Solid State Phenomena
ISSN: 1662-9779, Vols. 205-206, pp 157-162
doi:10.4028/www.scientific.net/SSP.205-206.157

Online: 2013-10-07

Fast and Slow Vacancies in Silicon

V. V. Voronkov [a] and R. Falster [b]

MEMC Electronic Materials, via Nazionale 59, 39012 Merano, Italy

[a]vvoronkov@memc.it, [b]rfalster@memc.it

Keywords: vacancy, self-interstitial, silicon, diffusion

Abstract. Vacancies (and probably also self-interstitials) in silicon appear to exist in several forms (atomic configurations) some of them being fast diffusers and other – slow diffusers. The data on enhanced self-diffusivity under proton irradiation, on vacancy and oxide precipitate profiles installed by Rapid Thermal Annealing, and on the self-diffusivity under equilibrium conditions suggest that there are at least two kinds of vacancy: 1) V_w - a fast-diffusing localized vacancy manifested in electron irradiated samples (Watkins vacancy), 2) V_s - a slow-diffusing extended vacancy manifested under hot proton irradiation. In RTA experiments, these two species behave as one equilibrated subsystem of a moderate effective diffusivity intermediate between those of V_w and V_s. There is also strong evidence in favor of a third kind of vacancy: V_f – a fast extended species, which controls the grown-in voids in silicon crystals.

Introduction

The diffusivity of vacancies (V) and of self-interstitials (I) in silicon is known from electron irradiation data [1-3] to be remarkably high even at cryogenic temperatures. The diffusivity of the neutral vacancy found by Watkins [3] is

$$D_w = (0.0012\ \text{cm}^2/\text{s}) \exp(-0.45\ \text{eV} / kT) \quad . \tag{1}$$

This expression is based on the loss of radiation-produced vacancies due to vacancy trapping by oxygen, to form VO defects (A-centres).

The diffusivity of self-interstitials is thought to be even higher than that of vacancies [2].

The data on the size and density of grown-in voids [4, 5] allow to deduce the vacancy diffusivity at the void nucleation temperature T_n (typically around 1100°C); the deduced value, $\approx 3\text{x}10^{-5}$ cm^2/s, is close to the value of $2.7\text{x}10^{-5}$ cm^2/s resulting from Eq.(1).

There are however some other data which indicate a very low diffusivity of V. The purpose of this paper is to discuss this problem and to explain why in some cases the vacancy diffusion appears to be fast while in other cases – very slow. The suggested explanation is that there are several different atomic configurations of a vacancy, and that different species are manifested in different experiments.

Enhanced self-diffusion under proton irradiation

The self-diffusion coefficient D_{sd} under hot irradiation with protons [6] was measured in thin Float-Zoned samples at some small distance from the sample surface. The quantity of interest is however the bulk value of D_{sd} reached outside a sub-micron near-surface region. The reported values of D_{sd} can be recalculated into bulk values [7]. The difference is not large but strongly dependent on the sample temperature during irradiation, 780 to 870°C. The diffusivity-concentration products, $S_V = D_V C_V$ (for vacancies) and $S_I = D_I C_I$ (for self-interstitials) have the same value (denoted by S) under steady-state conditions [7]. Since $D_{sd} = (f_I S_I + f_V S_V)/\rho$ - where ρ is the lattice site density and f_I and f_V are the correlation factors, then S is expressed as $\rho D_{sd} / (f_I + f_V)$. The sum $f_I + f_V$ is close to 1. On the other hand, S is determined by the balance of defect production at a rate G and annihilation at a rate $4 \pi r (D_I + D_V) C_I C_V$ for diffusion-limited annihilation. The reaction distance r is generally assumed to be about 0.5 nm. The resulting relation between the known value of S and the diffusivities reads:

$$D_{eff} = 1/ (1/D_V + 1/D_I) = 4 \pi r S^2 / G \quad . \tag{2}$$

The effective diffusivity D_{eff} is controlled by the smaller among D_V and D_I and was expected to be close to D_V. The values of D_{eff} calculated by Eq.(2) through known bulk values of S are plotted in Fig.1 (filled circles). The most striking feature is an extremely low D_{eff} in comparison to the vacancy diffusivity of Watkins shown by curve 1. The best-fit line, if drawn through the points, has a large slope of 2.7 eV but this activation energy is not well justified since the points in Fig.1 do not follow an Arrhenius temperature dependence. It is not clear if this is just a scatter or if there is a real change (reduction) in the slope upon increasing T. It cannot be excluded that the lower-T and higher-T portions of the $D_{eff}(T)$ curve in Fig.1 refer to different defects, for instance the former one may correspond to D_I while the latter one to D_V. Using only the higher-T points (which are more reliable), and assuming a representative value for the diffusivity prefactor, 0.01 cm^2/s, the vacancy diffusivity may be drawn through these points (curve 2 in Fig.1).

Two strongly different values for the vacancy diffusivities – represented by the curves 1 and 2 – suggest that a vacancy exists in (at least) two different atomic configurations. One (curve 1) is a fast diffuser; the same species can be also manifested in void formation process at around 1100°C. The other vacancy species (curve 2) has a diffusivity smaller by 3 orders of magnitude. We mark the two kinds of vacancies with a subscript w (Watkins vacancy) and s (slow vacancy). The diffusivity of the slow species (curve 2) is

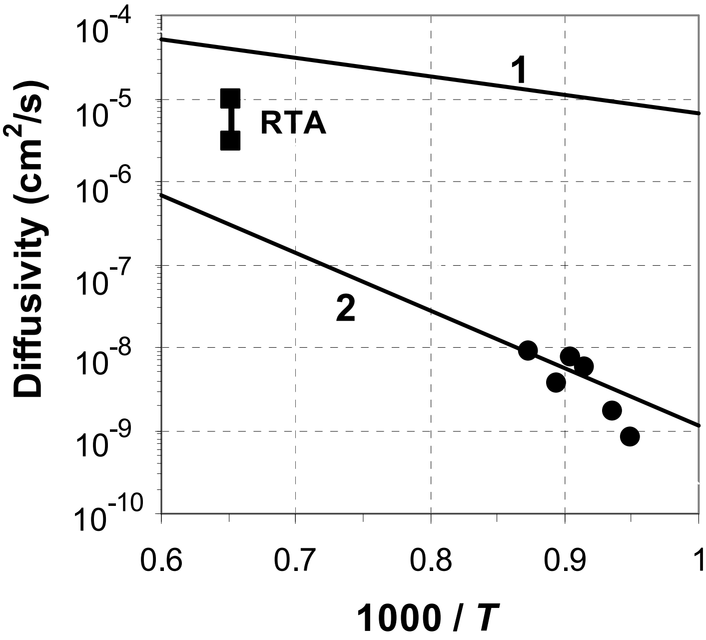

Fig.1 Effective diffusivity of intrinsic point defects V and I calculated by Eq.(2) using the measured self-diffusion coefficient (filled circles). The solid curve 1 shows the Watkins diffusivity, Eq.(1). The solid curve 2 is drawn through the higher-T points with a representative prefactor of 0.01 cm^2/s, Eq.(3). The vertical bar shows a range of apparent vacancy diffusivity deduced from RTA data.

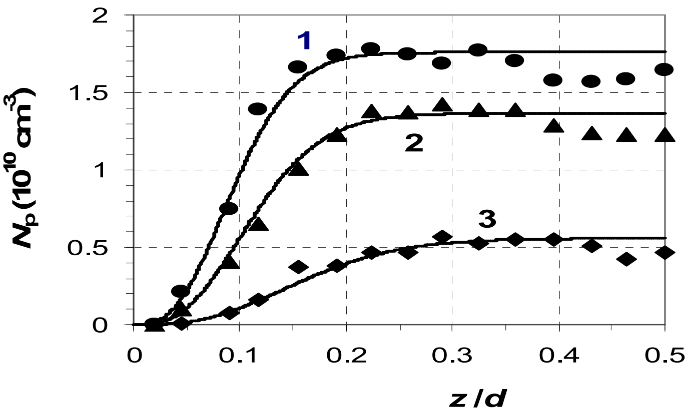

Fig.2 Precipitate density in wafers subjected to RTA in Ar at 1280°C after ref.[11] (circles, triangles and rhombic symbols correspond to the cooling rate of 70, 50 and 25 K/s, respectively). The density was averaged over the two halves of the wafer. The solid curves are plotted by Eq.(5) with the best-fit parameters: a = 0.15 cm^2/s and b = 0.004 Kcm^2/s.

$$D_s = (0.01\ cm^2/s)\exp(-1.38\ eV / kT)\ . \qquad (3)$$

If the lower-T points in Fig.1 are attributed to self-interstitials then also the self-interstitial exists in at least two forms, a fast one and a slow one.

Vacancy out-diffusion in wafers subjected to Rapid Thermal Annealing

Rapid Thermal Annealing (RTA) of silicon wafers at around 1240°C in Ar ambient is known to install a depth vacancy profile [8-10] that controls subsequent oxygen precipitation. The vacancy concentration was measured by Pt (or Au) diffusion, which results in a replacement of vacancies with substitutional metal atoms measurable by DLTS. The vacancy concentration is depleted near the wafer surface due to vacancy out-diffusion in the course of post-RTA fast cooling. The measured depth profiles of the vacancy concentration $C_V(z)$ are represented just by several significantly scattered points which makes it difficult to extract the vacancy diffusivity. A more reliable way for that is to use the depth profiles of oxygen precipitate density $N_p(z)$ developed in RTA-treated wafers after a conventional two-step anneal at 800 and 1000°C. The first step is used for a vacancy-assisted nucleation of oxide precipitates, and the second one – to enlarge the precipitate to a detectable size. The value of N_p was found [9, 10] to be proportional to C_V^m with m between 3 and 4. Therefore N_p is sensitive even to a small variation in C_V. Detailed depth profiles of N_p have been reported [11] and shown in Fig.2, for three different cooling rates $q = - dT/dt$. Vacancy out-diffusion is well manifested by a near-surface depletion of N_p. A significant reduction in the bulk value of N_p, at lower cooling rates (the curves 2 and 3 in Fig.2) is not related to vacancy out-diffusion. If it were the case, the depth profile $C_V(z)$ would be proportional to $\sin(\pi z / d)$ where d is the sample thickness, and the $N_p(z)$ profile would be sharply peaked at the wafer middle. Instead, the profile shape in Fig.2 remains practically flat excepting narrow near-surface zones. It is clear that out-diffusion affects only narrow near-surface regions while a reduction in C_V in the wafer middle is by some bulk process. One possible loss mechanism is a slow trapping of vacancies by oxygen, into VO_2 defects. In the past it was thought [10] that the formation of VO_2 is a fast process, and the vacancy defects that are active in oxide nucleation are VO_2. An alternative is that the active vacancy species are free slow vacancies V_s while the VO_2 species are inactive (not affecting the oxide particle nucleation).

Considering a vacancy loss, the vacancy profile developed in the course of fast post-RTA cooling will be approximately described by an equation:

$$\partial C_V/\partial t = D_V(T)\, \partial^2 C_V/\partial z^2 - \beta(T)\, C_V \quad . \tag{4}$$

where the last term corresponds to the vacancy loss assumed for simplicity to be proportional to C_V. The vacancy concentration at the surfaces, identical to the equilibrium concentration, decreases very fast upon lowering T, and can be taken to be zero while solving Eq.(4). The time derivative $\partial C_V/\partial t$ in this equation is expressed as $- q\, \partial C_V/\partial T$, and the solution for the final profile $C_V(z)$ is written as

$$C_V(z) = C_{V0} \exp(-a/q)\, \mathrm{erf}[z\, (q/b)^{1/2}] \quad . \tag{5}$$

where a is the integral of $\beta(T)$ over T, from low T up to the RTA temperature, and b is a similar integral of 4 $D_V(T)$. The coefficient C_{V0} is the vacancy concentration achieved in the limit of high q when the bulk loss is negligible. The two fitting parameters, a and b, are found independent one of the other: the former controls the bulk value of N_p while the latter controls the near-surface depleted zone. The $N_p(z)$ profiles - proportional to $C_V(z)^m$ and plotted with the best-fit coefficients a and b – are almost identical for $m = 3$ and $m = 4$. The results will be shown for $m = 3$. In this case the best-fit coefficients are $a = 15$ K/s and $b = 0.004$ K cm^2/s. The corresponding profiles (solid curves in Fig.2) provide a good fit to the data. The deduced value of $\int D_V dT$ is 10^{-3} Kcm^2/s. The effective temperature interval of integration corresponds to an appreciable reduction in D_V with respect to the value at T_{RTA}, and depends in particular on the activation energy for D_V. Roughly, we can estimate this interval as 100 to 300 K, and accordingly the apparent vacancy diffusivity at T_{RTA} is in a range 3×10^{-6} to 10^{-5} cm^2/s. This value – shown by a bar in Fig.1 – is essentially smaller than the Watkins diffusivity (curve 1) but essentially larger than the slow vacancy diffusivity (curve 2). A natural conclusion is that, in the vicinity of a relatively high RTA temperature, the two vacancy species, V_w and V_s, are fast converted from one to the other – and thus co-exist in the equilibrium ratio. The apparent (effective) vacancy diffusivity is then averaged over the two species:

$$D_V = (D_w C_{we} + D_s C_{se}) / (C_{we} + C_{se}) \approx D_w C_{we} / C_{se} \quad . \tag{6}$$

where the subscript e marks the equilibrium value of the two concentrations, C_w (of Watkins vacancy) and C_s (of slow vacancy). The last (simplified) expression is valid if D_V is well between the two limiting values, D_w and D_s.

The equilibrium vacancy concentrations

The vacancy contribution into silicon self-diffusivity, under equilibrium conditions, was found [4] to become dominant at $T < 950^oC$. On the other hand, the measured total self-diffusivity, in this range of T, was reported [12] to become well larger than the value extrapolated from high T. This low-T branch of self-diffusivity can be thus attributed to the vacancy. Moreover, it is likely [3, 7] that this self-diffusivity branch is due to highly mobile Watkins vacancy. In that case the diffusivity-concentration product deduced from the reported [12] self-diffusivity refers to D_wC_{we}:

$$D_wC_{we} = (2.3\text{x}10^{20}\ \text{cm}^{-1}\text{s}^{-1})\exp(-3.6\ \text{eV} / kT)\ . \tag{7}$$

With the known diffusivity, Eq.(1), the equilibrium concentration of Watkins vacancy is

$$C_{we} = (2\text{x}10^{23}\ \text{cm}^{-3})\exp(-3.15\ \text{eV} / kT)\ . \tag{8}$$

Here the prefactor only slightly exceeds the lattice site density ($\rho = 5\text{x}10^{22}\ \text{cm}^{-3}$) indicating a localized nature of the Watkins vacancy.

The value of C_{we} extrapolated to the melting point ($T_m = 1412^oC$) is $7.4\text{x}10^{13}\ \text{cm}^{-3}$. This is far smaller than the total vacancy concentration at T_m, estimated [4] to be about $10^{15}\ \text{cm}^{-3}$. It is thus concluded that this total concentration is most entirely due to slow vacancies, and then $C_{se}(T_m) \approx 10^{15}\ \text{cm}^{-3}$. We can also estimate C_{se} at the RTA temperature using Eq.(6). The RTA temperature was reported [11] to be 1280°C but the same absolute values of the precipitate density N_p were found for 1240°C [8, 9]. We accept for T_{RTA} a compromise average value, 1260°C. The D_wC_{we} product, by Eq.(7), is then $3.5\text{x}10^8\ \text{cm}^{-1}\text{s}^{-1}$. Adopting a tentative (average) value, $6\text{x}10^{-6}\ \text{cm}^2/\text{s}$, for the apparent diffusivity D_V in Eq.(6), we get a rough estimate: $C_{se} = 6\text{x}10^{13}\ \text{cm}^{-3}$ at $T_{RTA} = 1260^oC$. The equilibrium concentration drawn through these two points, T_m and T_{RTA}, appears to be

$$C_{se} = (1.8\text{x}10^{27}\ \text{cm}^{-3})\exp(-4.1\ \text{eV} / kT)\ . \tag{9}$$

A very large prefactor in this expression (in comparison to $\rho = 5\text{x}10^{22}\ \text{cm}^{-3}$) indicates an extended nature of the slow vacancy.

Contribution of various vacancy species into self-diffusivity

The contribution of slow vacancies into the self-diffusion, by Eq.(3) and (9), is much smaller than that by Watkins vacancy, Eq.(7), especially at lower T. The total contribution of the two vacancies into the D_VC_{Ve} product is about $5\text{x}10^9\ \text{cm}^{-1}\text{s}^{-1}$ at T_m.

On the other hand, the DC_e product due to vacancies, in a range of high T, was estimated [4, 13] to be much higher:

$$D_VC_{Ve} \approx (7\text{x}10^{23}\ \text{cm}^{-1}\text{s}^{-1})\exp(-4.4\ \text{eV} / kT)\ . \tag{10}$$

This expression comes from comparison of the measured self-diffusivity and the self-interstitial contribution into self-diffusion based on the known product D_IC_{Ie}. The D_IC_{Ie} product has a larger activation energy than D_{sd} and thus a significant vacancy contribution of a smaller activation energy is required, like that given by Eq.(10). At T_m, this product is about $5\text{x}10^{10}\ \text{cm}^{-1}\text{s}^{-1}$ – much larger than that due to $V_w + V_s$.

The discrepancy between a high value of $D_V C_{Ve}$ by Eq.(10) and a much smaller value of $D_V C_{Ve}$ due to $V_w + V_s$ can be explained in two ways. One is that the accuracy in measuring D_{sd} and $D_I C_{Ie}$ was insufficient to conclude the existence of a large vacancy contribution, Eq.(10). The other possibility (that seems to be more likely) is that there exists the third form of vacancy, different from V_w and V_s, and this third form is responsible for the high total $D_V C_{Ve}$ product at high T. The large value of the prefactor in Eq.(10) indicates an extended nature of this vacancy form. A large DC_e product suggests that this hypothetical kind of vacancy is highly mobile. Accordingly, this vacancy species will be denoted V_f.

The existence of V_f is further supported by considering void formation. In crystal growth, the V_w and V_s are likely to behave, like in RTA, as one community of an effective diffusivity specified by Eq.(6). This diffusivity is far smaller at the void nucleation temperature, 1100^{o}C, than the required above-mentioned value of 3x10^{-5} cm^2/s. It can be concluded that the vacancy species that controls void production is V_f – a fast extended species. During crystal growth – a process far slower than post-RTA cooling – all three kinds of vacancies may be in a mutual equilibrium – and behave as one species of an effective diffusivity – now averaged over all three types. The sum of all the DC_e products is controlled mostly by V_f while the sum of the equilibrium concentrations – mostly by V_s meaning that $D_V \approx D_f C_{fe}/C_{se}$. The $D_f C_{fe}$ product - for V_f - is given by Eq.(10). By Eq.(9) and (10), the apparent (averaged) vacancy diffusivity is expressed as

$$D_V = (3.9\text{x}10^{-4}\ \text{cm}^2/\text{s})\exp(-0.3\ \text{eV} / \text{kT}) \quad . \tag{11}$$

The activation energy for D_V is remarkably small, 0.3 eV, while an absolute value ≈ 3.1x10^{-5} cm^2/s at 1100^{o}C which is very close to the required value of 3x10^{-5} cm^2/s. In that case the conventional model of void production [5, 14] is not changed. Otherwise – if some vacancy species existed independent of the others – void production would become a much more complicated process.

Summary and discussion

The final list of different vacancy forms (configurations) inferred in the above analysis includes:

1) Fast localized species V_w (Watkins vacancy)
2) Slow extended species V_s (slow vacancy)
3) Fast extended species V_f (fast vacancy).

While the first two forms can be considered as definitely existing, the third form is hypothetical but highly probable.

In different experiments, the detected vacancies may belong to one of the three types, or to be a mixture of two or more forms.

- In electron irradiation at cryogenic temperatures [1-3], the V_w species are produced and traced. The V_s species can be also produced in parallel, but they exist independent of V_w, and escape detection.
- In proton irradiation at elevated temperatures [6], the major type of vacancy produced seems to be V_s. An exchange of V_s with other vacancy kinds is of a negligible rate and the observed enhanced self-diffusivity is controlled by V_s. An important part of this model is that the recombination of V_s with self-interstitials is fast (diffusion-limited) at around 820^{o}C. Hence it is equally fast at higher T.
- In RTA wafers, the two vacancy kinds, V_w and V_s, exist in a fast exchange, and behave as one species of an effective diffusivity intermediate between D_w and D_s. The initial vacancy concentration is first reduced due to a fast annihilation by self-interstitials. The remaining vacancies out-diffuse - but only to a small depth. Most likely, the out-diffusion lasts only for a limited range of T; below some temperature, the exchange between V_s and V_w becomes slow, and then the two species behave independent of each other. Further out-diffusion of V_s is then negligible. The V_w species are fast trapped by oxygen (just as it occurs at cryogenic temperatures [1]) and become VO and then VO_2. Also the V_s species convert into VO_2 but much more slowly. It is possible that the

conversion occurs indirectly, through slow conversion of V_s into V_w. As a result, the concentration C_s of free V_s species is reduced through the whole wafer depth in a manner that is not related to out-diffusion. The frozen-in V_s species control the nucleation of oxide precipitates at 800°C while the VO_2 species are inactive in this process.
- In crystal growth all the three vacancy species, V_w, V_s and V_f are likely to behave as one equilibrated community of an effective diffusivity D_V. The calculated D_V, at the void nucleation temperature, has a proper value ($3 x 10^{-5}$ cm^2/s) to account for the void size and density. The major contribution into D_V is by the fast species V_f.
- The elastic softening of silicon [15] was correlated to vacancy-populated crystal zones and claimed to be due to free vacancies. The relevant vacancy species is likely to be the slow vacancy V_s – which (contrary to V_w) is immune against trapping by oxygen into VO at T_{room}.

References

[1] G. D. Watkins, J. Phys. Soc. Japan, **18**, 22 (1963).

[2] G. D. Watkins, Material Science in Semiconductor Processing **3**, 227 (2000).

[3] G. D. Watkins, J. Appl. Phys. **103**, 106106 (2008).

[4] V. V. Voronkov, R. Falster, *Materials Science and Engineering*, **B 134** 227 (2006).

[5] V. V. Voronkov, ECS Transactions **18**, 945 (2009).

[6] H. Bracht, J. Fage-Peterson, N. Zangenberg, A. Nylandsted-Larsen, E.E.Haller, G.Lilly, M.Posselt, Phys. Rev. Lett. 91, 245502-1 (2003).

[7] V. V. Voronkov, R. Falster, *Material Science in Semiconductor Processing*, **15** 697 (2012).

[8] R. Falster, M. Pagani, D. Gambaro, M. Cornara, M. Olmo, G. Ferrero, P. Pichler and M. Jacob, Solid State Phenomena **57-58**, 129 (1997).

[9] R. Falster, V. V. Voronkov and F. Quast, Phys. Stat. Sol. **B222**, 219 (2000).

[10] V. V. Voronkov and R. Falster, J. Appl. Phys. **91**, 5802 (2002).

[11] M. Akatsuka, M. Okui, K.Sueoka, Nuclear Instruments and Methods in Physics Research, **B186**, 45 (2002).

[12] Y. Shimizu, M.Uematsu, K.M.Itoh, Phys. Rev. Lett. **98**, 095901-1 (2007).

[13] H. Bracht, E. E. Haller, R. Clark-Felps, Phys. Rev. Lett. **81**, 393 (1998).

[14] V. V. Voronkov, J. Crystal Growth **310**, 1307 (2008).

[15] H. Yamada-Kaneta, T. Goto, Y. Saito, Y. Nemoto, K. Sato, K. Kakimoto, S. Nakamura, Materials Sci. and Eng. **B134**, 240 (2006).