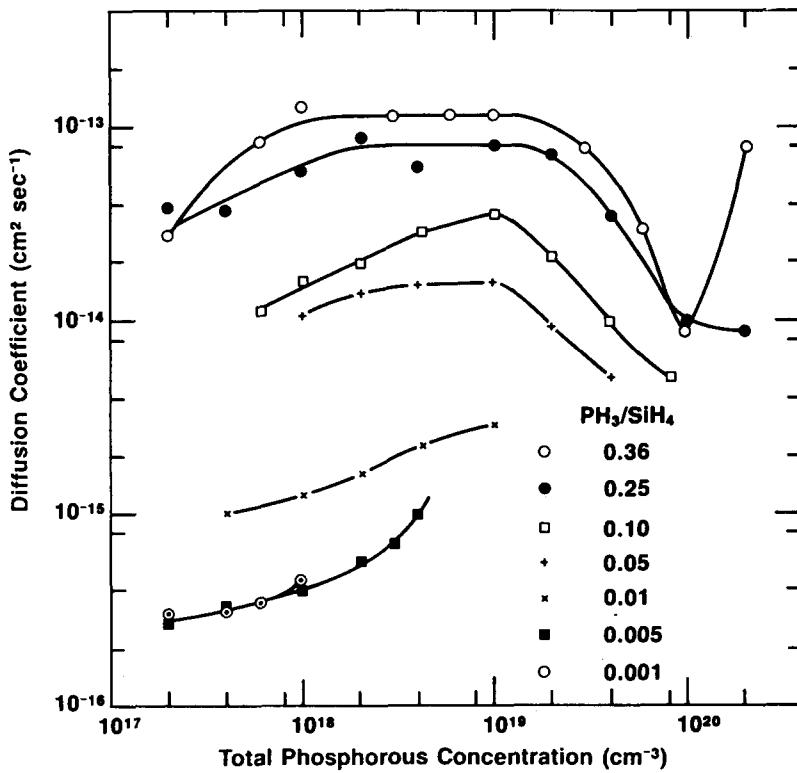


Fig. 22. The diffusion coefficient of phosphorus as a function of local and surface concentrations, as obtained from Matano analysis of phosphorus concentration profiles having different surface concentration. Notice the very strong effect of surface concentration, enhancing the low local concentration diffusivity by as much as more than two orders of magnitude (taken from Yoshida et al. [286]).

Theirs was the first experiment to use a buried layer monitor for long-range supersaturation of point defects. They reported a point defect supersaturation of more than 100, and a range of more than $10 \mu\text{m}$. Yoshida et al. [286,292] found that an extremely large supersaturation of point defects, which they considered as vacancies, extends to more than $20 \mu\text{m}$ deep, and that this supersaturation is affected much more by the surface concentration than by the local concentration of phosphorus. They suggested that phosphorus diffuses via the vacancy–phosphorus pair, often referred to as an E center (a nomenclature dating back to 1959 [293]). They suggested [286,292] that the vacancy–phosphorus pairs dissociate after entering the bulk of the substrate, releasing vacancies which become supersaturated. They have performed numerical simulations based on this model, and have been able to demonstrate the occurrence of a long diffusion tail as a consequence. But the plateau and kink of the phosphorus profile could not be simulated. Earlier, Schwettmann and Kendall [294,295] investigated the plateau-and-kink-and-tail feature of phosphorus diffusion profiles. They observed that the processes occurring in the highly concentrated region from 0–4000 Å play a major role in the tail formation. For example, when this region was removed by anodization (to a concentration of $1.1 \times 10^{20} \text{ cm}^{-3}$) and the slice annealed at 600°C for 150 h, no tail was observed. A comparable deposition with no surface removal annealed at the same time gave an increased junction penetration of $0.64 \mu\text{m}$. Furthermore, they observed that the kink usually occurs at a concentration corresponding to a Fermi level that approximately coincides with the energy level of the negative charge state of the E center. Schwettmann and Kendall's model of phosphorus diffusion comprises the following

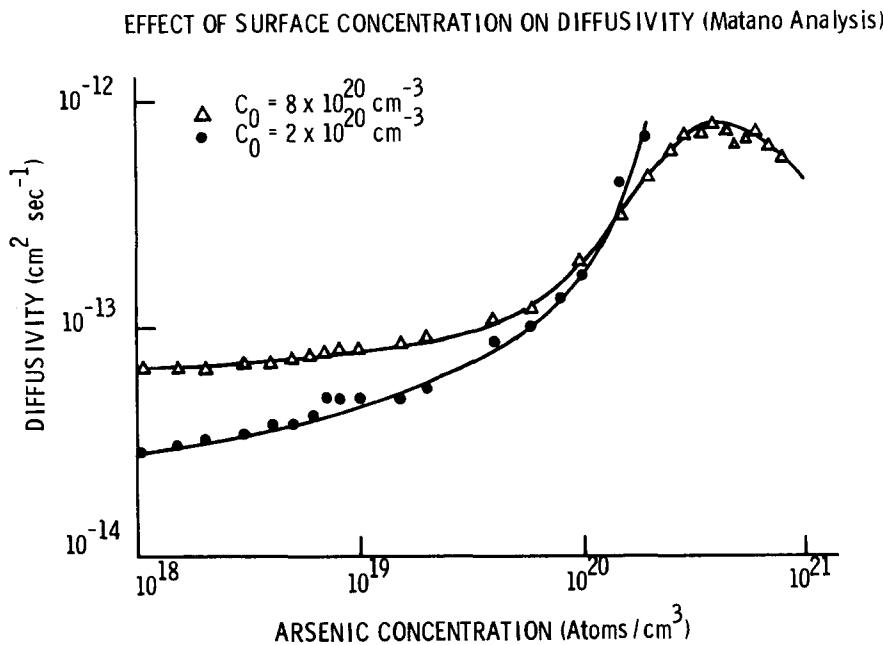


Fig. 23. The diffusivity of arsenic as a function of local and surface concentrations, as obtained from Matano analysis of two arsenic concentration profiles having different surface concentrations. Notice that the effect of surface concentration on the low local concentration diffusivity is much weaker (no more than a factor of 3 enhancement) than that for phosphorus (taken from Hu [290]).

three elements. (1) Phosphorus diffuses via the E center in its two charge states, E^- and E^0 ; the latter species diffuses much faster than the former. (2) Electrical compensation by the negatively charged E centers accounts for the inactive phosphorus in the concentration plateau in the surface region. (3) The kink occurs when, at a characteristic concentration corresponding to a specific Fermi level, the E center loses an electron to become E^0 , which then diffuses at a much faster rate to produce a long tail. But this model does not explain the emitter push effect. So Fair and Tsai [296] suggested that, at that characteristic concentration of the kink where E^- becomes E^0 , E^0 dissociates much more readily, producing a supersaturation of vacancies. However, instead of solving reaction-diffusion equations such as attempted by Yoshida et al. [286,292], or by balancing the vacancy generation from E center dissociation with vacancy out diffusion to the surface (as done in the quasi-steady analysis of chemical pump in Section 9.3.2), Fair and Tsai obtained a steady state supersaturation of vacancies simply by means of a series of mass action equilibrium relations. Apparently they did not even consider matter conservation to be essential, so that the concentration of E^- in the narrow surface region will simply translate into a corresponding concentration of E^0 in the half space beyond the kink. A series of mass action maneuvers culminates in their Eq. (17) [296] which expresses the vacancy supersaturation as

$$\frac{C_{V^-}}{C_{V^-}^*} = \frac{n_s^3}{n_e^2 n_i} \left[1 + \exp\left(\frac{0.3eV}{kT}\right) \right] \quad (82a)$$

where n_s is the surface electron concentration, n_e is the electron concentration corresponding to the Fermi level of the E^- state and n_i is the intrinsic electron concentration. The above expression gives a supersaturation that is independent of how fast the vacancies are pumped into the bulk, in balance against leaking back to the surface. Therefore, it cannot explain

why only fast diffusing impurities exhibit diffusion tails and the emitter push effect. Furthermore, it does not contain a local concentration factor, and therefore cannot lead to a transition to a low surface concentration case where the diffusivity is mainly affected by the local concentration.

Many later observations have established the identity of the point defect in supersaturation as the self-interstitial, rather than the vacancy. One is the observation by Strunk et al. [115] that helical screw dislocations which grow in regions of phosphorus diffusion exhibit the sense that is indicative of being driven by an interstitial supersaturation. The second is the observations by Harris and Antoniadis [297] and Fahey et al. [116] that the diffusion of buried antimony layers is retarded from in-diffusion of phosphorus. In Refs. [115 and 297], the diffusion of phosphorus from POCl_3 sources occurred under oxidizing conditions, so that the excess self-interstitial could have come from the thermal oxidation. This, in fact, was the explanation given by Strunk et al. [115]. But, from the number of turns per unit length of the helices, Gösele and Frank [259] estimated a supersaturation of $C_1/C_1^* \approx 140$, more than an order of magnitude larger than typically obtained during oxidation, and concluded that the influence of thermal oxidation on the self-interstitial supersaturation is not significant. The experiment of Strunk et al. explains the enhanced diffusion of a boron base under a phosphorus emitter (the "emitter push" effect) as due to an interstitial supersaturation. The experiment of Fahey et al. [116] was designed to remove the role of thermal oxidation, in order to test the four possible sources of excess self-interstitials as suggested by Hu et al. [298]. Their experiments involved incremental phosphorus diffusion carried out under non-oxidizing conditions, to produce the retarded diffusion of the buried antimony (see Fig. 24). The results clearly showed that self-interstitials are carried into the bulk of silicon by phosphorus atoms entering the silicon lattice, and are released in the bulk and become supersaturated. The diffusing entity is a complex of a phosphorus atom

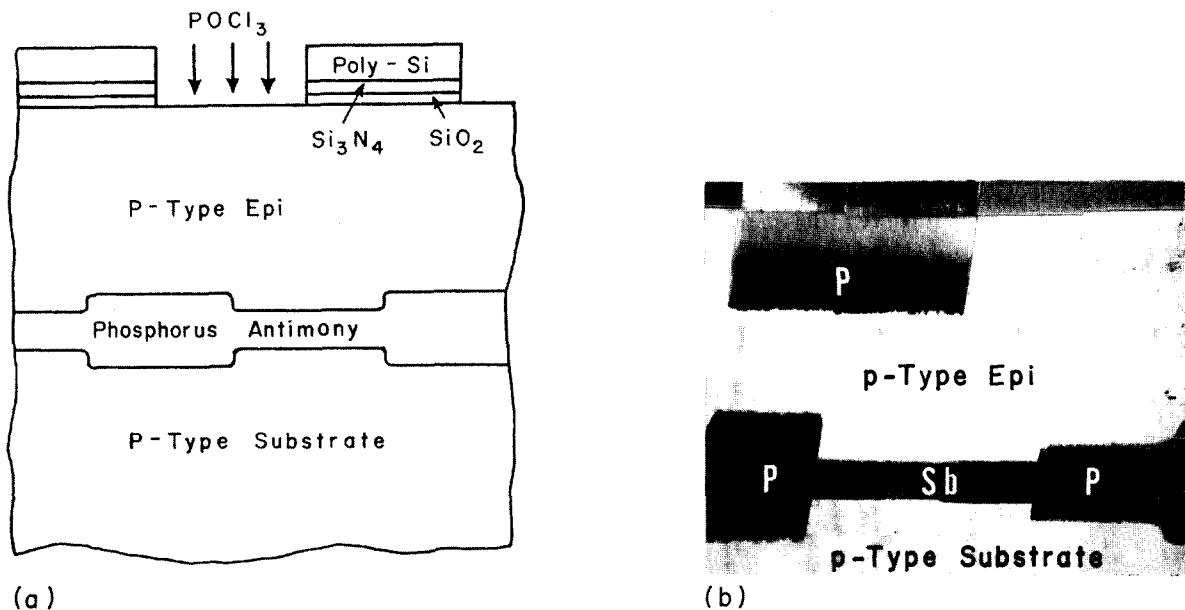


Fig. 24. Experiment to identify the cause and the identity of the point defect in supersaturation from phosphorus diffusion. (a) Starting structure with adjoining phosphorus and antimony buried layers. (b) Photomicrograph of beveled and stained sample showing an enhanced diffusion of the phosphorus buried layer and a retarded diffusion of the antimony buried layer (taken from Fahey et al. [116]).

and a self-interstitial as a pair, which could alternatively be viewed as a phosphorus interstitialcy. The third is the observation of Claeys et al. [195] that bulk stacking faults would grow under the in-diffusion of phosphorus in a nitrogen ambient.

The excess self-interstitials responsible for the growth of stacking faults and helical dislocations, as well as the emitter push effect, is apparently responsible for the occurrence of long tails in the phosphorus diffusion profiles.

9.2. Chemical pump and vacancy/interstitialcy winds

Seitz [10,299] showed that, if a substitutional impurity denoted as A diffuses faster than the lattice atom denoted as B, then the flux of A will cause a counter flux of vacancies, assuming that only the vacancy mechanism operates. He called this the “chemical pump”. He arrived at this conclusion through the following argument.

1. The sum of all fluxes of substitutional species, including the vacancy, is zero for conservation of lattice sites. For the case in which the lattice comprises constituents A, B and V, this, in one-dimensional diffusion, is expressed as

$$J_A + J_B + J_V = 0 \quad (83)$$

2. The flux of each of A, B and V is given according to Fick's law as

$$J_k = -D_k \frac{\partial C_k}{\partial x}, \quad k = A, B, V \quad (84)$$

3. For the case of a uniform vacancy concentration, the following relationship holds:

$$\frac{\partial C_B}{\partial x} = -\frac{\partial C_A}{\partial x} \quad (85)$$

because the concentration of total lattice sites, $C_A + C_B + C_V$, is a constant. C_V is negligibly small compared to C_A and C_B . Combining the above three equations, one gets

$$J_V = (D_A - D_B) \frac{\partial A}{\partial x} \quad (86)$$

The above equations says that, even when there is no vacancy concentration gradient, there will be a vacancy flux that is in the opposite direction of the impurity flux (but is not of equal magnitude), if $D_A > D_B$. A corollary can also be stated for the reverse situation, in which a vacancy flux would induce a counter flux of impurity atoms in the opposite direction. This is the “vacancy wind” effect.

The arguments leading to Eq. (86) are flawed on two counts. The first flaw in the derivation of Eq. (86) is the application of Fick's law [8]. There are two correct ways of analyzing the vacancy and/or interstitialcy winds. The first way proceeds in the irreversible thermodynamic formalism, by using a set of phenomenological equations of diffusion of all lattice components with due considerations of all cross terms. This approach was adopted by Darken [300], and by Bardeen and Herring [301] in their analysis of the Kirkendall effect [302], a phenomenon in which diffusion across a heterostructure, such as is formed by joining two slabs of binary alloy of different compositions, causes the interface to move relative to some stationary markers. The main problem with this approach is that the cross term coefficients are taken as already known, but the phenomenological approach does not provide a way of determining these coefficients. The second approach is through detailed

atomistic modeling. All the interactions, such as the formation of defect–impurity complexes are physically modeled, and the vacancy/interstitialcy winds will automatically come out in the solution. This is the approach that we will take. An antecedent of the present atomistic approach is the kinetic approach of Fisher et al. [303] and Seitz [10,299]. The kinetic approach of this earlier work does not consider the specifics of diffusion mechanism such as the binding of point defects to impurity atoms, and hence cannot be expected to yield correct answers to the vacancy/interstitialcy wind phenomenon.

Seitz' equation of chemical pumping is based on the fact that the impurity atom diffuses faster than the host atom. But this fact does not exist independently of other phenomena. It exists because there is an attractive interaction between the vacancy and the impurity atom [8,37]. This attraction causes a greater probability for any one impurity atom to be next to a vacancy than for any one host atom. A strong affinity between the vacancy and the impurity atom would also cause the vacancy and the impurity atom to diffuse as a pair, rather than to diffuse in opposite directions [8]. The pair diffusion was qualitatively suggested by Johnson much earlier. Assuming that the flux sum is valid, this means simply that J_B is given by $-(J_A + J_V)$, not by $-D_B \partial C_B / \partial x$ according to Fick's law. In other words, the driving force for B diffusion is dominated by cross coefficients with A and V, rather than by its concentration gradient [8].

The second flaw relates to the axiomatic assumption of the flux sum rule, Eq. (86), which we shall show to be not rigorously correct. (This may not be an important issue in the present context; but it has been so widely accepted that it is felt that the issue should be set straight here.) To derive the correct flux sum rule [304], we start with the assumption that the concentration of the lattice sites, C_S , remains unchanged:

$$\begin{aligned} 0 &= \frac{\partial C_S}{\partial t} = \frac{\partial C_A}{\partial t} + \frac{\partial C_B}{\partial t} + \frac{\partial C_V}{\partial t} \\ &= -\frac{\partial}{\partial x} (J_A + J_B + J_V) \end{aligned} \quad (87)$$

The above equation expresses just the condition of continuity, and is unrelated to Fick's law of diffusion. The general solution to the above equation is

$$J_A + J_B + J_V = \frac{u(t)}{\Omega} \quad (88)$$

where Ω is the atomic volume and $u(t)$ is the velocity of the movement of the whole lattice relative to some stationary markers. $u(t)$ is, obviously, a function of time only, and is independent of position x , since it is implicit in our derivation that the solid is not deformed by diffusion. $u(t)$ may be negative, zero or positive, and can only be obtained with cross coefficients determined from atomistic modeling. The above equation expresses the Kirkendall effect. It is possible that $u(t)$ may indeed be very small in silicon diffusion systems; but there has been no experimental data for the Kirkendall effect in any silicon diffusion system.

Seitz also gave a flux sum rule involving the self-interstitial in the mediation of diffusion, a mechanism which he named "interstitialcy", now a standard terminology. This sum rule is:

$$J_A + J_B - J_I = 0 \quad (89)$$

The above equation has been taken to mean that an "interstitialcy wind" would induce an impurity flux in the same direction [305]. Now, it is possible (though not necessary)

that an interstitialcy wind may cause an impurity flux in the same direction as the interstitialcy flux. However, when this happens, it is not because of Eq. (89), for the same reason given in the discussion of vacancy wind. Actually, Eq. (89) is incorrect, because an interstitialcy, defined by Seitz as a lattice site having an extra atom, occupies a lattice site and is a constituent of the lattice. This is readily recognized in the case of an interstitialcy having a split interstitial configuration. Its concentration, as is the concentration of a substitutional atom or a vacancy, is hence subject to the constraint of constant concentration of lattice sites. Thus, for the case in which both vacancies and interstitialcies are present, Eq. (88) is generalized to

$$J_A + J_B + J_V + J_I = \frac{u(t)}{\Omega} \quad (90)$$

Failure to recognize this relationship may lead to incorrect formulation of diffusion systems. For example, in analyzing the forces on charged defects in semiconductor heterostructures, Tersoff [306] asserted that moving a dopant atom across the boundary of a heterostructure entails moving a lattice atom across the boundary in the opposite direction, apparently not recognizing the roles of point defects and the existence of the Kirkendall effect. (He also asserted erroneously [306] that the strain effect, being short-range in atomic scale, plays a negligible role in diffusion across a heterostructure. The fact is, heterostructures typically exhibit very large epitaxial strains, which significantly raise the chemical potentials of solutes of misfitting atoms [307–310]. But this is outside the scope of the present review.)

Vacancy winds and/or interstitialcy winds may occur in substrates subjected to bombardments by energetic particles, including ion implantations. The observed uphill diffusion of boron and phosphorus due to proton irradiation of silicon [311,312] cannot be explained simply in terms of radiation-enhanced diffusion [313,314]. Flux interactions have to be invoked. The result of a study [315] of proton irradiation of silicon with a uniform initial boron profile showed the boron to diffuse uphill and away from, rather than toward, the proton projected range. At the approximate location of the proton projected range, thousands of vacancies and an equal number of silicon interstitials are generated by each bombarding ion. Some of these vacancies and interstitials will recombine locally, and the rest of them will diffuse in both directions away from the location where they were generated, before they recombine. (This is a case where there should be no argument that $C_I C_V \gg C_I^* C_V^*$. Since vacancies and self-interstitials are generated and annihilated in equal numbers, the supersaturation of one species necessarily means the supersaturation of the other species as well.) Gösele and Strunk [305] argued that the observation of Loualiche et al. [315] is evidence of the interstitialcy mechanism of boron diffusion in silicon. Their reasoning goes as follows: in the concept of vacancy wind according to Eq. (83), one should have expected the boron to be pumped to the location of defect generation, in contradiction to the observed results just mentioned. On the other hand, in the concept of interstitialcy wind according to Eq. (89), one would expect the boron to be carried away from the location of defect generation, which is in agreement with the experimental observation. These arguments of Gösele and Strunk [305] are invalid because they rest on Eq. (89), which we said is invalid. But these arguments are invalid also because they contradict later experimental results of [316] Kozlovskii et al., who observed antimony to diffuse away from the location of point defect generation, in exactly the same manner as boron. From this observation, their arguments would have led to the conclusion that antimony also diffuses in silicon via a dominant interstitialcy mechanism. But it has been established, through the oxidation-

retarded diffusion, that antimony diffusion in silicon is dominated by the vacancy mechanism, with only about 2% of interstitialcy component [3,46]. Thus, it is quite clear that the direction of the flux of substitutional atoms relative to that of the point defect mediating the diffusion is determined mainly by the attractive interaction between the impurity atom and the point defect, as already pointed out in the author's 1969 paper [8]. We will analyze in detail the radiation-induced redistribution of impurities in Section 11.

9.3. Chemical pump and diffusion enhancement

9.3.1. Analysis of a simple diffusion system

For simplicity of analysis in this section, we shall assume the impurity and the point defect to be electrically neutral. The fact that the activation energy of impurity diffusion is smaller than that of self-diffusion indicates that there is a binding energy between the impurity atom and the native point defect [8,37,317]. Hu has presented [37] a relationship between the point defect-impurity atom binding energy, ΔE_b , and the difference between the activation energies of self-diffusion and impurity diffusion in diamond lattice, with specific numerical relationship for the case when the binding energy is mainly due to electrostatic interaction. But an abundance of experimental observations has indicated that elastic interactions are very important, and, in fact, will mainly decide whether a substitutional impurity atom diffuses via a dominantly interstitialcy, or a dominantly vacancy mechanism. A consequence of the binding between the substitutional dopant atom and the point defect is that a diffusing dopant atom carries with it a point defect from the surface to the interior of the silicon substrate, releases the point defect in the interior, and itself settles down and becomes a normal substitutional dopant atom. The formation and dissociation of the defect-dopant pair follows the mass reaction:



(For a vacancy mechanism of diffusion, I is replaced by V in the above expression.) In the surface region, where C_A is large, the reaction shifts to the right for formation of more dopant-interstitial pairs. Since only AI, but not isolated A, can move, the interior region will be swamped with an overpopulation of AI, tipping the local balance and causing the above reaction to shift to the left. Self-interstitials (or vacancies in a vacancy mechanism) are released and become supersaturated in the interior. In this way, point defects are pumped by the dopant diffusion into the interior, and reflux to the surface (see Fig. 25). The local supersaturation of point defects in this process can be, found from the solution of a system of simultaneous equations comprising the above reaction, and the continuity equations for the point defect and the AI complex.

Since only AI, but not A, can physically move in the lattice, the diffusion equation of dopant A is not, rigorously speaking, given by

$$\frac{\partial C_A}{\partial t} = \frac{\partial}{\partial x} \left(D_A \frac{\partial C_A}{\partial x} \right) \quad (92)$$

The above equation can be regarded as an *effective* diffusion equation of A only under the condition of point defects being at equilibrium concentration, or at least when the concentration $C_I(x, t)$ is given, or can be obtained independently of the diffusion of A, such as when $C_I(x, t)$ is dominated by an independent process, e.g. the thermal oxidation. Furthermore,

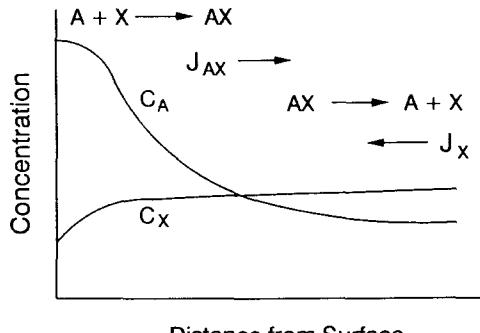


Fig. 25. Chemical pump of point defects by pair diffusion: mobile complex AX is formed in the surface region and diffuses inward; in the interior, where C_A is low, AX dissociates, releasing point defect X as A settles down; the excess concentration of X builds up, causing X to diffuse to the surface, which is assumed to be the major sink. Thus, point defects are being pumped into the interior against the C_X gradient. Note that a "kick-out" mechanism has an exactly identical mathematical formulation, and will produce exactly identical results.

there must be a near equilibrium between the defect-dopant reaction of Eq. (91):

$$C_{AI} \approx K_{AI} C_A C_I \quad (93)$$

so that, while physically D_A is zero, an effective diffusivity of A is given by

$$D_A \approx D_{AI} \frac{C_{AI}}{C_A} \approx D_{AI} K_{AI} C_I \quad (94)$$

The equilibrium relationship of Eq. (93) between A and I is an acceptable approximation for the purpose here. In the presence of a strong chemical pumping of point defects, however, the diffusion process must be written as physically taking place. This involves the solution of at least the following three simultaneous equations:

$$\frac{\partial C_{AI}}{\partial t} = - \frac{\partial J_{AI}}{\partial x} + k_{AI}^f C_A C_I - k_{AI}^r C_{AI} \quad (95)$$

$$\frac{\partial C_I}{\partial t} = - \frac{\partial J_I}{\partial x} + k_{AI}^r C_{AI} - k_{AI}^f C_A C_I \quad (96)$$

$$\frac{\partial C_A}{\partial t} = k_{AI}^r C_{AI} - k_{AI}^f C_A C_I \quad (97)$$

More rigorously, even if the dopant diffuses via a dominantly interstitialcy mechanism ($D_{AI} C_{AI} \gg D_{AV} C_{AV}$), we should also include the participation of the vacancy in the following reactions:



where S denotes silicon atoms at lattice sites. Thus, in addition to an extra equation of continuity for the vacancy, the reaction terms must also be modified accordingly. The effect of the participation of the vacancy in this case is indirect, through its influence on both C_I and C_{AI} . When both the interstitialcy and the vacancy mechanism must be considered, then there is an additional reaction involving the formation and dissociation of the AV pair. In typical conditions, $C_{AI} \ll C_A$, so that, if $C_V \approx C_I$ and $D_V \approx D_I$, the reaction rate of

Eq. (99) will be negligible compared to that of Eq. (91), assuming diffusion-limited reactions. Thus, it is quite reasonable to omit these additional reactions.

9.3.2. Quasi-steady state analysis of chemical pump

Under certain conditions, a simple but rather satisfactory method of solving the point defect transients caused by chemical pumping is the quasi-steady state approximation, first made by Schaake [318], and subsequently by others [319–321]. Furthermore, a local equilibrium between the diffusing intermediate complex and their parents is assumed. With $\partial C_{AI}/\partial t \approx 0$ and $\partial C_I/\partial t = 0$, we obtain from Eqs. (95) and (96)

$$J_{AI} + J_I = 0 \quad (100)$$

The participation of the vacancy has been neglected in this example. Not included is an integration constant which is a function of time only. For many reasonable boundary conditions, this constant can be taken to be zero. With the assumption of the local equilibrium of the diffusing complex AI, one can write

$$J_{AI} = -D_{AI} \frac{\partial}{\partial x} K_{AI} C_A C_I \quad (101)$$

$$J_I = -D_I \frac{\partial C_I}{\partial x} \quad (102)$$

Combining Eqs. (100)–(102), one obtains:

$$\frac{dC_I}{C_I} = \frac{-dC_A}{\alpha + C_A} \quad (103)$$

with $\alpha = D_I/K_{AI}D_{AI}$. Under the boundary condition of $C_I(0) = C_I^*$, the above equation is easily integrated to give

$$C_I(x) = C_I^* \frac{\alpha + C_A(0)}{\alpha + C_A(x)} \quad (104)$$

In obtaining Eq. (104), it has been assumed that the annihilation of excess self-interstitials at the surface is diffusion-controlled. Eq. (104) can be regarded as an explicit expression of the local dopant concentration dependence of C_I , and hence of the local dopant concentration dependence of diffusivity. This would enable a diffusion equation such as Eq. (92) to be integrated directly. By use of the definition of the effective diffusivity D_A in Eq. (94), we can recast the above equation as

$$\frac{C_I(x)}{C_I^*} = \frac{\frac{D_S}{D_A} f_I^S C_S + C_A(0)}{\frac{D_S}{D_A} f_I^S C_S + C_A(x)} \quad (105)$$

where D_S is the silicon self-diffusivity, f_I^S is its fractional interstitialcy component and C_S is the concentration of silicon atoms. An inspection of Eq. (105) indicates that the point defect supersaturation will be greater the larger D_A and $C_A(0)$. This explains why phosphorus diffusion exhibits prominent profile tails and produces a strong emitter push effect, whereas arsenic produces no detectable emitter push effect. The diffusivity of arsenic in silicon is typically two orders of magnitude smaller than the diffusivity of phosphorus. In simple

words, the arsenic diffusion rate is too low to pump in point defects at a sufficiently high rate against the leaking of these point defects back to the silicon surface.

The self-interstitial supersaturation caused by the diffusion process itself must imply that the diffusion takes place through a larger interstitialcy component than the vacancy component. It also implies that diffusion via the interstitialcy mechanism does not occur only because of nonequilibrium point defects injected by extraneous sources such as thermal oxidation. One sees already in this simplified analysis that the super-saturation of self-interstitials (1) increases with the surface impurity concentration ($C_A(0)$ in the numerator); (2) decreases with the local impurity concentration ($C_A(x)$ in the denominator), and (3) increases with the impurity diffusivity. For small impurity diffusivity, as the term $(D_s/D_A)^{1/2} C_s$ becomes dominant in both the numerator and the denominator, C_i/C_i^* tends to unity. These are the characteristics of the nonlocal effect we already mentioned, which also explain why this effect is small for such a slow diffusant as arsenic in silicon (compare Figs. 23 and 24).

A more realistic boundary condition may not be $C_i(0) = C_i^*$, but

$$D_i \left(\frac{\partial C_i}{\partial x} \right)_{x=0} = k_{\text{ann}} [C_i(0) - C_i^*] \quad (106)$$

where k_{ann} is the rate constant of the surface annihilation of excess self-interstitials. The surface concentration is then not known a priori, but is given by

$$C_i(0) = C_i^* \frac{1}{1 + \frac{D_i/k_{\text{ann}}}{\alpha + C_A(0)} \left(\frac{\partial C_A}{\partial x} \right)_{x=0}} \quad (107)$$

which now replaces C_i^* in Eq. (104). But, unlike Eq. (104), Eq. (107) is not an explicit expression of the local dopant concentration dependence of C_i . Its usage can, however, be implemented by successive approximations.

The above analysis is intended only as a demonstration of how diffusion pumping of point defects can lead to point defect supersaturation, and how the supersaturation can be fed back to enhance the diffusion process itself. In semiconductors, point defects have charge states within the bandgap, and are charged according to Fermi statistics. It has been shown thermodynamically [5–8] that the equilibrium concentration of neutral point defects is a constant (at a given temperature and stress) and is unaffected by the Fermi level; the equilibrium concentrations of charged point defects increase or decrease from the concentration of neutral point defects according to the Fermi statistics. This underlies the well-established diffusion modeling under equilibrium conditions. Under nonequilibrium conditions one can no longer assume the concentration of neutral point defects to be a constant. Because the electronic transition between various charge states is so fast compared with defect migration, the relative concentrations of point defects of various charge states can be assumed to be in proportionality given by the Fermi statistics. But it is the total concentration of a point defect species (the vacancy or the self-interstitial), inclusive of all charge states, that is subject to the description by the equation of continuity. With this understanding, the inclusion of all charge states introduces only trivially more complications, by requiring an effective diffusivity and reaction rate constants (or equilibrium constants, through binding energies) of a defect species to be replaced by charge state weighted averages, which is a function of the local Fermi level, and hence a function of local dopant

concentrations. We still need only the same number of equations of continuity, for the same number of defect species. Some quasi-steady state analyses have been given for phosphorus and boron diffusion in silicon by including the consideration of defect charge states [319–321]. The quasi-steady state approach seems to accomplish a reasonably adequate job in modeling the special features of phosphorus diffusion, especially the long diffusion tail. No doubt it should also be able to model the emitter push effect very well. The plateau of the high concentration surface region, however, has not been simulated satisfactorily.

9.3.3. Limitation of quasi-steady state analysis of chemical pump

In many situations, the quasi-steady state approximation is invalid. We see that Eq. (100), the basis of quasi-steady state approximation, contradicts the flux sum rule of Eq. (90). It amounts to neglecting the flux of lattice atoms, J_B , as well as the Kirkendall effect $u(t)/\Omega$. The Kirkendall effect may well be negligible in silicon. But J_B will not be negligible, and may actually be dominant, under strongly nonequilibrium conditions such as in ion implantation, short time diffusion, rapid thermal/optical annealing, etc. Proton irradiation-induced redistribution of impurities in the experiments of Loualiche et al. [315] and Kozlovskii et al. [316] provides a prime example of a case in which the quasi-steady state approach is invalid. It should be apparent that, in this case, the fluxes of self-interstitials, vacancies and defect–impurity complexes will all be in the same direction, all flowing away from the position of point defect concentration peak. Opposite and balancing all these fluxes must then be the flux of the host atoms, J_B . In Section 11 we will discuss an especially simple method, based on a simplified concept, for the analysis of the case of radiation-induced impurity redistribution.

A quasi-steady state may not be attained also in situations like short time or low-temperature annealing. Furthermore, in these situations, the mobile defect–impurity complex AI (or AV) will not be in local equilibrium with its parents. This topic is discussed in detail in Section 10.

The more generally valid way to model the diffusion process in presence of nonequilibrium point defects is to describe it as a set of fully coupled diffusion–reaction equations involving the dopants, the point defects, and their complexes such as outlined by Hu et al. [298]. Their formulation will be expanded into a more general form, and then discussed more fully in the next section. In such formulations, the diffusing intermediate complexes of the impurity atoms and the point defects are no longer assumed to be in local equilibrium with their parents. The complexity of such formulations inevitably requires numerical solutions. Numerical solutions of some cases of such reaction–diffusion systems have recently been reported by Mulvaney and Richardson [322], Baccus et al. [87], Hane and Matsumoto [88] and others. Undoubtedly, this approach will become the trend and the standard of the future.

10. Point defects, impurities and impurity–defect complexes in diffusion

10.1. Diffusion via defect–impurity complexes

The activation energy of diffusion of substitutional impurities is typically smaller than that of self-diffusion. This can be explained only if there exists a binding energy between the impurity atom and the native point defect that mediates the diffusion [8,37,317].

Consequently, the impurity atom will migrate as an intermediate species a substantial distance before reversion to ancestral form [8,37]. It should be stated at the outset that the kick-out mechanism of diffusion could also provide a different, and possibly also smaller, activation energy for impurity diffusion than for self-diffusion. In this paper, the choice of the defect-impurity pair as the diffusing intermediate is for convenience and to avoid duplication of analyses; it is not meant to exclude the viability of the kick-out mechanism. The mathematical formulations for diffusion via paired complexes and via the kick-out mechanism are identical (compare, in later sections, Eqs. (118) and (119) for the former, and Eqs. (128) and (129) for the latter). When in local equilibrium, the concentration of the intermediate complex is given by $C_{AX} = KC_A C_X$, in accordance with the following reaction:



Here, A, X and AX are respectively the impurity, the point defect (either a vacancy, V, or a self-interstitial, I) and their complex. It is customary to assume that the effective diffusivity of the impurity is then given by

$$D_A = D_{AX} \frac{C_{AX}}{C_A} = D_{AX} K C_X \quad (109)$$

where D is the diffusivity, C the concentration and K the equilibrium constant. As modern microelectronic devices tend to be very small, and modern processing tends to be low-temperature and short time, the diffusing intermediate complex will no longer be in local equilibrium. For the annealing of implanted impurities, another common occurrence of nonequilibrium intermediate species, various researchers [323–325] have adopted a time-dependent “effective diffusivity” of

$$D_A(t) = D_A^i + D_A^e e^{-t/\tau} \quad (110)$$

where D_A^i and D_A^e are respectively the “intrinsic” and “enhanced” diffusivities of A and τ is a time constant.

In two processes that are technologically very important now, these intermediate complexes will not be in local equilibrium with their component species: (a) in rapid thermal processing where the diffusion time is short, and (b) in low-temperature annealing of implanted impurities, which are massively complexed with implantation-generated point defects and become an instantaneous source. We will show that the concept of effective diffusivity is no longer valid in these cases. We will present a set of reaction-diffusion equations with which the diffusion of impurities and nonequilibrium point defects can be analyzed rigorously. For insight, we will present analytical solutions for these two cases, using a reduced set of the reaction-diffusion equations. The solution for the first case, that of a constant diffusion source boundary condition, is valid for all diffusion time; but the short time limit tends to exhibit an exponential tail or its derivative from convolution. This exponential tail is independent of the point defect species and of whether the intermediate is a kicked-out atom or a complexed pair. The solution for the case of annealing with an instantaneous source of implanted atoms leads to an asymptotic distribution at long time. Even when the intermediate complexes are in local equilibrium with their parent species, the traditional form of effective diffusivity is incomplete.

10.2. Formulation of diffusion via complex intermediates

Diffusion via an intermediate complex should rigorously be formulated as a set of simultaneous reaction-diffusion continuity equations. A set of such equations was proposed

by Hu et al. in 1983 [298] as an approach to analyze anomalous diffusion phenomena such as found in phosphorus diffusion in silicon. In that formulation, phosphorus was considered, based on many experimental observations, to proceed via a mainly interstitialcy mechanism. The phosphorus atom was perceived to react with a silicon self-interstitial, producing a phosphorus–self-interstitial complex PI (which may also be called a phosphorus interstitialcy, and was denoted as P_i in the work of Hu et al. [298]). The Hu–Fahey–Dutton formulation can easily be expanded to include the impurity–vacancy complex, AV, for any impurity that diffuses via a dual mechanism. The extended set of simultaneous reaction–diffusion equations are:

$$\frac{\partial C_V}{\partial t} = \frac{\partial}{\partial x} \left(D_V \frac{\partial C_V}{\partial x} \right) - k_{VI}(C_V C_I - C_V^* C_I^*) + k_{AV}^r C_{AV} - k_{AV}^f C_A C_V \quad (111)$$

$$\frac{\partial C_I}{\partial t} = \frac{\partial}{\partial x} \left(D_I \frac{\partial C_I}{\partial x} \right) - k_{VI}(C_V C_I - C_V^* C_I^*) + k_{AI}^r C_{AI} - k_{AI}^f C_A C_I \quad (112)$$

$$\frac{\partial C_{AV}}{\partial t} = \frac{\partial}{\partial x} \left(D_{AV} \frac{\partial C_{AV}}{\partial x} \right) - k_{AV}^r C_{AV} + k_{AV}^f C_A C_V \quad (113)$$

$$\frac{\partial C_{AI}}{\partial t} = \frac{\partial}{\partial x} \left(D_{AI} \frac{\partial C_{AI}}{\partial x} \right) - k_{AI}^r C_{AI} + k_{AI}^f C_A C_I \quad (114)$$

$$\frac{\partial C_A}{\partial t} = k_{AV}^r C_{AV} - k_{AV}^f C_A C_V + k_{AI}^r C_{AI} - k_{AI}^f C_A C_I \quad (115)$$

k_{AX}^f and k_{AX}^r are the forward (complexing) and the reverse (dissociating) rate constants for the reaction of Eq. (108). X is to be replaced by V or I in the above set of more specific equations. The reason for putting the diffusivities of various species inside the parentheses is to take care of the situation where the silicon is inhomogeneously doped (say, by dopant diffusion) such that the average defect diffusivity weighted by various charge states becomes position dependent. The reaction terms included in the above equations are by no means complete. For example, there may also exist indirect recombinations between the vacancy and the self-interstitial via an impurity atom, for example,



Impurities besides A, such as oxygen, which is present in concentrations as high as 10^{18} atoms cm^{-3} , may also participate as a recombination intermediary.

By adding bulk generation terms when present, such as from precipitation or particle irradiation, and by applying appropriate boundary conditions, the solution of the above reaction–diffusion equations should provide answers to:

1. anomalous diffusion phenomena, such as the emitter push effect and the phosphorus profile tail (the chemical pump effect)
2. effect of particle irradiation on impurity distribution (vacancy/interstitialcy winds)
3. short time diffusion
4. annealing of implanted dopants

We expect the coupled equations to be solved numerically, something which is beyond the scope of this article. Numerical solutions of such reaction-diffusion systems (somewhat different from given above) have been made only very recently [87,322,326].

10.3. Analytical solution of diffusion with nonequilibrium intermediates

Our objective here is to investigate the characteristics of diffusion when the intermediate mobile species is not in local equilibrium. To gain an insight into diffusion of this kind, it suffices to choose an impurity which diffuses via only one defect species, the vacancy or the interstitial, which we shall denote by X. We will also assume that the impurity diffusion is not that fast to cause a significant departure from equilibrium point defect concentrations. Assuming that C_X is approximately constant, Eqs. (118) and (119) become linear and uncoupled from the continuity equation for C_X . Then we need consider only the following subset of the above reaction-diffusion equations:

$$\frac{\partial C_{AX}}{\partial t} = D_{AX} \frac{\partial^2 C_{AX}}{\partial x^2} + k_{AX}^f C_A C_X - k_{AX}^r C_{AX} \quad (118)$$

$$\frac{\partial C_A}{\partial t} = k_{AX}^r C_{AX} - k_{AX}^f C_A C_X \quad (119)$$

D_{AX} and D_X are considered to be constant.

10.3.1. Solution for constant source

We consider here the boundary condition of a constant source. The simultaneous equations (4) and (5) can be solved analytically by means of Laplace transformation. A somewhat tricky inverse transformation has been obtained by Lapidus and Amundson [327] for the mathematically analogous problem of adsorption in ion exchange and chromatographic columns (without the flow term in our case). For the following initial and boundary conditions

$$C_A(0, t) = C_A^0 \quad (120a)$$

$$C_{AX}(0, t) = C_{AX}^0 = K C_A^0 \quad (120b)$$

$$C_A(x, 0) = C_{AX}(x, 0) = 0 \quad \text{for } x > 0 \quad (120c)$$

the solution is

$$\frac{C_{AX}(x, t)}{C_{AX}^0} = F(x, t) + k_{AX}^f C_X \int_0^t F(x, t') dt' \quad (121)$$

$$C_A(x, t) = k_{AX}^r \exp(-k_{AX}^f C_X t) \int_0^t \exp(k_{AX}^f C_X t') C_{AX} dt' \quad (122)$$

with

$$F(x, t) = \exp(-k_{AX}^f C_X t) \int_0^t I_0[2[k_{AX}^r k_{AX}^f C_X t'(t-t')]^{1/2}] \times \exp\left[-\frac{x^2}{4D_{AX}t'} - (k_{AX}^r - k_{AX}^f C_X)t'\right] t'^{-3/2} dt' \quad (123)$$

Here, I_0 is the modified Bessel function of the first kind, of order zero. One may take $C_{AX}^0 = (k_{AX}^f C_X / k_{AX}^r) C_A^0$. For the case $k_{AX}^f C_X = 0$, $I_0(0) = 1$, and Eq. (123) integrates to

$$\begin{aligned} \frac{C_{AX}}{C_{AX}^0} &= \frac{1}{2} e^{-x/L} \operatorname{erfc} \left[\frac{x}{2(D_{AX}t)^{1/2}} - (k_{AX}^r t)^{1/2} \right] \\ &\quad + \frac{1}{2} e^{x/L} \operatorname{erfc} \left[\frac{x}{2(D_{AX}t)^{1/2}} + (k_{AX}^r t)^{1/2} \right] \end{aligned} \quad (124)$$

where $L = (D_{AX}/k_{AX}^r)^{1/2}$. But, for the case of $k_{AX}^f C_X = 0$, there is a simpler way to solve the problem, because Eq. (118) then reduces to the irreversible reaction-diffusion equation

$$\frac{\partial C_{AX}}{\partial t} = D_{AX} \frac{\partial^2 C_{AX}}{\partial x^2} - k_{AX}^r C_{AX} \quad (125)$$

which is uncoupled from Eq. (119) and readily yields the textbook solution of Eq. (124) [328]. A negligible local C_A would have the same effect as a negligibly small $k_{AX}^f C_X$. Typically, for a diffusion time of the order of 1 s or more, $(k_{AX}^r t)^{1/2}$ would dominate the arguments of the erfc functions in Eq. (124), which then approaches a steady state distribution of

$$C_{AX} = C_{AX}^0 e^{-x/L} \quad (126)$$

C_A corresponding to this steady state C_{AX} is then obtained by setting $k_{AX}^f C_X = 0$ in Eq. (122) and integrating:

$$C_A = k_{AX}^r \int_0^t C_{AX} dt = C_A^0 k_{AX}^f C_X t e^{-x/L} \quad (127)$$

C_A is seen to have an exponential distribution, and the entire profile increases linearly with time. Numerical simulations have shown approximately such a feature within a certain time limit [87]. Obviously, C_A cannot increase indefinitely. As local C_A builds up with time, the irreversible reaction approximation becomes invalid. Then, only Eqs. (121)–(123), the solutions to the reversible reaction-diffusion equations Eqs. (118) and (119), will give correct results.

The exponential tail in an impurity profile is a hallmark of diffusion in which the intermediate species far exceeds local equilibrium. In 1973, Schwettmann [329] showed that low-temperature annealing of implanted arsenic in silicon creates exponential tails in the arsenic profiles. He recognized that the exponential tails are caused by a transient intermediate species, which he speculated [329] to be the interstitial arsenic atom A_i . His result is shown in Fig. 26. Recently, Cowern et al. reported [330–333] a similar observation of exponential tails in short time, low-temperature annealing of boron in silicon. They, too, recognized that this is caused by a transient intermediate species, which they suggested to be an interstitially wandering impurity atom after being kicked out from its substitutional site by a self-interstitial.

For the kick-out mechanism, one may write:

$$\frac{\partial C_{A_i}}{\partial t} = D_{A_i} \frac{\partial^2 C_{A_i}}{\partial x^2} + k_{AX}^f C_A C_I - k_{AX}^r C_{A_i} \quad (128)$$

$$\frac{\partial C_A}{\partial t} = k_{AX}^r C_{A_i} - k_{AX}^f C_A C_I \quad (129)$$

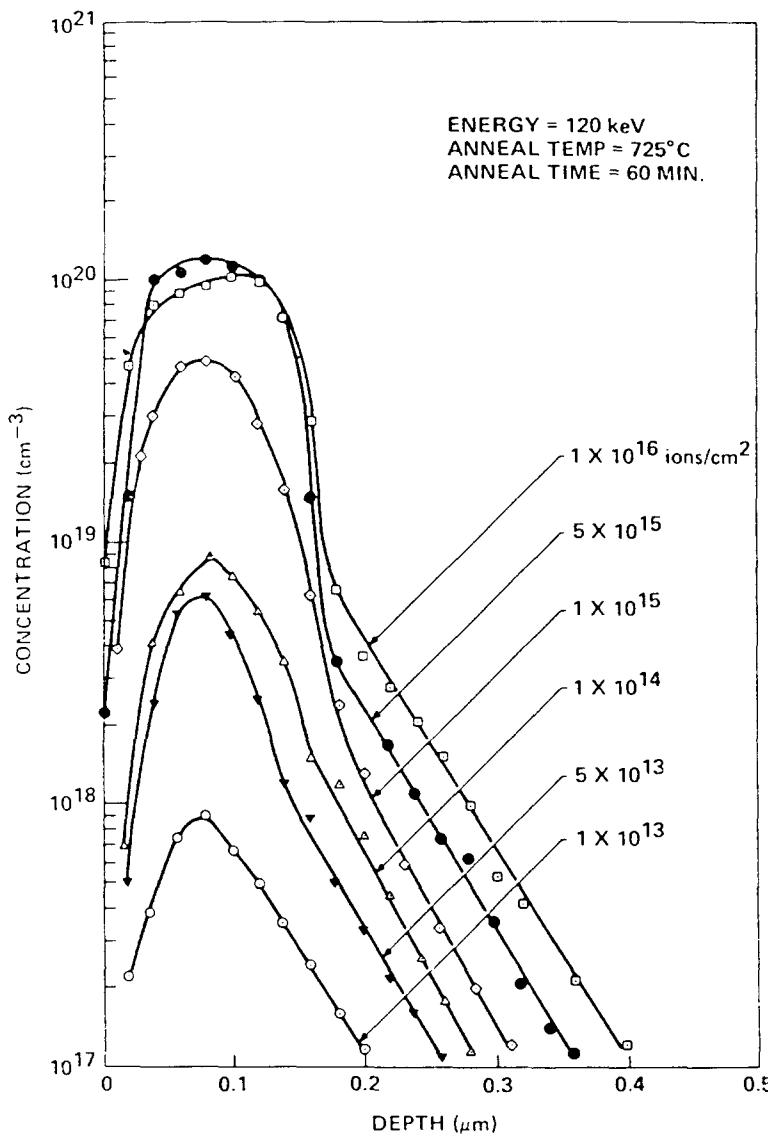


Fig. 26. Effect of implantation dose on the penetration of the electrically active arsenic tail after annealing at 725 °C for 60 min. Notice that the concentration at the kink increases with the dose, rather than being a function of temperature supposed of a solubility limit (taken from Schwettmann [329]).

It is clear that Eqs. (128) and (129) are exactly identical to Eqs. (118) and (119), with the renaming of AX by A_i and X by I . This was also recognized and discussed by Cowern et al. [330]. For a given set of initial and boundary conditions, the solutions to both sets of equations are exactly identical. Hence, the same $C_A(x, t)$ is obtained regardless of whether the diffusing intermediate is A_i or AX .

Baccus et al. [87] have numerically solved a superset of the Hu–Fahey–Dutton formulation, including charge carriers, of diffusion–reaction equations without approximations. They assumed AI and AV , but not A_i , as the intermediates. The exponential spread for short diffusion time in their simulation is clearly illustrated in Fig. 2(b) of their paper [87]. Hence the exponential spread of impurity profiles is not evidence of boron diffusing via the kick-out mechanism or, for that matter, an interstitialcy mechanism. Cowern et al. [331,333]

analyzed their experiments of boron diffusion in oxidizing ambients, and found that only their parameter g (equivalent to k_{AX}^f of Eq. (128)) is enhanced, while λ (equivalent to L of Eq. (127)) remains essentially constant. They correctly interpreted this to be indicative of a diffusion process mediated predominantly by the silicon self-interstitial. This process can take place via one of two mechanisms: an interstitialcy mechanism in which a self-interstitial and a substitutional boron atom diffuse as a pair, or a kick-out mechanism in which the kicked out boron atom wanders as an isolated interstitial through the lattice. From their articulations [330–333], Cowern et al. appear to advocate the kick-out mechanism. But no experimental observations to this date, including those of their own, can lend themselves to decide which of these two mechanisms actually operates.

Intending for conceptual microscopic visualization, Cowern et al. called the mechanism of diffusion via an intermediate species “intermittent diffusion” [332], in which an impurity atom switches alternately between a mobile intermediate and an immobile substitutional. Each switch to a mobile intermediate, which may wander through the lattice by taking multiple elementary random walk steps, is referred to as a “migration step”. We note, however, that this is already implicit (or perhaps explicit) in our reaction-diffusion continuity equations, e.g. Eqs. (118) and (119). When the reaction is formulated as irreversible by omitting the term $k_{AX}^f C_A C_X$ in Eqs. (118) and (119), each intermediate complex is allowed just a single migration step, after which the complex will break up forever. Formulation of a reversible reaction allows switching back and forth between A and AX once or more times as weighted by Poisson’s distribution. Clearly, a short annealing time tends to limit each intermediate to just one migration event, and consequently causes the impurity to spread exponentially.

In accord with the purpose of short-time annealing, which is to minimize the spread of the impurity profile, the amount of impurity getting into the exponential tail should be very small compared to the amount of impurity in the initial profile. Hence, provided the initial profile is sharp, the boundary condition is reasonably approximated by a constant source.

10.3.2. Low-temperature annealing of predeposited impurities – the case of an instantaneous source

Two technologically important such cases are the annealing of implanted dopants [323–325,334–336], and the low-temperature annealing of dopants deposited at high temperatures [337,338].

Shibayama et al. [337] carried out low-temperature annealing of an arsenic profile predeposited at 1000 °C. Fig. 27(a) shows the further movement of the arsenic profile after heat treatments at 700 °C in nitrogen for various lengths of time. They extracted the effective diffusion coefficient from the data of Fig. 27(a) and plotted it against heat treatment time, as shown in Fig. 27(b) [337]. It can be seen that the effective diffusivity is orders of magnitude higher than expected by extrapolation from its normal diffusion coefficient at high temperatures (dashed line). In their second series of experiments [338], they again deposited arsenic at 1000 °C, but on a substrate with a boron base, forming an n⁺-p-n structure. They then carried out heat treatments at various temperatures and investigated the evolution of the arsenic profile as well as the movement of the boron base underneath it. The result is shown in Fig. 28(a). Fig. 28(b) shows the effective diffusion coefficients of arsenic and boron underneath the arsenic-diffused layer as a function of temperature for 24 h heat treatments. One sees a greatly enhanced movement of the boron base under the arsenic

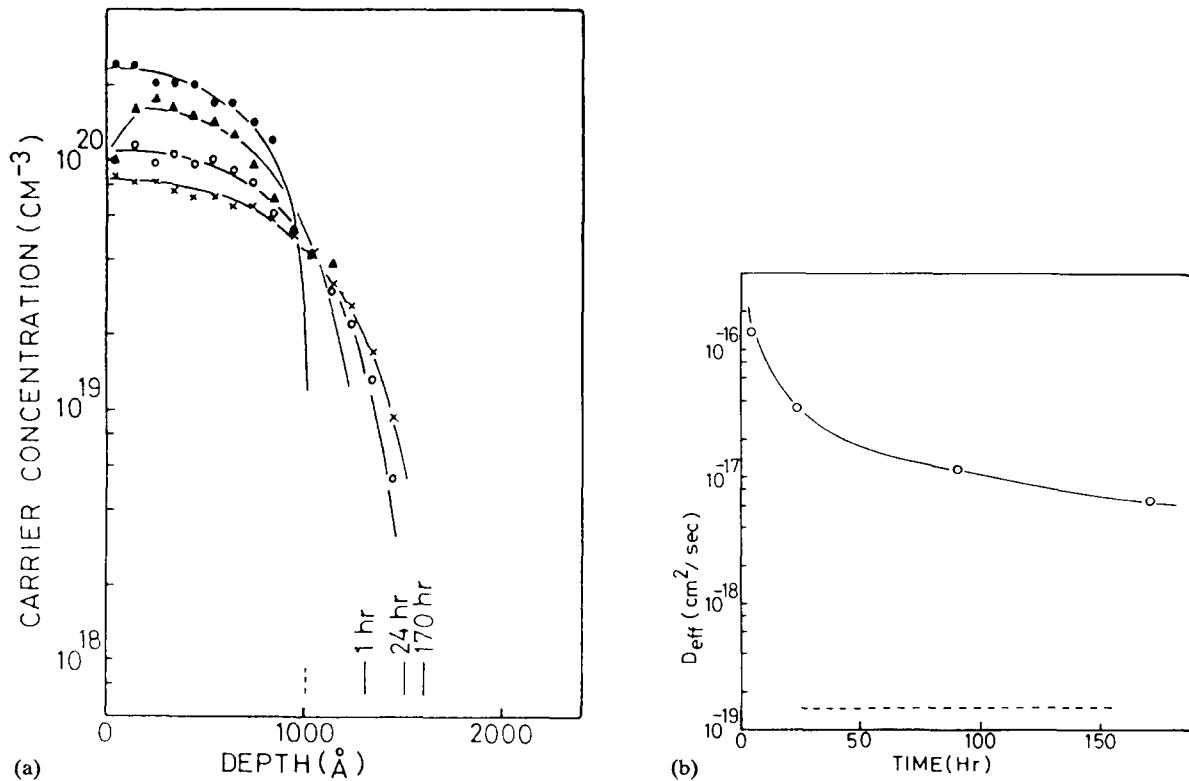


Fig. 27. (a) Effect of heat treatments at 700 °C in nitrogen on carrier profiles of an arsenic 1000 °C diffused layer. Closed circles: before heat treatment; triangles, solid circles and crosses are respectively heated for 1 h, 24 h and 170 h. (b) The effective diffusion coefficients of arsenic as a function of heat treatment time at 700 °C, extracted from the data of Fig. 27(a). The dashed line indicates the normal diffusion coefficient as extrapolated from high-temperature data (taken from Shibayama et al. [337]).

emitter: the emitter push effect. Having a very small diffusivity (two orders of magnitude smaller than phosphorus), arsenic does not normally exhibit a significant emitter push effect, for the reason already discussed in Section 9.3.2, especially following Eq. (105). Apparently, low-temperature annealing can release a very large amount of excess point defects, which decay relatively slowly.

A basically similar phenomenon was observed by Park and Law [339] in their study of low-temperature annealing of low-dose implanted phosphorus. They found a greatly enhanced transient diffusivity, which decays with time. The enhancement is greater and the decay is slower at lower annealing temperatures. This can be understood in terms of frozen-in defect-impurity complexes which dissociate more slowly at lower annealing temperature. According to Hu [37], the migration energy of a defect-impurity pair is generally smaller than the activation of dissociation, roughly by an amount equal to the potential energy at the third coordination site. This is because in migration, the point defect need only to go as far as the third coordination site of the impurity atom, while in dissociation it needs to leave the impurity atom entirely. Therefore, at lower temperatures, the rate of migration of the defect-impurity complex is not reduced as much as the rate of dissociation, leading to a net increase of effective diffusivity.

In all these cases, the primordial defect-dopant complexes at the beginning of annealing were those frozen in from ion implantation, or from high-temperature deposition. The time-

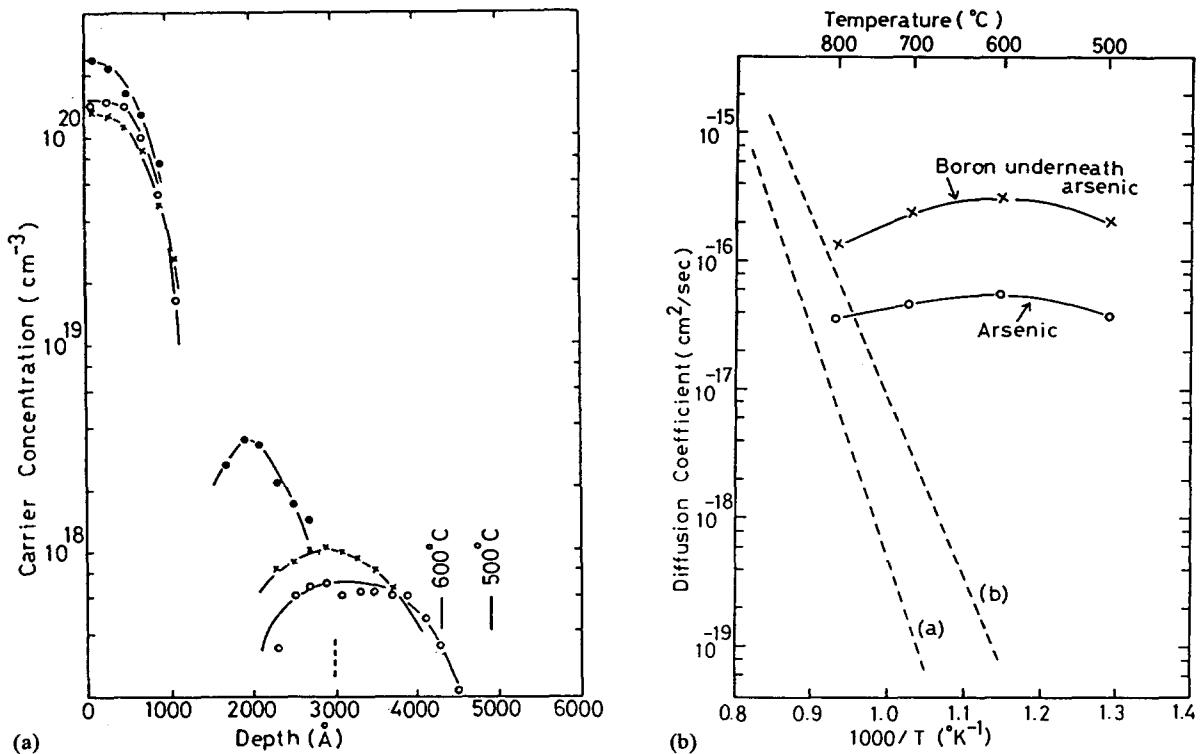


Fig. 28. (a) Effect of heat treatment at low temperature on carrier profiles of an n^+ -p-n structure. The n- and p-type dopants are respectively arsenic and boron. The arsenic was deposited at 1000 °C. Solid circles: before heat treatment; crosses: heat treated at 600 °C for 45 h; open circles: heat treated at 500° for 336 h. The three vertical lines mark the junction depths before and after heat treatments. (b) Effective diffusion coefficients of arsenic and boron underneath the arsenic-diffused layer as a function of temperature. Heat treatment is carried out for 24 h in nitrogen. Dashed lines (a) and (b) are respectively arsenic and boron diffusion coefficients extrapolated from high temperatures (taken from Shibayama et al. [338]).

zero (during annealing) concentration of these complexes far exceeds its equilibrium value at the annealing temperature. During annealing, these complexes will break up essentially irreversibly. Therefore, Eq. (125) becomes an accurate formulation of the reaction-diffusion process for any t . (Here we will restrict our analysis to cases in the absence of implantation-induced amorphization.) It can be shown that, for an impulse of unit strength of defect-impurity complexes (cm^{-2}), at $x=0$ in an infinite space, the solution for Eq. (125) is

$$C_{AX}(x, t) = \frac{1}{2(\pi D_{AX}t)^{1/2}} \exp(-k_{AX}^r t) \exp\left(-\frac{x^2}{4D_{AX}t}\right) \quad (130)$$

Compare Eq. (130) with Eq. (126). In contrast with constant source diffusion, instantaneous source diffusion does not lead to a steady state distribution of the mobile intermediate. The impurity profile corresponding to C_{AX} of Eq. (130) is

$$C_A(x, t) = k_{AX}^r \int_0^t C_{AX} dt = \frac{1}{4L} \left\{ 2e^{-|x|/L} - e^{x/L} \operatorname{erfc}\left[(k_{AX}^r t)^{1/2} + \frac{x}{2(D_{AX}t)^{1/2}}\right] \right. \\ \left. - e^{-x/L} \operatorname{erfc}\left[(k_{AX}^r t)^{1/2} - \frac{x}{2(D_{AX}t)^{1/2}}\right] \right\} \quad (131)$$

As $t \rightarrow \infty$, C_A approaches a stationary distribution of

$$C_A(x, \infty) = \frac{1}{2L} e^{-|x|/L} \quad (132)$$

The existence of an asymptotic distribution is the most distinct feature of low-temperature annealing of implanted profiles [325,334–336], and of profiles initially formed by diffusion at high temperatures [337,338]. The evolution of an initial impurity distribution $C_{AX}(x, 0)$ is given by the convolution $C_{AX}(x, 0)$ with $C_A(x, t)$ of Eq. (131).

It is of considerable practical interest to analyze the asymptotic distribution of an initial gaussian profile, as implanted with a range straggle of ΔR_p . For simplicity, we designate the position of the gaussian peak as $x=0$ in an infinite space. Let the impurity dose be Φ_A . Assume that a fraction of A exists as AX, the fraction being determined by the equilibrium constant of mass action at the end of implantation, or high-temperature deposition. We may then assign ϕ_A and ϕ_{AX} as the doses of A and AX, respectively. Let $\lambda = 2^{1/2} \Delta R_p$. The asymptotic profile after annealing is then given by

$$\begin{aligned} C_A(x, \infty) &= \frac{\phi_A}{\pi^{1/2} \lambda} e^{-(x/\lambda)^2} + \frac{\phi_{AX}}{2\pi^{1/2} \lambda L} \int_{-\infty}^{\infty} e^{-(x'/\lambda)^2} e^{-|x-x'|/L} dx' \\ &= \frac{\phi_A}{\pi^{1/2} \lambda} e^{-(x/\lambda)^2} + \frac{\phi_{AX}}{4L} e^{(\lambda/2L)^2} \left[e^{-x/L} \left(2 - \text{erfc} \left(\frac{x}{\lambda} - \frac{\lambda}{2L} \right) \right) + e^{x/L} \text{erfc} \left(\frac{x}{\lambda} + \frac{\lambda}{2L} \right) \right] \end{aligned} \quad (133)$$

Fig. 29 shows the asymptotic behavior of the redistribution of an implanted profile [336].

The profile given by the above equation exhibits a kink, another notable feature of low-temperature annealing. It is commonly held that the kink indicates the solubility limit. Now, individual atoms would not know that they are in a supersaturated solution until

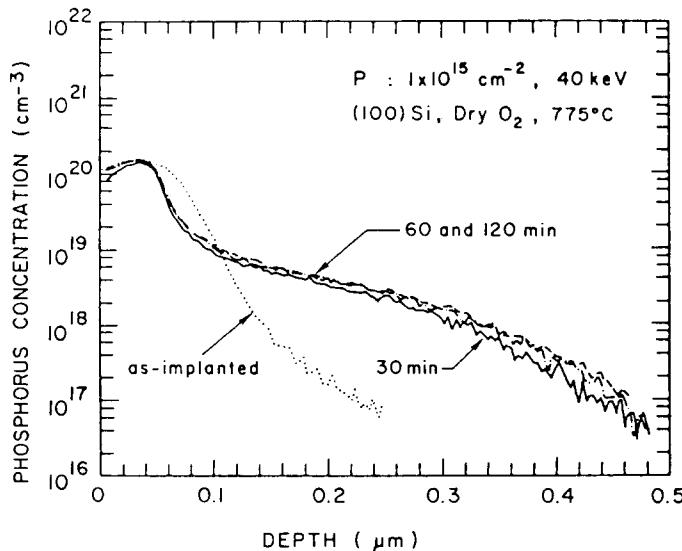


Fig. 29. SIMS-measured profiles for implanted phosphorus ($1 \times 10^{15} \text{ cm}^{-2}$, 40 keV), and annealed in dry oxygen at 775°C for 30, 60 and 120 min. Notice that the profile becomes stationary after annealing for between 30 and 60 min (taken from Kim et al. [336]).

running into precipitates. But the nucleation of precipitates usually requires an incubation period far longer than the short annealing time of our interest; otherwise, the annealing would undesirably lead to dopant inactivation. In our model here, atoms not already complexed with point defects at the start of annealing would never move during that annealing. The concentration at the kink in the resultant profile is therefore determined by the concentration of defect-impurity complexes originally present, not by the annealing conditions, e.g. the annealing temperature. This inference is supported by Schwettmann's [329] experimental results showing the kink concentration to increase with implant dose, at the same annealing temperature of 725 °C. Note also that the profiles in Schwettmann's experiments were measured electrically, thus voiding the precipitation argument for the kink. Note further that the exponential tails reported by Schwettmann are apparently stationary distributions in the manner of Eq. (133), since they are stable during annealing up to 1000 °C [329]. Fig. 29, from Kim et al. [336], also shows such a kink.

10.4. On the effective diffusivity concept in nonequilibrium

We will now show that the concept of "effective diffusivity" is inapplicable in transient diffusion. Fundamentally, the effective diffusivity such as given by Eq. (110) is inapplicable because the time dependence of local C_{AX} is presumed based on a first-order reaction without diffusion. The lack of self-consistency is inevitable in treating the diffusion separately. Let us examine the resultant diffusion profile from the use of the effective diffusivity of Eq. (110). Define a new independent variable (nonlinear time) θ as

$$\theta = \frac{D_A^i}{D_A^e} t + \tau(1 - e^{-t/\tau}) \quad (134)$$

With the effective diffusivity defined by Eq. (110), the diffusion equation can be written as

$$\frac{\partial C_A}{\partial \theta} = D_A^e \frac{\partial C_A^2}{\partial x^2} \quad (135)$$

which is mathematically identical to Fick's second law with linear time. Noting $\theta=0$ for $t=0$ and $\theta \rightarrow \infty$ as $t \rightarrow \infty$, it can easily be shown that the same initial gaussian profile as considered above will evolve into

$$C_A(x, \theta) = \frac{\Phi_A}{2[\pi(\lambda^2 + 4D_A^e \theta)]^{1/2}} \exp\left(-\frac{x^2}{\lambda^2 + 4D_A^e \theta}\right) \quad (136)$$

If D_A^i is negligibly small and can be dropped as an approximation, then, setting θ to τ in the above expression gives the asymptotic distribution. Thus, firstly, the effective diffusivity model does not produce a kink in the profile as commonly observed. Secondly, the distribution of Eq. (136) is gaussian, regardless of the parameter λ in the original profile. In contrast, the distribution of Eq. (133) has a distinct heritage of exponential trait. For very small λ , Eq. (133) reduces to

$$C_A(x, \infty) = \phi_A \delta(x) + \frac{\phi_{AX}}{2L} e^{-|x|/L} \quad (137)$$

Even when the intermediate species is in local chemical equilibrium, the concept of "effective diffusivity" is not strictly correct, and must be used with caution. From Eqs. (118)

and (119), it follows that

$$\frac{\partial}{\partial t} (C_A + C_{AX}) = D_{AX} \frac{\partial^2 C_{AX}}{\partial x^2} \quad (138)$$

Since C_{AX} is much smaller than C_A , it can be omitted from Eq. (122). Putting $C_{AX} = KC_A C_X$ into Eq. (122), one obtains

$$\frac{\partial C_A}{\partial t} = \frac{\partial^2}{\partial x^2} (D_{AX} K C_X) C_A \quad (139)$$

With the effective impurity diffusivity given by Eq. (109), Eq. (139) implies the following flux expression

$$J = - \frac{\partial}{\partial x} (DC) \quad (140)$$

instead of the traditional flux expression of $J = -D(\partial C / \partial x)$. Flux expression in the form of Eq. (140) has been used by some scientists dealing with hot electron phenomena [340,341]. It would seem that Eq. (140) can be dismissed simply by reductio ad absurdum: If Eq. (140) were valid, then a mere existence of a nonuniform diffusivity, as in composite materials or graded structures, will cause particle fluxes even if there is no concentration gradient. But a deeper reflection shows this to be a more complex issue. This subject has been discussed by Landauer and Woo [342] and by Landsberg and Hope [343]. It is also related to the Ito–Stratonovich dilemma [344]. The simple answer to this problem is that there is nothing wrong with the form of Eq. (140) mathematically, and it can also be derived from the Fokker–Planck equation. Physically, it simply means that an “effective diffusivity” may implicitly contain a driving force in the form of chemical potential gradient that gives rise to the position-dependent effective diffusivity. One example, to be discussed in detail in a later section, is the radiation-induced redistribution of an initially uniform impurity concentration. The position-dependent effective diffusivity in that case comes from the formation of local mobile complexes with highly localized point defects.

Eq. (139) can be written as

$$\frac{\partial C_A}{\partial x} = \frac{\partial}{\partial x} \left[D_{AX} K \left(C_X \frac{\partial C_A}{\partial x} + C_A \frac{\partial C_X}{\partial x} \right) \right] \quad (141)$$

Thus, the effective impurity diffusivity as traditionally used amounts to omitting the term $C_A(\partial C_X / \partial x)$ from the above expression. This term may not always be negligible. A gradient of C_X may arise from chemical pump, as discussed in Section 9. We see in the case of phosphorus diffusion a large supersaturation of self-interstitials ($C_I/C_I^* \approx O(10^2)$) at a depth of the phosphorus diffusion length, about 1 μm or a little more, and decaying toward the surface. In this case, the only way to calculate the diffusion profile is to abandon the effective diffusivity approach and instead solve the reaction–diffusion equations as discussed in Section 9, the chemical pump, because the gradient $\partial C_X / \partial x$ cannot be obtained otherwise. $\partial C_X / \partial x$ can also arise from other sources, such as the effect of the Fermi level on the equilibrium concentrations of point defects [6,7,289], and in particle irradiation. In the case of the Fermi level dependence of equilibrium point defects, $\partial C_X / \partial x$ can be reduced to a function of local dopant concentration, and the concept of effective diffusivity is then permissible. The problem, however, is that one cannot categorically decide that the chemical pump effect is negligible in a given case.

10.5. An incidental note on rapid thermal processing

Lastly, we should note that in nonisothermal diffusion such as rapid thermal processing, the traditional solution of diffusion problems by using a time-averaged diffusivity, as widely adopted [345,346], is no longer valid when the intermediate species departs from local equilibrium. The validity of the time-averaged diffusivity rests on the condition that the diffusion equation is Fickian (without any defect-impurity reaction term), and that the temperature dependence of the effective diffusivity can be separated from other dependencies (such as concentration or position). Only that condition allows $D_A(T)$ to be grouped with ∂t of the term $\partial C_A/\partial t$ in the continuity equation, with temperature T being a known function of time [347].

11. Effect of particle irradiation on impurity distribution

11.1. Experimental observations and some past interpretations

Silicon device processing at times involves particle irradiation, such as in plasma etching and ion implantation. Radiation from energetic ionic particles produces Frenkel pairs and thereby raises the concentrations of both the vacancy and the self-interstitial above their equilibrium values. This gives rise to radiation-enhanced diffusion (RED) of impurities that diffuse via the vacancy and/or the interstitialcy mechanism. For atomic particles (vis-à-vis electrons, for example), the production of Frenkel pairs is spatially quite nonuniform, and tends to peak around the projected range. This inhomogeneous generation of point defects gives rise to vacancy and self-interstitial fluxes in the outward directions from the position of their peak concentrations. Thus, particle irradiation additionally produces the vacancy wind and/or the interstitialcy wind effect. Earlier studies view the effect of proton irradiation simply in terms of enhanced diffusion [313,314], due mainly to the limitation of the experimental technique used for characterizing impurity profile evolution. More recent studies generally utilize characterization techniques, such as differential capacitance-voltage ($C-V$) and secondary ion mass spectrometry (SIMS), that are capable of fine resolution of impurity profiles. One characteristic of irradiation with protons [311,315,316,348] and heavier ions [349,350] is the creation of a dip in the initial impurity profile. Impurities would diffuse uphill, away from the location where Frenkel defects are generated at peak rate. Fig. 30 [315] shows the redistribution of an initially uniform boron distribution in a substrate subjected to 500 keV H_2^+ irradiation at a substrate temperature of 750 °C, for 20, 40 and 60 min. Notice that the peak-to-valley distance L_{vp} is not time dependent, but the magnitude of the peak is. Fig. 31 [311] shows the redistribution of an implanted phosphorus profile in a wafer subjected to a 80 keV proton bombardment at a substrate temperature of 700 °C for 3, 10, 30 and 120 min. Again notice that the peak-to-valley distance L_{vp} is not time dependent, but the magnitude of the peak is. Heavy ion implantation, such as with nitrogen and argon, gives very similar redistributions of impurity profiles [349,350].

In general, no redistribution of impurity has been observed for substrate temperature held at ≤ 600 °C. The silicon vacancy and the self-interstitial are quite mobile at temperatures well below 600 °C; but their complexes with impurities are not. High-current ion implantation can, if there is no effort to cool the substrate, lead to a temperature increase able to remove most or all of the damage during the early stage of implantation [351]. This single-

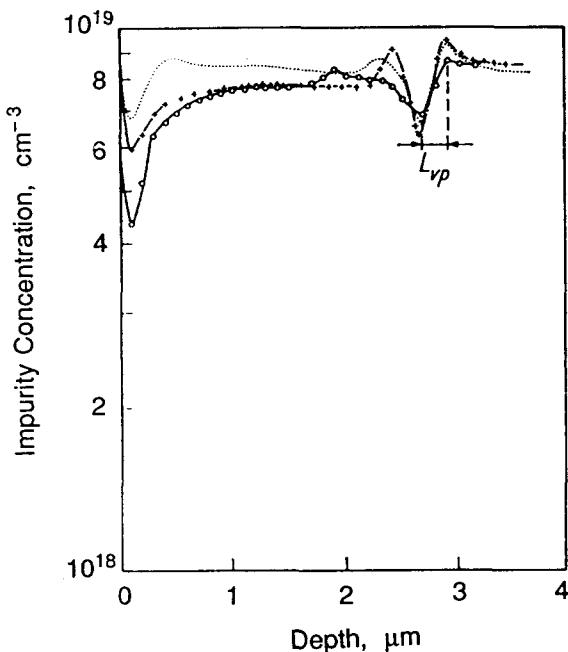


Fig. 30. Initial uniform boron profile ($8.5 \times 10^{18} \text{ cm}^{-3}$) irradiated with H_2^+ ions at 500 keV, 320 nA cm^{-2} , 750°C (+ + +: 20 min, ····: 40 min, ···: 60 min). Notice that the valley-to-peak distance L_{vp} does not change with irradiation time (taken from Loualiche et al. [315]).

step process is called “self-annealing implantation” (SAI). Redistribution that is quite similar to intentional particle irradiation has been reported for, for example, arsenic implantation [352]. Under the usual implantation condition in which the substrate is cooled, it may be expected that the frozen defect–impurity complexes would become mobile during the subsequent annealing stage, causing an impurity redistribution in essentially the same way as during hot implantation. In Section 11.2, we will propose a simple concept to analyze the redistribution of impurity during particle bombardment at suitably high temperatures, or during low-temperature implantation followed by annealing at a suitably high temperature. Recently, Sadata et al. [353] reported the observation of a dip in the boron base in a transistor after arsenic implantation followed by annealing. The magnitude of this dip is larger than normally observed from the effect of the emitter-induced internal electric field based on the Hu–Schmidt theory [289], which usually agrees well with experimental results from diffusion without implantation damage. The dip can also be produced by implantation with neutral ions (that is, ions of Groups IV and VIII elements which become, charge-neutral after entering silicon). Sadana et al. interpreted their result to be one of segregation of boron into the region of implantation-induced damage. We believe that this interpretation is incorrect; we believe that their observations are really the same radiation-induced redistribution reported by many earlier researchers. Giles [354] reported that implantation with argon or silicon causes phosphorus to exhibit uphill diffusion upon annealing in nitrogen at 800°C for 30 min. The phenomenon observed by Giles is essentially similar to that observed by Sadana et al.

Traditionally, the radiation-induced redistribution of impurities is discussed in terms of vacancy or interstitial wind, and the discussions are usually very cursory and qualitative, and have relied on the arguments of flux sum rule. But that approach does not really have a solid foundation, and can easily lead to incorrect conclusions, let alone its incapability

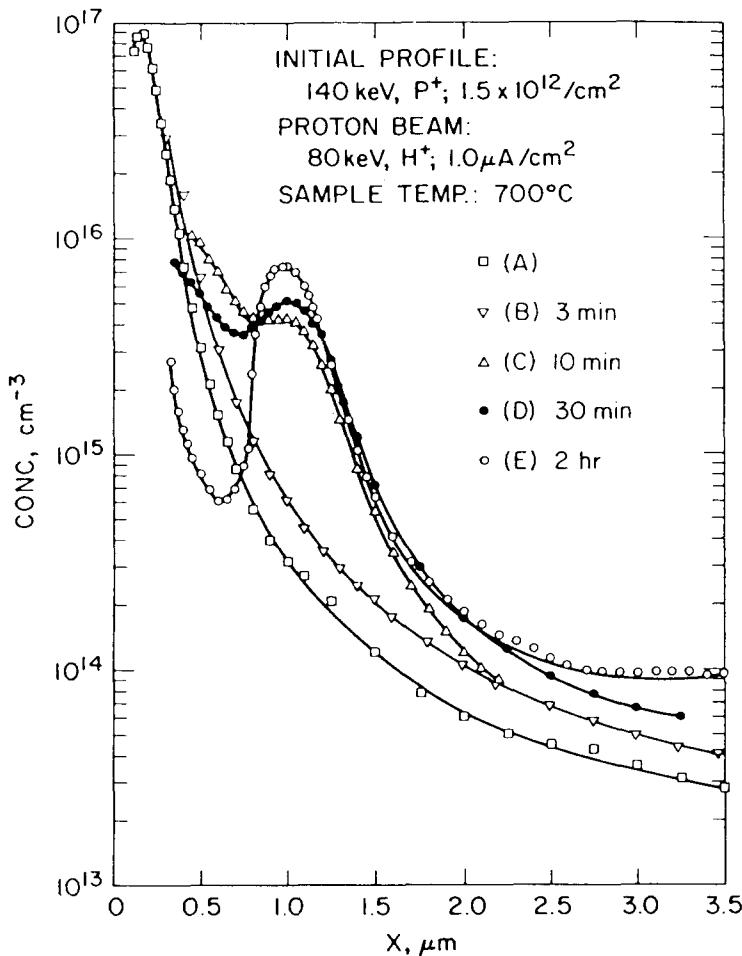


Fig. 31. The redistribution of phosphorus profile due to bombardment with 80 keV protons ($1.0 \mu\text{A cm}^{-2}$, 700°C) for 3 min (B), 10 min (C), 30 min (D) and 2 h. (A) is the initial profile. The profiles were obtained using the differential capacitance-voltage technique. Notice that the peak position does not change with bombardment time, but the valley position appears to move toward the surface, apparently a consequence of the skewed initial profile (vis-à-vis the constant valley position during bombardment of an initially uniform profile shown in Fig. 30) (taken from Akutagawa et al. [311]).

of providing quantitative information. For example, Gösele and Strunk have argued [305] that, based on the flux sum rules of Eqs. (88) and (89) for the vacancy wind and the interstitialcy wind, respectively, the observation of boron diffusing away from the location where vacancies and self-interstitials are generated is evidence of boron diffusing via an interstitialcy mechanism. They argued that the flux sum rule of Eq. (89) demands the impurity and the interstitial fluxes to be in the same direction. Therefore, when excess self-interstitials diffuse away from the peak position of their generation, boron atoms will also diffuse away in the same direction, producing a dip as illustrated in panel c of Fig. 32. On the other hand, Eq. (88) would demand the impurity and the vacancy fluxes to be in the opposite direction. As excess vacancies diffuse away from the peak position of their generation, impurity atoms would diffuse in the opposite direction and pile up at the peak position of Frenkel pair generation, as illustrated in panel b of Fig. 32. These arguments are invalid for a number of reasons:

1. The arguments rest on Eq. (89) which, as already discussed in Section 9.2, is invalid.

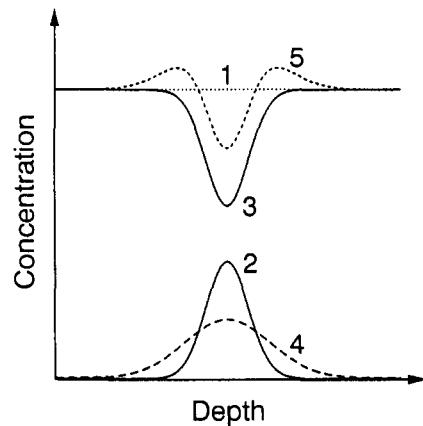
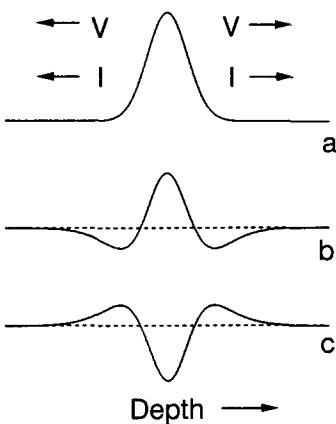


Fig. 32. The vacancy/interstitialcy wind effect in proton irradiation according to Gösele and Strunk [305]. (a) Shows a schematic distribution of Frenkel pairs generated by energetic protons. (b) A hypothetical redistributed impurity profile for a vacancy mechanism of diffusion; this redistribution has never been observed. (c) A redistributed impurity profile that would obtain only in the interstitialcy mechanism of diffusion. However, Kozlovskii et al. [316] have shown that this distribution is obtained for antimony, which is known to diffuse in silicon via a vacancy mechanism.

Fig. 33. A simplified concept of the vacancy/interstitialcy wind effect in proton irradiation. Curve 1: the initially uniform impurity distribution; Curve 2: profile of mobile defect-impurity complexes formed after irradiation; Curve 3: profile of impurity atoms that are left uncomplexed; Curve 4: profile of defect-impurity complexes after spreading out within their lifetime; Curve 5: final redistributed total impurity profile.

2. They ignore the flux of host atoms J_B which may be very large and in the opposite direction.
3. The flux sum rule is never predicated on any particular diffusion mechanism; it is simply based on the conservation of lattice sites.
4. They contradict the experiment of Kozlovskii et al. [316], who observed antimony to diffuse away from the location of point defect generation, exactly in the same manner as boron. From this observation, their arguments would have led to the conclusion that antimony also diffuses in silicon via a dominant interstitialcy mechanism. But it has been well established, through the observation of oxidation-retarded diffusion, that antimony diffusion in silicon is dominated by the vacancy mechanism, with only about a 2% interstitialcy component [3,46].
5. The hypothetical peak in panel b of Fig. 32 has never been observed experimentally in any diffusion system. Thus, it is quite clear that the direction of the flux of substitutional atoms relative to that of the point defect mediating the diffusion is determined mainly by the attractive interaction between the impurity atom and the point defect, as already pointed out in the author's 1969 paper [8].

11.2. A simplified conceptual model of radiation-induced impurity redistribution

The easiest way to gain an insight into the mechanism of the radiation-induced redistribution of impurity atoms is through the following simplified conceptual explanation: refer now to Fig. 33. The entire complex process comprises defect generation, diffusion, recombination and reacting with impurity atoms to produce mobile intermediate species which diffuse and then dissociate. The idealized process is imagined to consist of a sequence of discrete steps. We start with an initially uniform impurity distribution, as in curve 1 of Fig. 33. Irradiation with energetic particles generates a spatial distribution of Frenkel pairs

which then do three things simultaneously: self-annihilation at a finite rate, diffusing away from the peak position of their generation, and reacting with impurity atoms to produce a steady state distribution of mobile intermediate species, as given by curve 2 in Fig. 33. The identity of the intermediate species is unimportant; it may be an impurity-vacancy pair, an impurity-self-interstitial pair or an impurity atom displaced into an interstice by a self-interstitial. Curve 3, obtained by subtracting curve 2 from curve 1, is now the concentration profile of the substitutional impurity atoms that remain isolated and immobile. The intermediate species has a certain lifetime (say, τ_{AX}) in which to diffuse; its profile evolves from curve 2 to curve 4 at the end of its life span. Thereafter it dissociates and reverts to the immobile isolated substitutional species. The resulting profile of the total immobile isolated substitutional impurity atoms is now given by curve 5, which is the sum of curves 3 and 4. (In this picture, the uniform background concentration of thermal point defects is assumed to be so low compared with the radiation-generated point defects that it can be neglected. It can be trivially included if one wishes.)

It is readily apparent from this model that the position of the valley of the redistributed profile should coincide with the peak of the generation profile of point defects. This feature has indeed been observed in many investigations. It contradicts the model of uphill diffusion mechanism in proton-irradiated silicon proposed by Morikawa et al. [312], who found in their calculations the valley position to be substantially shallower than the peak of vacancy distribution. The position of the peaks of the redistributed profile is a complicated function of many parameters, and should be dependent on the profile of Frenkel pair generation, the initial impurity distribution, and, most of all, the reaction and transport parameters of the point defects. In the discretized model given above, the peak occurs at $\partial(C_4 - C_2)/\partial x = 0$, where C_4 and C_2 are the concentration profiles of the intermediate species given as curves 4 and 2 in Fig. 33. When the impurity concentration is substantially higher than the defect concentration, C_2 is a linearly scaled image of the point defect generation profile. We will show that this conceptual model is capable of producing a simple expression of the peak position that is in agreement with essential features of experimental observations.

If the lifetime is short compared to the radiation/implantation time, then a steady state distributions of point defects and intermediate species will obtain. An immediately apparent result is that the peak-to-valley distance is invariant with radiation time, but the peak height (or valley depth) increases linearly with radiation time. In the discretized conceptual model, this simply means that C_2 as well as C_4 scale with time. Put in another way, the C_2 to C_4 evolution cycle is repeated for each increment of irradiation/implantation time on the order of the lifetime of the mobile intermediate, and the results accumulate with time. The peak-to-valley distance, L_{vp} , is then approximately given by the diffusion length of the intermediate, given by $(D_{AX}\tau_{AX})^{1/2}$. Composite results from different sources and gathered by Loualiche et al. [315] appear to support this conclusion. Loualiche et al. reported that L_{vp} follows an Arrhenius behavior, as shown in Fig. 34. While this would seem reasonable, their interpretation of this phenomenon as an indication of the migration energy of the silicon vacancy or self-interstitial is debatable. The peak-to-valley distance L_{vp} is a complicated function of the distribution function of Frenkel pair generation, the diffusivities of the vacancy, the self-interstitial and defect-impurity complexes, the rate constants of vacancy-interstitial annihilation-generation and defect-impurity complex reaction dissociation. But let us for now continue pursuing the above simplified discretized model a little further in order to gain an insight into the most important parameters affecting the quantity L_{vp} . Let us assume that the Frenkel pair generation profile gives rise to steady state profiles

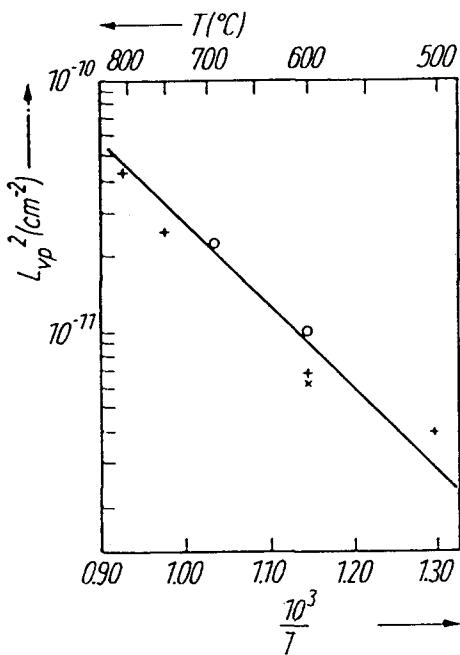


Fig. 34. The valley-to-peak distance of redistributed impurity profiles in proton-bombarded silicon as a function of temperature (taken from Loualiche et al. [315]).

of vacancies and self-interstitials via coupled reaction-diffusion processes. The steady state profiles for the vacancies and for the self-interstitial will not in general be equal, because of the difference in their diffusivities. Now assume that impurity reacts predominantly with defect X. The steady state concentration of defect X is, for simplicity, assumed to be gaussian and characterized by a spreading parameter L_0 . L_0 is characteristically similar to a diffusion length; but it should be noted that L_0 is not given by $2(D_X\tau)^{1/2}$ because it depends, perhaps more strongly, on the spread of generation profile of the Frenkel pairs. Through the law of mass action, a scaled profile of AX defect-impurity complexes is produced, with the same spreading parameter L_0 . This corresponds to curve 2 of Fig. 33, and may be written as

$$C_{AX}(x, 0) = \frac{Q}{\pi^{1/2}L_0} \exp\left[-\frac{(x-x_0)^2}{L_0^2}\right] \quad (142)$$

where x_0 is the valley position, assumed to coincide with the peak position of the Frenkel pair generation profile, and Q is the total quantity of complexes formed per cm^2 . Q is cumulative over the total radiation dose. In their lifetime, the complexes move a distance of $2(D_{AX}\tau_{AX})^{1/2}$. Thus, at the end of their lifetime, the profile of the complexes is characterized by a spreading parameter of $(L_0^2 + 4D_{AX}\tau_{AX})^{1/2}$:

$$C_{AX}(x, \tau_{AX}) = \frac{Q}{[\pi(L_0^2 + 4D_{AX}\tau_{AX})]^{1/2}} \exp\left[-\frac{(x-x_0)^2}{L_0^2 + 4D_{AX}\tau_{AX}}\right] \quad (143)$$

which corresponds to curve 4 of Fig. 33. After the dust has settled, the final impurity concentration profile is given by

$$C_A = C_A(0) - C_{AX}(x, 0) + C_{AX}(x, \tau_{AX}) \quad (144)$$

The position of the peak is then found from

$$\frac{\partial C_A}{\partial x} = 0 \quad (145)$$

from which we readily find the peak-to-valley distance to be

$$\begin{aligned} L_{vp} &= (x - x_0)_{\partial C_A / \partial x = 0} \\ &= L_0 \left[\frac{3}{2} \left(1 + \frac{L_0^2}{4D_{AX}\tau_{AX}} \right) \ln \left(1 + \frac{4D_{AX}\tau_{AX}}{L_0^2} \right) \right]^{1/2} \end{aligned} \quad (146)$$

Thus, according to this simplified model, the defect diffusion length (or the complex diffusion length) alone does not determine the peak-to-valley position; the generation profile of Frenkel pairs, characterized by the parameter L_0 , plays a role that is at least as important. As given by Eq. (146), L_{vp} is independent of irradiation time (t) and dose (αQ), in agreement with experimental observations.

11.3. Rigorous analysis of radiation-induced redistribution

The vacancy/interstitialcy wind effect can be analyzed rigorously. We would start with the same coupled diffusion reaction equations as before in our discussion of diffusion through intermediate complexes (Section 10.2). But we need to add appropriate terms for bulk sources to Eqs. (111) and (112):

$$\frac{\partial C_V}{\partial t} = \frac{\partial}{\partial x} \left(D_V \frac{\partial C_V}{\partial x} \right) + g_V(x, t) - k_{VI}(C_V C_I - C_V^* C_I^*) + k_{AV}^r C_{AV} - k_{AV}^f C_A C_V \quad (147)$$

$$\frac{\partial C_I}{\partial t} = \frac{\partial}{\partial x} \left(D_I \frac{\partial C_I}{\partial x} \right) + g_I(x, t) - k_{VI}(C_V C_I - C_V^* C_I^*) + k_{AI}^r C_{AI} - k_{AI}^f C_A C_I \quad (148)$$

$g_V(x, t)$ and $g_I(x, t)$ are the generation rates of vacancies and interstitials, respectively. Strictly speaking, they are not equal because, in addition to atomic displacement, there will be atomic replacement of lattice atoms by implant atoms. It is very difficult, if not impossible to measure the Frenkel pair generation profile experimentally. This is because any damage profile measured is the compound result of Frenkel pair generation and vacancy/interstitial reaction-diffusion. Examples of some such measurements are given by Chu et al. [355], Crowder and Title [356] and Wondrak et al. [357]. Probably, as of now, g_V and g_I may best be calculated theoretically, as described in many papers [358–367]. The reaction terms included in the above equations are by no means complete. For example, one can include the indirect recombination between the vacancy and the self-interstitial via an impurity atom, for example,



Impurities besides A, such as oxygen, which is present in concentrations as high as 10^{10} atoms cm^{-3} , may also participate as a recombination intermediary. We expect the coupled equations to be solved numerically, something which is beyond the scope of this article.

Considerable simplification of the coupled defect-impurity reaction-diffusion equations is possible. Let us limit ourselves to the time span of irradiation or implantation, during which almost all of the impurity redistribution due to the defect-wind effect takes place. We assume that the concentration of the defect-impurity pairs (or any other intermediate species) is substantially smaller than the concentration of point defects, so that a steady state can be quickly approached in a constant rate of Frenkel pair generation. We further assume, which is easily justified, that the vacancies and the self-interstitials, being generated at a high rate, are much farther from equilibrium than are the defect-impurity complex. The last assumption allows us to drop the defect-impurity reaction terms (forward-reverse cancellation) from the coupled equations. This also makes better the assumption of steady state defect distributions. We now need only to solve the following two coupled nonlinear ordinary differential equations:

$$D_V \frac{\partial^2 C_V}{\partial x^2} + g_V(x) - k_{VI}(C_V C_I - C_V^* C_I^*) = 0 \quad (151)$$

$$D_I \frac{\partial^2 C_I}{\partial x^2} + g_V(x) - k_{VI}(C_V C_I - C_V^* C_I^*) = 0 \quad (152)$$

Once $C_V(x)$ and $C_I(x)$ have been obtained, one can easily obtain a solution to the impurity diffusion by use of an effective diffusivity that is position dependent (see Eq. (139)).

12. Conclusion

Point defects play an important role in atomic transport process and the formation of extended defects. When we need only dealing with point defects at thermal equilibrium, many phenomena such as diffusion can be understood easily and quite fully. Analytical models for simulation and prediction are mature based on equilibrium point defects. When nonequilibrium point defects are present, the situation becomes extremely complicated. This is because of the large variety of sources of nonequilibrium point defects. In this report, we have discussed most of the important ones, ones that are likely to be encountered in silicon device processing. Many of these sources are still not quite thoroughly understood; but in general they are understood well enough to allow us to analyze them quantitatively. We have shown that the approach of formulating diffusion and annealing problem in terms of fully coupled reaction-diffusion equations is the only rigorous approach. In some special situations, simplified quasi-steady state approximation may serve the purpose. But in the emerging technologies of low-temperature and rapid thermal annealing, phenomena associated with nonequilibrium point defects become very complicated and profound. We have shown our best recent understanding of these phenomena. But this is far from being a mature field.

References

- [1] S.M. Hu, in D. Shaw (ed.), *Atomic Diffusion in Semiconductors*, Plenum, London, 1973, p. 217.
- [2] S.M. Hu, *J. Appl. Phys.*, 45 (1974) 1567.
- [3] P.M. Fahey, P.B. Griffin and J.D. Plummer, *Rev. Mod. Phys.*, 61 (1989) 289.
- [4] W. Schottky, *Z. Phys. Chem. B*, 29 (1935) 335.
- [5] R.L. Longini and R.F. Greene, *Phys. Rev.*, 102 (1956) 992.

- [6] W. Shockley and J.T. Last, *Phys. Rev.*, **107** (1957) 392.
- [7] W. Shockley and J.L. Moll, *Phys. Rev.*, **119** (1960) 1480.
- [8] S.M. Hu, *Phys. Rev.*, **180** (1969) 773.
- [9] S.M. Hu, *J. Appl. Phys.*, **40** (1969) 4413.
- [10] F. Seitz, *Acta Cryst.*, **3** (1950) 355.
- [11] A. Seeger and K.P. Chik, *Phys. Stat. Solidi*, **29** (1968) 455.
- [12] G.N. Will, *Solid State Electron.*, **12** (1969) 133.
- [13] K.E. Bean and P.S. Gleim, *Proc. IEEE*, **57** (1969) 1469.
- [14] T.C. Chan and C.C. Mai, *Proc. IEEE*, **58** (1970) 588.
- [15] R.A. Kovalev, V.B. Bernikov, Yu. I. Pashintsev and V.A. Marasanov, *Sov. Phys. Solid State*, **11** (1970) 1571.
- [16] M. Okamura, *Jpn. J. Appl. Phys.*, **9** (1970) 849.
- [17] L.E. Katz, in *Silicon Device Processing*, Natl. Bur. Standards Spec. Publ. 337, Washington DC, 1970, p. 192.
- [18] W.G. Allen and K.W. Anand, *Solid State Electron.*, **14** (1971) 397.
- [19] H. Higuchi, M. Maki and Y. Takano, Electrochem. Soc. Meeting, Washington, D.C., May 9–13 1971, Extended Abstract No. 78.
- [20] D.J.D. Thomas, *Phys. Stat. Solidi*, **3** (1963) 2261.
- [21] H.J. Queisser and P.G.G. van Loon, *J. Appl. Phys.*, **35** (1964) 3066.
- [22] R.J. Jaccodine and C.M. Drum, *Appl. Phys. Lett.*, **8** (1966) 29.
- [23] G.R. Booker and W.J. Tunstall, *Phil. Mag.*, **13** (1966) 71.
- [24] W.A. Fisher and J.A. Amick, *J. Electrochem. Soc.*, **113** (1966) 1054.
- [25] P.B. Hirsch, A. Howie, R.B. Nicholson and D.W. Pashley, *Electron Microscopy of Thin Crystals*, Butterworths, London, 1965, p. 234ff.
- [26] H. Hashimoto, A. Howie and M.J. Whelan, *Proc. Roy. Soc. A*, **269** (1962) 80.
- [27] S.M. Hu, *J. Appl. Phys.*, **51** (1980) 3666.
- [28] W.B. Rogers, H.Z. Massoud, R.B. Fair, U.M. Gösele, T.Y. Tan and G.A. Rozgonyi, *J. Appl. Phys.*, **65** (1989) 4215.
- [29] W.B. Rogers and H.Z. Massoud, *J. Electrochem. Soc.*, **138** (1991) 3492.
- [30] S. Mizuo and H. Higuchi, *Jpn. J. Appl. Phys.*, **20** (1981) 739.
- [31] D.A. Antoniadis and I. Moskowitz, *J. Appl. Phys.*, **53** (1982) 6788.
- [32] S. Mizuo and H. Higuchi, *Jpn. J. Appl. Phys.*, **21** (1982) 281.
- [33] Y. Hayafuji, K. Kajiwara and S. Usui, *J. Appl. Phys.*, **53** (1982) 8639.
- [34] P. Fahey, R.W. Dutton and M. Moslehi, *Appl. Phys. Lett.*, **43** (1983) 683.
- [35] S. Mizuo, T. Kusaka, S. Shintani, M. Nanba and H. Higuchi, *J. Appl. Phys.*, **54** (1983) 3860.
- [36] P. Fahey, G. Barbuscia, M. Moslehi and R.W. Dutton, *Appl. Phys. Lett.*, **46** (1985) 784.
- [37] S.M. Hu, *Phys. Stat. Solidi B*, **60** (1973) 595.
- [38] K.C. Pandey, *Phys. Rev. Lett.*, **57** (1986) 2287.
- [39] R.F. Peart, *Phys. Stat. Solidi*, **15** (1966) K119.
- [40] B.J. Masters and J.M. Fairfield, *Appl. Phys. Lett.*, **8** (1966) 280.
- [41] H.J. Mayer, H. Mehrer and K. Maier, in N.B. Urli and J.W. Corbett (eds.), *Radiation Effects in Semiconductors 1976*, Institute of Physics, London, 1977, p. 186.
- [42] L. Kalinowski and R. Seguin, *Appl. Phys. Lett.*, **35** (1979) 211.
- [43] J. Hirvonen and A. Antilla, *Appl. Phys. Lett.*, **35** (1979) 703.
- [44] F.J. Demond, S. Kalbitzer, H. Mannsperger and H. Damjantschitsch, *Phys. Lett.*, **93A** (1983) 503.
- [45] G. Hettich, H. Mehrer and K. Maier, in J.H. Albany (ed.), *Defects and Radiation Effects in Semiconductors 1978*, Institute of Physics, Bristol, 1979, p. 500.
- [46] W. Frank, U. Gösele, H. Mehrer and A. Seeger, in G.E. Murch and A.S. Nowick (eds.), *Diffusion in Crystalline Solids*, Academic, New York, 1984, p. 64.
- [47] P. Fahey, S.S. Iyer and G.J. Scilla, *Appl. Phys. Lett.*, **54** (1989) 843.
- [48] G.D. Watkins, in *Radiation Damage in Semiconductors*, Dunod, Paris, 1965, p. 97.
- [49] S.T. Sah, L. Forbes, L.L. Rosier and A.F. Tasch, *Solid-State Electron.*, **13** (1970) 759.
- [50] D.V. Lang, *J. Appl. Phys.*, **45** (1974) 3014.
- [51] J.C. Brabant, M. Pugnet, J. Barbolla and M. Brousseau, *J. Appl. Phys.*, **47** (1976) 4809.
- [52] L.C. Kimerling, in N.B. Urli and J.W. Corbett (eds.), *Radiation Effects in Semiconductors 1976*, Institute of Physics, London, 1977, p. 221.
- [53] G.D. Watkins, in F.A. Huntly (ed.), *Lattice Defects in Semiconductors 1974*, Institute of Physics, London, 1975, p. 23.
- [54] K. Weiser, *Phys. Rev.*, **126** (1962) 1427.
- [55] P.W. Anderson, *Phys. Rev. Lett.*, **34** (1975) 953.
- [56] M.J. Makin, *Phil. Mag.*, **18** (1968) 637.
- [57] M.J. Makin, *Phil. Mag.*, **20** (1969) 1133.
- [58] G.D. Watkins, in F.L. Vook (ed.), *Radiation Damage in Semiconductors*, Plenum, New York, 1968, p. 67.
- [59] G.D. Watkins, J.R. Troxell and A.P. Chatterjee, in J.H. Albany (ed.), *Defects and Radiation Effects in Semiconductors 1978*, Institute of Physics, Bristol, 1979, p. 16.

- [60] G.D. Watkins and J.R. Troxell, *Phys. Rev. Lett.*, **44** (1980) 593.
- [61] J.L. Newton, A.P. Chatterjee, R.D. Harris and G.D. Watkins, *Physica*, **116B** (1983) 219.
- [62] T.L. Chiu and N.H. Ghosh, *IBM J. Res. Dev.*, **15** (1971) 472.
- [63] R.B. Fair, in F.F.Y. Wang (ed.), *Impurity Doping Processes in Silicon*, North-Holland, New York, 1981, p. 315.
- [64] J.P. John and M.E. Law, *Appl. Phys. Lett.*, **62** (1993) 1388.
- [65] S. Dannefaer, P. Mascher and D. Kerr, *Phys. Rev. Lett.*, **56** (1986) 2195.
- [66] M. Okaji, *Int. J. Thermophys.*, **9** (1988) 1101.
- [67] Y. Okada and Y. Tokumaru, *J. Appl. Phys.*, **56** (1984) 314.
- [68] Y. Okada, *Phys. Rev. B*, **41** (1989) 10741.
- [69] K. Taniguchi, D.A. Antoniadis and Y. Matsushita, *Appl. Phys. Lett.*, **42** (1983) 961.
- [70] K. Taniguchi and D.A. Antoniadis, *Appl. Phys. Lett.*, **46** (1985) 944.
- [71] P.B. Griffin, P.M. Fahey, J.D. Plummer and R.W. Dutton, *Appl. Phys. Lett.*, **47** (1985) 319.
- [72] E. Scheid and P. Chenevier, *Phys. Stat. Solidi A*, **93** (1986) 523.
- [73] S.T. Ahn, P.B. Griffin, J.D. Shott, J.D. Plummer and W.A. Tiller, *J. Appl. Phys.*, **62** (1987) 4745.
- [74] P.B. Griffin, S.T. Ahn, W.A. Tiller and J.D. Plummer, *Appl. Phys. Lett.*, **51** (1987) 115.
- [75] W.B. Rogers and H.Z. Massoud, *J. Electrochem. Soc.*, **138** (1991) 3483.
- [76] F.C. Frank and D. Turnbull, *Phys. Rev.*, **104** (1956) 617.
- [77] U. Gösele, W. Frank, and A. Seeger, *Appl. Phys.*, **23** (1980) 361.
- [78] F. Morehead, N.A. Stolwijk, W. Meyberg and U. Gösele, *Appl. Phys. Lett.*, **42** (1983) 690.
- [79] F. Morehead, *MRS Symp. Proc.*, **104** (1988) 99.
- [80] S. Coffa, L. Calcagno, S.U. Campisano, G. Calleri and G. Ferla, *J. Appl. Phys.*, **64** (1988) 6291.
- [81] C. Boit, F. Lau and R. Sittig, *Appl. Phys. A*, **50** (1990) 197.
- [82] H. Zimmermann and H. Ryssel, *Appl. Phys. A*, **55** (1992) 121.
- [83] H. Zimmermann and H. Ryssel, *J. Electrochem. Soc.*, **139** (1992) 256.
- [84] D. Mathiot, *Phys. Rev. B*, **45** (1992) 13345.
- [85] G.B. Bronner and J.D. Plummer, *J. Appl. Phys.*, **61** (1987) 5286.
- [86] M.E. Law, *IEEE Trans. Computer-Aided Des.*, **10** (1991) 1125.
- [87] B. Baccus, T. Wada, Shigyo, M. Norishima, H. Nakajima, K. Inou, T. Iinuma and H. Iwai, *IEEE Trans. Electron Dev.*, **39** (1992) 648.
- [88] M. Hane and H. Matsumoto, *IEEE Trans. Electron Dev.*, **40** (1993) 1215.
- [89] R.A. Swalin, *J. Phys. Chem. Solids*, **18** (1961) 290.
- [90] A. Scholz and A. Seeger, *Phys. Stat. Solidi*, **3** (1963) 1480.
- [91] R.R. Hasiguti, *J. Phys. Soc. Jpn.*, **21** (1966) 1927.
- [92] I.P. Batra, F.F. Abraham and S. Ciraci, *Phys. Rev. B*, **35** (1987) 9552.
- [93] M.I. Baskes, J.S. Nelson and A.F. Wright, *Phys. Rev. B*, **40** (1989) 6085.
- [94] P.J. Ungar, T. Takai, T. Halicioglu and W.A. Tiller, *J. Vac. Sci. Technol. A*, **11** (1993) 224.
- [95] D. Maroudas and R.A. Brown, *Appl. Phys. Lett.*, **62** (1993) 172.
- [96] S.T. Pantelides, I. Ivanov, M. Scheffler and J.P. Vigneron, *Physica*, **116B** (1983) 18.
- [97] W. Kohn and L.J. Sham, *Phys. Rev.*, **140** (1965) A1138.
- [98] G.A. Baraff and M. Schlüter, *Phys. Rev. Lett.*, **41** (1978) 892.
- [99] J. Bernholc, N.O. Lipari and S.T. Pantelides, *Phys. Rev. Lett.*, **41** (1978) 895.
- [100] N.O. Lipari, J. Bernholc and S.T. Pantelides, *Phys. Rev. Lett.*, **43** (1979) 1354.
- [101] J. Bernholc, N.O. Lipari and S.T. Pantelides, *Phys. Rev. B*, **21** (1980) 3545.
- [102] G.A. Baraff, E.O. Kane and M. Schlüter, *Phys. Rev. B*, **21** (1980) 5662.
- [103] R. Car, P.J. Kelly, A. Oshiyama and S.T. Pantelides, *Phys. Rev. Lett.*, **52** (1984) 1814.
- [104] G.A. Baraff and M. Schlüter, *Phys. Rev. B*, **30** (1984) 3460.
- [105] R. Car, P.J. Kelly, A. Oshiyama and S.T. Pantelides, *Phys. Rev. Lett.*, **54** (1985) 360.
- [106] P.J. Kelly and R. Car, *Phys. Rev. B*, **45** (1992) 6543.
- [107] Y. Bar-Yam and J.D. Joannopoulos, *Phys. Rev. Lett.*, **52** (1984) 1129.
- [108] Y. Bar-Yam and J.D. Joannopoulos, *Phys. Rev. B*, **30** (1984) 2216.
- [109] Y. Bar-Yam and J.D. Joannopoulos, *Phys. Rev. B*, **30** (1984) 1844.
- [110] A. Antonelli and J. Bernholc, *Phys. Rev. B*, **40** (1989) 10643.
- [111] C.S. Nichols, C.G. Van de Walle and S.T. Pantelides, *Phys. Rev. B*, **40** (1989) 5484.
- [112] O. Sugino and A. Ashiyama, *Phys. Rev. B*, **46** (1992) 12335.
- [113] D. Mathiot and J.C. Pfister, *J. Phys. Lett. (Paris)*, **43** (1982) 453.
- [114] G.D. Watkins, R.P. Messmer, C. Weigel, D. Peak and J.W. Corbett, *Phys. Rev. Lett.*, **27** (1971) 1573.
- [115] H. Strunk, U. Gösele and B.O. Kolbesen, *Appl. Phys. Lett.*, **34** (1979) 530.
- [116] P. Fahey, R.W. Dutton, and S.M. Hu, *Appl. Phys. Lett.*, **44** (1984) 777.
- [117] K. Nishi and D.A. Antoniadis, *J. Appl. Phys.*, **56** (1984) 3428.
- [118] E. Nygren, M. Aziz, D. Turnbull, J.M. Poate, D.C. Jacobson and R. Hull, *Appl. Phys. Lett.*, **47** (1985) 105.
- [119] J. Furthmüller and M. Fähnle, *Phys. Rev. B*, **46** (1992) 3839.

- [120] T.S. Plaskett, *Trans. AIME*, 233 (1965) 809.
- [121] T. Abe, T. Samizo and S. Maruyama, *Jpn. J. Appl. Phys.*, 5 (1966) 458.
- [122] A.J.R. de Kock, *Philips Res. Rep. Suppl.* 1 (1973).
- [123] H. Föll and B.O. Kolbesen, *Appl. Phys.*, 8 (1975) 319.
- [124] S.M. Hu, *J. Vac. Sci. Technol.*, 14 (1977) 17.
- [125] F.R.N. Nabarro, *Proc. R. Soc. (London)*, A175 (1940) 519.
- [126] F.R.N. Nabarro, *Proc. Phys. Soc. (London)*, 52 (1940) 90.
- [127] U. Gösele, W. Frank and A. Seeger, *Solid State Commun.*, 45 (1982) 45.
- [128] A.J.R. de Kock and W.M. van de Wijgert, *J. Cryst. Growth*, 49 (1980) 718.
- [129] P.M. Petroff and A.J.R. de Kock, *J. Cryst. Growth*, 36 (1976) 4.
- [130] H. Föll, U. Gösele and B.O. Kolbesen, *J. Cryst. Growth*, 40 (1977) 90.
- [131] P.J. Roksnoer, *J. Cryst. Growth*, 68 (1984) 596.
- [132] A.J.R. de Kock, W.T. Stacy and W.M. van de Wijgert, *Appl. Phys. Lett.*, 34 (1979) 611.
- [133] P.J. Roksnoer and M.M.B. van den Boom, *J. Cryst. Growth*, 53 (1981) 563.
- [134] L.M. Sorokin, A.A. Sitnikova, I.F. Chervonyi, and É.S. Fal'kevich, *Sov. Phys. Solid State*, 33 (1991) 1824.
- [135] H. Yamagishi, I. Fusegawa, N. Fujimaki and M. Katayama, *Semicond. Sci. Technol.*, 7 (1992) 135.
- [136] V.V. Voronkov, *J. Cryst. Growth*, 59 (1982) 625.
- [137] C.A. Pimentel and B.C. Brito-Filho, *J. Cryst. Growth*, 62 (1983) 129.
- [138] A. Harada, T. Abe and J. Chikawa, in H.R. Huff, T. Abe and B. Kolbesen (eds.), *Semiconductor Silicon 1986*, Electrochemical Society, Pennington, 1986, p. 76.
- [139] C.J. Gallagher, *Phys. Rev.*, 100 (1955) 1259.
- [140] G. Bemski and C.A. Dias, *J. Appl. Phys.*, 35 (1964) 2893.
- [141] L. Elstner and W. Kamprath, *Phys. Status Solidi*, 22 (1967) 541.
- [142] B.I. Boltaks and S.I. Budarina, *Sov. Phys. Solid State*, 11 (1969) 330.
- [143] J.D. Gerson, L.J. Cheng and J.W. Corbett, *J. Appl. Phys.*, 48 (1977) 4821.
- [144] Y.H. Lee, R.L. Kleinhenz and J.W. Corbett, *Appl. Phys. Lett.*, 31 (1977) 142.
- [145] E. Weber and H.G. Riotti, *Appl. Phys. Lett.*, 33 (1978) 433.
- [146] H. Feichtinger, J. Waltl and A. Gschwandtner, *Solid-State Commun.*, 27 (1978) 867.
- [147] H.J. Rijks, J. Bloem and L.J. Giling, *J. Appl. Phys.*, 50 (1979) 1370.
- [148] E. Weber and H.G. Riotti, *J. Appl. Phys.*, 51 (1980) 1484.
- [149] K. Graff and H. Pieper, *J. Electrochem. Soc.*, 128 (1981) 669.
- [150] K.D. Glinchuk and N.M. Litovichenko, *Phys. Status Solidi (a)*, 66 (1981) K75.
- [151] C.B. Collins and R.O. Carlson, *Phys. Rev.* (1957) 1409.
- [152] N.F. Mott, *Proc. R. Soc. (London) A*, 376 (1981) 207.
- [153] N.F. Mott, *S. Afric. J. Phys.*, 9 (1986) 40.
- [154] N.F. Mott, *Trans. Faraday Soc.*, 43 (1947) 429.
- [155] N. Cabrera and N.F. Mott, *Rep. Progr. Phys.*, 12 (1948/49) 163.
- [156] D.A. Vermilyea, *Acta Metall.*, 5 (1957) 492.
- [157] W.A. Tiller, *J. Electrochem. Soc.*, 130 (1983) 501.
- [158] J.P. Hirth and A.A. Tiller, *J. Appl. Phys.*, 56 (1984) 947.
- [159] S.T. Dunham and S.D. Plummer, *J. Appl. Phys.*, 59 (1986) 2541.
- [160] D.G. Cahill and Ph. Avouris, *Appl. Phys. Lett.*, 60 (1992) 326.
- [161] G. Hetherington, K.H. Jack and J.C. Kennedy, *Phys. Chem. Glasses*, 5 (1964) 130.
- [162] T.Y. Tan and U. Gösele, *Appl. Phys. Lett.*, 39 (1981) 86.
- [163] L.M. Landsberger and W.A. Tiller, *Appl. Phys. Lett.*, 51 (1987) 1416.
- [164] S.M. Hu, *J. Appl. Phys.*, 64 (1988) 323.
- [165] E.P. EerNisse, *Appl. Phys. Lett.*, 35 (1979) 8.
- [166] S.M. Hu, *J. Appl. Phys.*, 70 (1991) R53.
- [167] A. Mayer, *RCA Rev.*, 31 (1970) 414.
- [168] S.M. Hu, *Appl. Phys. Lett.*, 27 (1975) 165.
- [169] C.L. Claeys, E.L. Laes, G.J. Declerck and R.J. van Overstraeten, in H.R. Huff and E. Sirtl (eds.), *Semiconductor Silicon 1977*, Electrochemical Society, Pennington NJ, 1977, p. 773.
- [170] S.P. Murarka and G. Quintana, *J. Appl. Phys.*, 48 (1977) 46.
- [171] S.P. Murarka, *J. Appl. Phys.*, 49 (1978) 2513.
- [172] H. Hashimoto, H. Shibayama, H. Masaki and H. Ishikawa, *J. Electrochem. Soc.*, 123 (1976) 1899.
- [173] Y. Sugita, H. Shimizu, A. Yoshinaka and T. Aoshima, *J. Vac. Sci. Technol.*, 14 (1977) 44.
- [174] H. Shimizu, A. Yoshinaka and Y. Sugita, *Jpn. J. Appl. Phys.*, 17 (1978) 767.
- [175] C.L. Claeys, G.J. Declerck and R.J. van Overstraeten, *Appl. Phys. Lett.*, 35 (1979) 797.
- [176] A.M. Lin, R.W. Dutton, D.A. Antoniadis and W.A. Tiller, *J. Electrochem. Soc.*, 128 (1981) 1121.
- [177] S.M. Hu, in J. Narayan and T.Y. Tan (eds.) *Defects in Semiconductors*, North-Holland, New York, 1981, p. 333.
- [178] S.M. Hu, *Appl. Phys. Lett.*, 43 (1983) 449.

- [179] S.M. Hu, *J. Appl. Phys.*, **57** (1985) 4527.
- [180] M. Hamasaki, *Solid State Electron.*, **25** (1982) 1.
- [181] P. Fahey and P. Griffin, in H.R. Huff, K.G. Barraclough and J. Chikawa (eds.), *Semiconductor Silicon 1990*, Electrochemical Society, Pennington, 1990, p. 486.
- [182] Y. Shibata, K. Taniguchi and C. Hamaguchi, *J. Appl. Phys.*, **70** (1991) 4846.
- [183] A.M. Lin, D.A. Antoniadis and R.W. Dutton, *J. Electrochem. Soc.*, **128** (1981) 1131.
- [184] N. Jeng and S.T. Dunham, *J. Appl. Phys.*, **72** (1992) 2049.
- [185] Y. Shibata, S. Hashimoto, K. Taniguchi and C. Hamaguchi, *J. Electrochem. Soc.*, **139** (1992) 231.
- [186] S.T. Dunham, *J. Electrochem. Soc.*, **136** (1989) 250.
- [187] S.T. Dunham, *J. Appl. Phys.*, **71** (1992) 685.
- [188] K. Taniguchi, Y. Shibata and C. Hamaguchi, *J. Appl. Phys.*, **65** (1989) 2723.
- [189] K. Taniguchi, Y. Shibata and C. Hamaguchi, *IEEE Trans. Computer-Aided Des.*, **9** (1990) 1177.
- [190] B.E. Deal and A.S. Grove, *J. Appl. Phys.*, **36** (1965) 3770.
- [191] M.P. Murrell, C.J. Sofield and S. Sugden, *Phil. Mag. B*, **63** (1991) 1277.
- [192] H. Shiraki, *Jpn. J. Appl. Phys.*, **15** (1) (1976) 83.
- [193] T. Hattori, *J. Electrochem. Soc.*, **123** (1976) 945.
- [194] Y. Nabeta, T. Uno, S. Kubo and H. Tsukamoto, *J. Electrochem. Soc.*, **123** (1976) 1714.
- [195] C.L. Claeys, G.J. Declerck and R.J. van Overstraeten, in P.A. Barnes and G.A. Rozgonyi (eds.), *Semiconductor Characterization Techniques*, Electrochem. Soc., Pennington, 1978, p. 366.
- [196] M. Wittmer and K.N. Tu, *Phys. Rev. B*, **29** (1984) 2010.
- [197] I. Ohdomari, M. Akiyama, T. Maeda, M. Hori, C. Takebayashi, A. Ogura, T. Chikyo, I. Kimura, K. Yoneda and K.N. Tu, *J. Appl. Phys.*, **56** (1984) 2723.
- [198] M. Wittmer, P.M. Fahey, G.J. Scilla, S.S. Iyer and M. Tejwani, *Phys. Rev. Lett.*, **66** (1991) 632.
- [199] M. Wittmer, P.M. Fahey, J. Cotte, S.S. Iyer and G.J. Scilla, *Phys. Rev. Lett.*, **66** (1991) 632.
- [200] P. Pichler and M. Orlowski, *Appl. Phys. Lett.*, **60** (1992) 1205.
- [201] P.A. Ronsheim and M. Tejwani, *Phys. Rev. Lett.*, **71** (1993) 947.
- [202] A. Ishitani, A. Karen, Y. Nakagawa, M. Uchida, M. Hatada, K. Okuno and F. Soeda, *Proc. 8th Int. Conf. SIMS* (1992) 315.
- [203] A. Karen, Y. Nakagawa, M. Uchida, M. Hatada, K. Okuno, F. Soeda and A. Ishitani, *Proc. 8th Int. Conf. SIMS* (1992) 339.
- [204] S.M. Hu, *Appl. Phys. Lett.*, **51** (1987) 308.
- [205] D.S. Wen, P.L. Smith, C.M. Osburn and G.A. Rozgonyi, *Appl. Phys. Lett.*, **51** (1987) 1182.
- [206] A. Ajmera and G.A. Rozgonyi, *Appl. Phys. Lett.*, **49** (1986) 1260.
- [207] K. Maeck, R. de Keersmaecker, C. Claeys, J. Vanhellemont and P.F.A. Alkemade, in H.R. Huff, T. Abe and B. Kolbesen (eds.), *Semiconductor Silicon 1986*, Electrochemical Society, Pennington NJ, 1986, p. 346.
- [208] I. Ohdomari, K. Konuma, K. Takano, T. Chikyow, H. Kawarada, J. Nakanishi and T. Ueno, *Mater. Res. Soc. Symp. Proc.*, **54** (1986) 63.
- [209] J.W. Honeycutt and G.A. Rozgonyi, *Appl. Phys. Lett.*, **58** (1991) 1302.
- [210] S.M. Hu and W.J. Patrick, *J. Appl. Phys.*, **46** (1975) 1869.
- [211] S.M. Hu, *Appl. Phys. Lett.*, **31** (1977) 53.
- [212] J.T. Yue and H.J. Ruiz, in H.R. Huff and E. Sirtl (eds.), *Semiconductor Silicon 1977*, Electrochemical Society, Pennington NJ, 1977, p. 596.
- [213] K. Sumino, H. Harada and I. Yonegawa, *Jpn. J. Appl. Phys.*, **19** (1980) L49.
- [214] H.-D. Chiou, J. Moody, R. Sandfort and F. Shimura, in K.E. Bean and G.A. Rozgonyi (eds.), *VLSI Science and Technology/1984*, Electrochemical Society, Pennington NJ, 1984, p. 59.
- [215] G.A. Rozgonyi, R.P. Deysher and C.W. Pearce, *J. Electrochem. Soc.*, **123** (1976) 1910.
- [216] T.Y. Tan, E.E. Gardner and W.K. Tice, *Appl. Phys. Lett.*, **30** (1977) 175.
- [217] G.A. Rozgonyi and C.W. Pearce, *Appl. Phys. Lett.*, **32** (1978) 747.
- [218] W.J. Patrick, S.M. Hu and W.A. Westdorp, *J. Appl. Phys.*, **50** (1979) 1399.
- [219] G.H. Schwuttke, *J. Electrochem. Soc.*, **108** (1961) 163.
- [220] L. Fiersmans and J. Vennik, *Phys. Stat. Solidi*, **12** (1965) 277.
- [221] S.M. Hu, *J. Appl. Phys.*, **51** (1980) 5945.
- [222] J.E. Olson and F. Shimura, *J. Appl. Phys.*, **66** (1989) 1353.
- [223] A. Borghesi, B. Pivac and A. Sassella, *Appl. Phys. Lett.*, **60** (1992) 871.
- [224] S.M. Hu, *Appl. Phys. Lett.*, **36** (1980) 561.
- [225] H.F. Schaake, S.C. Barber and R.F. Pinizzotto, in H.R. Huff, R.J. Kriegler and Y. Takeishi (eds.), *Semiconductor Silicon 1981*, Electrochemical Society, Pennington, 1981, p. 344.
- [226] R.A. Craven, in H.R. Huff, R.J. Kriegler and Y. Takeishi (eds.), *Semiconductor Silicon 1981*, Electrochemical Society, Pennington, 1981, p. 254.
- [227] G.S. Oehrlein, J.L. Lindstroem and J.W. Corbett, *Appl. Phys. Lett.*, **40** (1982) 241.

- [228] J.O. Borland, in W.M. Bullis and L.C. Kimerling (eds.), *Defects in Silicon*, Electrochemical Society, Pennington NJ, 1983, p. 236.
- [229] T.Y. Tan and C.Y. Kung, in H.R. Huff, T. Abe and B. Kolbesen (eds.), *Semiconductor Silicon 1986*, Electrochemical Society, Pennington, 1986, p. 864.
- [230] J.S. Kang and D.K. Schroder, *J. Appl. Phys.*, **64** (1988) 6673.
- [231] H. Yamanaka, Y. Aoki and T. Samizo, *Jpn. J. Appl. Phys.*, **29** (1990) 2450.
- [232] M. Kohketsu and S. Isomae, *Jpn. J. Appl. Phys.*, **30** (1991) L1337.
- [233] F. Shimura, *J. Appl. Phys.*, **72** (1992) 1642.
- [234] H.W. Kennel and J.D. Plummer, in H.R. Huff, K.G. Barraclough and J. Chikawa (eds.), *Semiconductor Silicon 1990*, Electrochemical Society, Pennington, 1990, p. 496.
- [235] A. Borghesi, A. Piaggi, A. Sassella, A. Stella and B. Pivac, *Phys. Rev. B*, **46** (1992) 4123.
- [236] A. Bourret, J. Thibault-Desseaux and D.N. Seidman, *J. Appl. Phys.*, **55** (1984) 825.
- [237] H. Bender, *Phys. Status Solidi A*, **86** (1984) 245.
- [238] W. Bergholz, P. Pirouz and J.L. Hutchinson, in L.C. Kimerling and M.M. Parsey (eds.), *Proc. 13th Int. Conf. Defects in Semiconductors*, The Metallurgical Society, AIME, Warrendale PA, 1984, p. 717.
- [239] M. Olivier, D. Lafeuille, M. Dupuy, G. Rolland and B. Guinet, *Extended Abstr.*, **75-1** (1975) 343.
- [240] W.K. Tice and T.Y. Tan, *Appl. Phys. Lett.*, **28** (1976) 564.
- [241] T.Y. Tan and W.K. Tice, *Phil. Mag.*, **34** (1976) 615.
- [242] D.M. Maher, A. Staudinger and J. Patel, *J. Appl. Phys.*, **47** (1976) 3813.
- [243] W.J. Patrick, E. Hearn, W.A. Westdorp and A. Bohg, *J. Appl. Phys.*, **50** (1979) 7156.
- [244] K. Tempelhoff, F. Spiegelberg, R. Gleichmann and D. Wruck, *Phys. Status Solidi A*, **56** (1979) 213.
- [245] F. Shimura, H. Tsuya and T. Kawamura, *J. Appl. Phys.*, **51** (1980) 269.
- [246] C. Claeys, H. Bender, G. Declerck, J. Van Landuyt, R. Van Overstraeten and S. Amelincks, in E. Sirtl and J. Goorissen (eds.), *Aggregation Phenomena of Point Defects in Silicon*, Electrochemical Soc., Pennington NJ, 1983, p. 74.
- [247] K.H. Yang, H.F. Kappert and G.H. Schwuttke, *Phys. Status Solidi A*, **50** (1978) 221.
- [248] F. Shimura, *J. Cryst. Growth*, **54** (1981) 588.
- [249] F.A. Ponce, T. Yamashita and S. Hahn, *Appl. Phys. Lett.*, **43** (1983) 1051.
- [250] H.L. Tsai, *J. Appl. Phys.*, **58** (1985) 3775.
- [251] J.J. Wortman and R.A. Evans, *J. Appl. Phys.*, **36** (1965) 153.
- [252] W.A. Brantley, *J. Appl. Phys.*, **44** (1973) 534.
- [253] J.F. Nye, *Physical Properties of Crystals*, Clarendon Press, Oxford, 1957, p. 146.
- [254] R.J. Jaccodine, *J. Electrochem. Soc.*, **110** (1963) 524.
- [255] F.R. Boyd and J.L. England, *J. Geophys. Res.*, **65** (1960) 749.
- [256] S.M. Hu, *Appl. Phys. Lett.*, **48** (1986) 115.
- [257] U. Gösele and T.Y. Tan, *Appl. Phys. A*, **28** (1982) 79.
- [258] H. Alexander, H. Eppenstein, H. Gottschalk and S. Wendler, *J. Microscop.*, **118** (1980) 1.
- [259] U. Gösele and W. Frank, in J. Narayan and T.Y. Tan (eds.), *Defects in Semiconductors*, North-Holland, New York, 1981, p. 55.
- [260] T.Y. Tan and U. Gösele, *J. Appl. Phys.*, **53** (1982) 4767.
- [261] S.M. Hu, *J. Appl. Phys.*, **52** (1981) 3974.
- [262] R.C. Newman, M.J. Binns, W.P. Brown, F.M. Livingston, S. Messoloras, R.J. Stewart and J.G. Wilkes, *Physica B*, **116** (1983) 264.
- [263] F.S. Ham, *J. Phys. Chem. Solids*, **6** (1958) 335.
- [264] M.C. Arst and J.G. de Groot, *J. Electron. Mater.*, **13** (1984) 763.
- [265] H. Shirai, A. Yamaguchi, and F. Shimura, *Appl. Phys. Lett.*, **54** (1989) 1748.
- [266] S.M. Hu, U.S. Patent 4,053,335, October 11 1977.
- [267] M.C. Chen and V.J. Silvestri, *J. Electrochem. Soc.*, **129** (1982) 1294.
- [268] H. Yamanaka, H. Kaneko, Y. Haruta and Y.Y. Iwamoto, *Jpn. J. Appl. Phys.*, **32** (1993) PL738.
- [269] E. Sirtl, in H.R. Huff and E. Sirtl (eds.), *Semiconductor Silicon 1977*, Electrochemical Society, Pennington NJ, 1977, p. 4.
- [270] T.Y. Tan and U. Gösele, *Appl. Phys. Lett.*, **40** (1982) 616.
- [271] D.A. Antoniadis, *J. Electrochem. Soc.*, **129** (1982) 1093.
- [272] T.Y. Tan, U. Gösele and F.F. Morehead, *Appl. Phys. A*, **31** (1983) 97.
- [273] S.M. Hu, *J. Appl. Phys.*, **57** (1985) 1069.
- [274] H.R. Yeager and R.W. Dutton, *IEEE Trans. Electron Dev.*, **32** (1985) 1964.
- [275] T.R. Waite, *Phys. Rev.*, **107** (1957) 463.
- [276] R.M. Noyes, in G. Porter (ed.), *Progress in Reaction Kinetics*, Vol. 1, Pergamon, New York, 1961, pp. 129–160.
- [277] B.J. Masters, in J.H. Albany (ed.), *Defects and Radiation Effects in Semiconductors 1978*, Institute of Physics, Bristol, 1979, p. 545.
- [278] L.E. Miller, in H. Gato (ed.), *Properties of Elemental and Compound Semiconductors*, Interscience, New York, 1960, p. 303.

- [279] G.E. Moore, in E. Keonjian (ed.), *Microelectronics*, McGraw-Hill, New York, 1963, p. 282.
- [280] Y. Sato and H. Harata, *Jpn. J. Appl. Phys.*, **3** (1964) 511.
- [281] R. Gereth, P.G. G. Van Loon and W. Williams, *J. Electrochem. Soc.*, **112** (1965) 323.
- [282] K.H. Nicholas, *Solid State Electron.*, **9** (1966) 35.
- [283] J.E. Lawrence, *J. Appl. Phys.*, **37** (1966) 4106.
- [284] A.F.W. Willoughby, *J. Mater. Sci.*, **3** (1968) 89.
- [285] J. Makris and B.J. Masters, *J. Appl. Phys.*, **42** (1971) 3750.
- [286] M. Yoshida, E. Arai, H. Nakamura and Y. Terunuma, *J. Appl. Phys.*, **45** (1974) 1498.
- [287] S. Solmi, G. Cellotti, D. Nobili and P. Negrini, *J. Electrochem. Soc.*, **123** (1976) 654.
- [288] J. Crank, *The Mathematics of Diffusion*, Oxford University Press, London, 1956, pp. 232–233.
- [289] S.M. Hu and S. Schmidt, *J. Appl. Phys.*, **39** (1968) 4272.
- [290] S.M. Hu. *IBM Internal Memo*, 1968.
- [291] D.B. Lee and A.F.W. Willoughby, *J. Appl. Phys.*, **43** (1972) 245.
- [292] M. Yoshida, *J. Appl. Phys.*, **48** (1977) 2169.
- [293] G.D. Watkins, J.W. Corbett and R.M. Walker, *J. Appl. Phys.*, **30** (1959) 1198.
- [294] F.N. Schwettmann and D.L. Kendall, *Appl. Phys. Lett.*, **19** (1971) 218.
- [295] F.N. Schwettmann and D.L. Kendall, *Appl. Phys. Lett.*, **21** (1972) 2.
- [296] R.B. Fair and J.C.C. Tsai, *J. Electrochem. Soc.*, **124** (1977) 1107.
- [297] R.M. Harris and D.A. Antoniadis, *Appl. Phys. Lett.*, **43** (1983) 937.
- [298] S.M. Hu, P. Fahey and R.W. Dutton, *J. Appl. Phys.*, **54** (1983) 6912.
- [299] F. Seitz, *Phys. Rev.*, **74** (1948) 1513.
- [300] L.S. Darken, *Trans. AIME*, **175** (1948) 184.
- [301] J. Bardeen and C. Herring, in W. Shockley, J.H. Hollomon and R. Maurer (eds.), *Imperfections in Nearly Perfect Crystals*, Wiley, New York, 1952, p. 261; Pocono Manor Symposium, Oct. 12–14, 1950.
- [302] A.C. Smigelskas and E.O. Kirkendall, *Trans. AIME*, **147** (1942) 104.
- [303] J.C. Fisher, J.H. Hollomon, and D. Turnbull, *Trans. AIME*, **175** (1948) 202.
- [304] S.M. Hu, *J. Electrochem. Soc.*, **139** (1992) 2066.
- [305] U. Gösele and H. Strunk, *Appl. Phys.*, **20** (1979) 265.
- [306] J. Tersoff, *Phys. Rev. Lett.*, **65** (1990) 887.
- [307] S.M. Hu, *Phys. Rev. Lett.*, **63** (1989) 2492.
- [308] S.M. Hu, D.C. Ahlgren, P.A. Ronsheim and J.O. Chu, *Phys. Rev. Lett.*, **67** (1991) 1450.
- [309] S.M. Hu, *Phys. Rev. B*, **45** (1992) 4498.
- [310] G.H. Loehelt, G. Tam, J.W. Steele, L.K. Knoch, K.M. Klein, J.K. Watanabe and J.W. Christiansen, *J. Appl. Phys.*, **74** (1993) 5520.
- [311] W. Akutagawa, H.L. Dunlap, R. Hart and O.J. Marsh, *J. Appl. Phys.*, **50** (1979) 777.
- [312] Y. Morikawa, K. Yamamoto and K. Nagami, *Appl. Phys. Lett.*, **36** (1980) 997.
- [313] P. Baruch, C. Constantin, J.C. Pfister and R. Saintesprit, *Discuss. Faraday Soc.*, **31** (1961) 76.
- [314] B.J. Masters and E.F. Gorey, *J. Appl. Phys.*, **49** (1978) 2717.
- [315] S. Loualiche, C. Lucas, P. Baruch, J.P. Gailliard and J.C. Pfister, *Phys. Stat. Solidi A*, **69** (1982) 663.
- [316] V.V. Kozlovskii, V.N. Lomasov, G.M. Gur'yanov and A.P. Kovarskii, *Sov. Phys. Semicond.*, **18** (1984) 598.
- [317] R.P. Johnson, *Phys. Rev.*, **56** (1939) 814.
- [318] H.F. Schaake, *J. Appl. Phys.*, **55** (1984) 1208.
- [319] D. Mathiot and J.C. Pfister, *Appl. Phys. Lett.*, **47** (1985) 962.
- [320] F.F. Morehead and R.F. Lever, *Appl. Phys. Lett.*, **48** (1988) 151.
- [321] F.F. Morehead and R.F. Lever, *J. Appl. Phys.*, **66** (1989) 5349.
- [322] W.B. Richardson and B.J. Mulvaney, *J. Appl. Phys.*, **65** (1989) 2243.
- [323] R.B. Fair, J.J. Wortman and J. Liu, *J. Electrochem. Soc.*, **131** (1984) 2387.
- [324] S. Guimarães, R. Kröger, E. Landi, S. Solmi and E. Wieser, *Phys. Stat. Sol. (a)*, **93** (1986) 549.
- [325] M. Miyake and S. Aoyama, *J. Appl. Phys.*, **63** (1988) 1754.
- [326] K. Ghaderi, G. Hobler, M. Budil, H. Pötzl, P. Pichler, H. Ryssel, W. Hansch, I. Eisele, C. Tian and H. Stingereder, *Electrochem. Soc. Meeting*, San Francisco, May 22–27, 1994.
- [327] L. Lapidus and N.R. Amundson, *J. Phys. Chem.*, **56** (1952) 984.
- [328] J. Crank, *The Mathematics of Diffusion*, Oxford University Press, London, 1956, pp. 124–131.
- [329] F.N. Schwettmann, *Appl. Phys. Lett.*, **22** (1973) 570.
- [330] N.E.B. Cowern, K.T.F. Janssen, G.F.A. van de Walle and G.J. Gravesteijn, *Phys. Rev. Lett.*, **65** (1990) 2434.
- [331] N.E.B. Cowern, G.F.A. van de Walle, G.J. Gravesteijn and C.J. Vriezema, *Phys. Rev. Lett.*, **67** (1991) 212.
- [332] N.E.B. Cowern and G.F.A. van de Walle, *Mod. Phys. Lett. B*, **5** (1991) 1555.
- [333] N.E.B. Cowern, G.F.A. van de Walle, P.C. Zalm and D.J. Oostra, *Phys. Rev. Lett.*, **69** (1992) 116.
- [334] W.K. Hofker, H.W. Werner, D.P. Oosthoek and H.A.M. de Grefte, *Appl. Phys.*, **2** (1973) 265.
- [335] A.E. Michel, W. Rausch, P.A. Ronsheim and R.H. Kastl, *Appl. Phys. Lett.*, **50** (1989) 416.
- [336] Y. Kim, H.Z. Massoud and R.B. Fair, *J. Electrochem. Soc.*, **137** (1990) 2599.

- [337] H. Shibayama, H. Masaki and H. Hashimoto, *Appl. Phys. Lett.*, **27** (197) 230.
- [338] H. Shibayama, H. Masaki, H. Ishikawa and H. Hashimoto, *J. Electrochem. Soc.*, **123** (1976) 742.
- [339] H. Park and M.E. Law, *Appl. Phys. Lett.*, **58** (1991) 732.
- [340] W.P. Dumke, *Appl. Phys. Lett.*, **11** (1967) 314.
- [341] P.N. Butcher, W. Fawcett and N.R. Ogg, *Br. J. Appl. Phys.*, **18** (1967) 755.
- [342] R. Landauer and J.W.F. Woo, in H. Haken (ed.), *Synergetics*, Teubner, Stuttgart, 1973, pp. 97–123.
- [343] P.T. Landsberg and S.A. Hope, *Solid State Electron.*, **20** (1977) 421.
- [344] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981, p. 291ff.
- [345] J.D. Fehribach, R. Ghez and G.S. Oehrlein, *Appl. Phys. Lett.*, **46** (1985) 433.
- [346] A. Arbel and M. Natan, *J. Appl. Phys.*, **61** (1987) 1209.
- [347] J. Crank, *The Mathematics of Diffusion*, Oxford University Press, London, 1956, p. 147.
- [348] P. Baruch, J. Monnier, B. Blanchard and C. Castaing, *Appl. Phys. Lett.*, **26** (1975) 77.
- [349] K. Holldack, H. Kerkow and W. Frentrup, *Phys. Stat. Solidi A*, **94** (1986) 357.
- [350] G.A. Kachurin, I.E. Tyshchenko, E. Wieser and Ch. Weise, *Phys. Stat. Solidi A*, **109** (1988) 141.
- [351] P.G. Merli and F. Zignani, *Radiat. Eff. Lett.*, **50** (1980) 115.
- [352] G. Lulli, P.G. Merli, R. Rizzoli, M. Berti and A.V. Drigo, *J. Appl. Phys.*, **66** (1989) 2840.
- [353] D.K. Sadana, B. Acovic, B. Davari, D. Grutzmacher, H. Hanafi and F. Cardone, *Appl. Phys. Lett.*, **61** (1992) 3038.
- [354] M.D. Giles, *Appl. Phys. Lett.*, **62** (1993) 1940.
- [355] W.K. Chu, R.H. Kastl, R.F. Lever, S. Mader and B.J. Masters, *Phys. Rev. B*, **16** (1977) 3851.
- [356] B.L. Crowder and R.S. Title, *Radiat. Eff.*, **6** (1970) 63.
- [357] W. Wondrak, K. Bethge and D. Silber, *J. Appl. Phys.*, **62** (1987) 3464.
- [358] P.V. Pavlov, D.I. Tetelbaum, E.I. Zorin and V.I. Alekseer, *Sov. Phys. Solid State*, **8** (1967) 2141.
- [359] P. Sigmund, *Radiat. Eff.*, **1** (1969) 15.
- [360] P. Sigmund, *Appl. Phys. Lett.*, **14** (1969) 114.
- [361] D.K. Brice, *Appl. Phys. Lett.*, **16** (1970) 1039.
- [362] D.K. Brice, *Radiat. Eff.*, **6** (1970) 77.
- [363] K.B. Winterbon, *Radiat. Eff.*, **46** (1980) 181.
- [364] V.V. Emtsev, T.V. Mashovets and N.A. Vitovskii, *Phys. Status Solidi A*, **90** (1985) 523.
- [365] V.V. Emtsev, T.V. Mashovets and V. Mikhovich, *Sov. Phys. Semicond.*, **26** (1992) 12.
- [366] A.M. Mazzone, *Phys. Stat. Solidi A*, **95** (1986) 149.
- [367] G. Hobler and S. Selberherr, *IEEE Trans. Computer-Assisted Des.*, *CAD-7* (1988) 174.