

# Doping in Semiconductors Diffusion

EE 2520

## 1 Introduction

Slide #2

1. One of the main challenges in designing a front end process for fabricating a device is the accurate control of the placement of active doping regions.
2. With the scaling of transistors, not only do the dimensions scale but the vertical dimensions such as the deep source/drain conducting junctions and the shallower tip or extension junctions also scale. There is a continuous drive to decrease the junction depth with each new technology generation

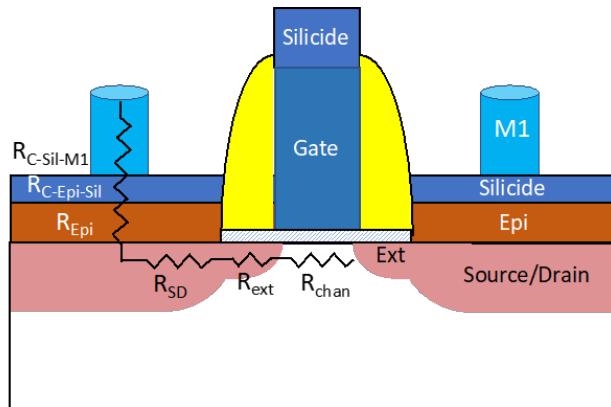


Figure 1: Junctions in MOSFET

3. There are two things that become very important and place constraints on junction doping profiles:
  - Junctions are getting shallower
  - The doping concentration is going up
4. Dopant concentration in MOS transistors have increased more than 100 fold over the past 20 years and are on the order of 1% of the silicon lattice density for current device technologies, which is the solubility limit of dopants in the lattice
5. The diffusion cycles required to electrically activate the implanted dopant atoms are often the limiting factor on the junction depth that can be obtained
6. Three types of resistance measurements are commonly used in semiconductor technology, conventional resistance  $R$  ( $\Omega$ ), resistivity ( $\Omega - cm$ ) and sheet resistance ( $\Omega/\square$ )

7. The conductivity of a doped region is defined as the relationship between the current density  $J$  and electric field  $E$

8.

$$J = nq\mu E = \sigma E = \frac{1}{\rho} E \quad (1)$$

9. The resistivity  $\rho$  defines the relationship between the electric field and the current density

$$\rho = \frac{E}{J} \quad (2)$$

10. This can be thought of as the resistance that would be measured between opposite sides of a cube of length 1 cm with units of  $\Omega - \text{cm}$

11. The sheet resistivity is resistivity per unit thickness. The units are  $\Omega/\square$

$$\rho_s = \frac{\rho}{x_j} \Omega/\square \quad (3)$$

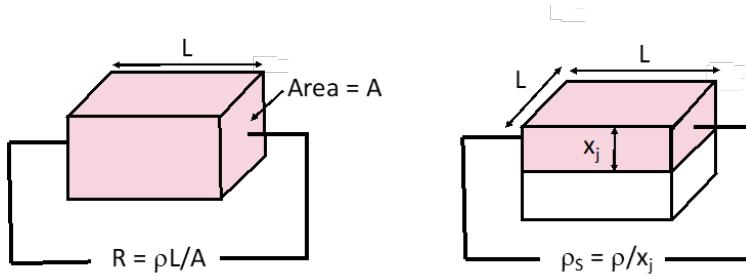


Figure 2: Various types of resistances

12. Dimensionally, the sheet resistance is expressed in ohms but has the geometrical significance that the sheet resistance is same for any square
13. A small square has less area for the current to flow through and thus higher current density but proportionally a higher field, giving the same measured resistance
14. It is extremely useful to be able to specify the resistance of the doped region, without having to specify the dimensions of the region or the depth of the junction
15. The sheet resistance of a doped region can be measured using the four point measurement technique.
16. The resistance of the source/drain region should not amount to more than a small percentage of channel resistance
17. It seems obvious that increasing the depth of the junction reduces resistance. However, in sub micron technologies, this allows drain to control the current flow. This effect is known as Drain Induced Barrier Lowering (DIBL)

18. Very shallow junctions with very high doping concentrations are required to simultaneously meet DIBL and  $\rho_s$  requirements

Slide #3

19. Junction formation using diffusion is typically a two step process

- Pre-deposition
- Drive-in

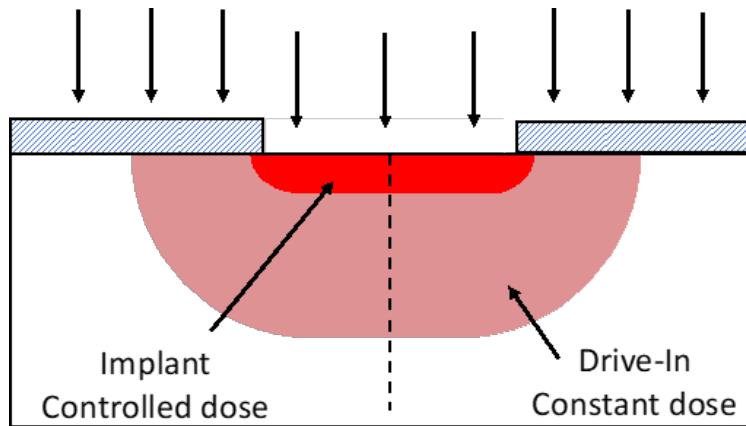


Figure 3: Steps in Diffusion

20. The dominant pre-depositions were solid phase diffusion from glass layers deposited on the wafer surface or high temperature gas phase depositions in which the wafers were placed in a furnace with a gas containing the desired doping species
21. In gas phase doping technique, the only way to make reproducible was to pin the surface concentration at the solid solubility level
22. One could always introduce an overpressure of dopant gas and the silicon would accept dopant up to its solubility and no more, so the boundary condition was fixed at the solubility
23. Small doses of dopant cannot be introduced using diffusion based doping mechanisms
24. Ion implantation is the alternative technique for doping because it provides much better control of the pre-deposition dose
25. The concept of ion implantation relies on accelerating dopant ions to high energy and shooting them into silicon. They come to stop by "billiard ball" collision with the silicon atoms
26. Damage is the big issue with ion implantation and it must be annealed at high temperatures to repair it. During this annealing step diffusion and redistribution of dopant atoms occur

## 2 Dopant Solid Solubility

1. Solid Solubility refers to the amount of an element that can dissolve in the solid without forming a separate phase
2. At low concentrations of the dopant, the dopant atoms tend to sit in the silicon lattice by substituting them
3. At high concentrations, not all the atoms are isolated and bound by four neighbouring silicon atoms and thus able to donate an electron
4. Above the solid solubility limit, the dopant atoms interact with each other as a result of electrochemical interactions and the strain fields caused by the atomic size mismatch of the dopant atoms and silicon lattice
5. This leads to the formation of clusters of dopant atoms which are not located on lattice sites and hence do not contribute to electrically active sites
6. The doping concentrations needed for current process technologies are already at the solid solubility limit. The solubility of various dopants is shown Fig.4

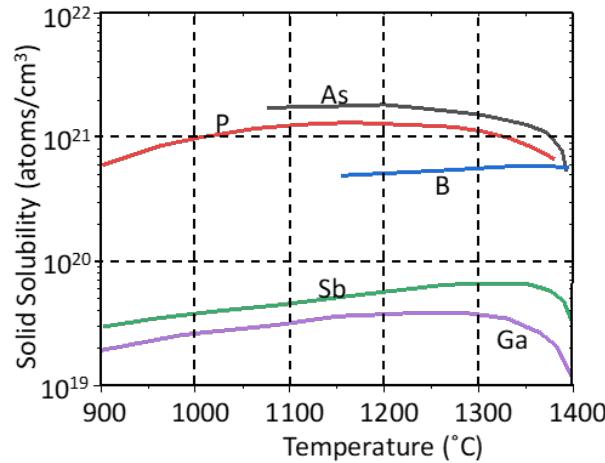


Figure 4: Dopant Solid Solubility of various dopants

7. Electrical solubility is a measure of how many free electrons or holes available as a result of doping process
8. From a practical point of view, the maximum electrically active doping concentration is around  $2 \times 10^{20} \text{ cm}^{-3}$
9. For boron, it is close to the solid solubility, but the electrical solubility is an order of magnitude below the solid solubility for phosphorous and arsenic
10. Dopants above the electrical soluble limit form an inactive complex that is electrically neutral and does not contribute to free carriers

11. For example, at high doping concentrations, arsenic atoms remain on the substitutional sites but adjoin a vacancy, which leave the arsenic three-fold coordinated with the silicon lattice while retaining two electrons in a dangling bond, for a full shell of eight electrons, thus rendering the complex electrically inactive as shown in Fig.5.

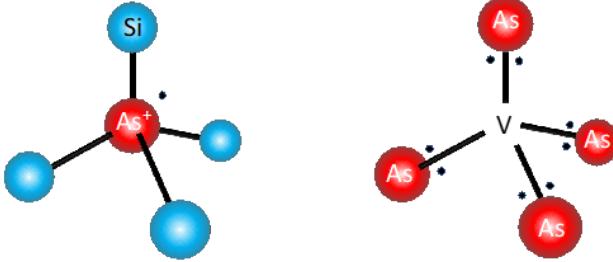


Figure 5: Inactive Complexes at higher concentrations

### 3 Diffusion from a Macroscopic point of view

Slide #4

1. The macroscopic point of view considers the overall motion of a dopant profile and predicts the amount of motion by solving a diffusion equation subject to boundary conditions
2. In microscopic behavior approach, the motion of the dopant atoms is considered at the atomic scale, and is related to the overall motion of the whole profile to the motion of unseen individual atoms based on the interactions of atoms and point defects in the lattice. This is needed to explain the complex behavior in modern devices
3. Consider a block of material in which the concentration is a function of distance. Fick first law relates the flux and the concentration gradient. Flux ( $\text{atoms cm}^{-2}\text{s}^{-1}$ ) is proportional to the change in the concentration gradient. The proportional constant is the diffusivity.

$$J = -D \frac{\partial C}{\partial x} \quad (4)$$

4. D is the diffusion coefficient and its units are  $\text{cm}^2/\text{V} - \text{s}$
5. The negative sign indicates that the flow is down the concentration gradient. Atoms diffuse from high concentration toward lower concentration
6. Fick's second law is a fundamental conservation of law for matter, which says that the increase in concentration in a cross section of unit area with time is the difference between the flux into the volume and the flux out of the volume

$$\frac{\Delta C}{\Delta t} = \frac{\Delta F}{\Delta x} = \frac{F_{in} - F_{out}}{\Delta x} \quad (5)$$

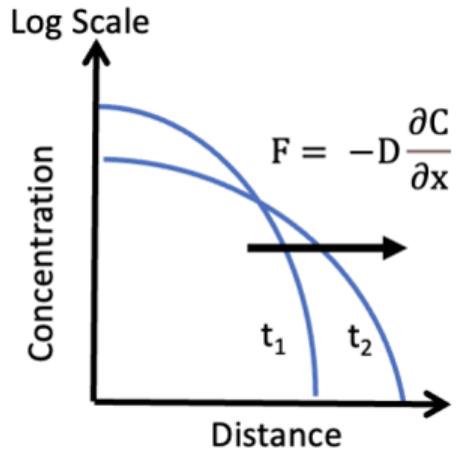


Figure 6: Fick's first law

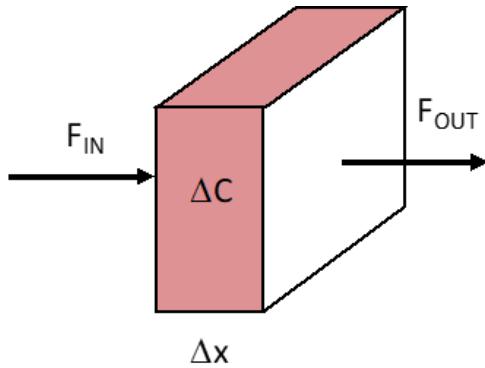


Figure 7: Fick's Second law

$$\boxed{\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}} \quad (6)$$

7. The analytical solutions with constant diffusivity are usually valid for concentrations below  $n_i$  (ex: at 1000 °C,  $n_i=7.14\times10^{18} \text{ cm}^{-3}$ , because the intrinsic population of electrons and holes dominates the mobile charges donated by the dopants).
8. At concentrations above  $n_i$ , the diffusion coefficient is concentration dependent.
9. In other words constant diffusion coefficient may not hold good in all circumstances
10. The mechanism of diffusion of boron or phosphorous is hypothesized to be through interactions with vacancies and interstitials
11. Constant dopant diffusivity implies that the vacancy or interstitial positions are constant with positions.
12. Point defect populations depend on the position of the Fermi level because of the charged defects in the material and thus depend on the doping level. If there is a high concentration

profile at the surface that falls off in the bulk, the population of defects will be higher in the high concentration region than the lower concentration region

13. The generalized three dimensional diffusion equation is given by

$$\boxed{\frac{\partial C}{\partial t} = \nabla \cdot F = \nabla \cdot (D \cdot \nabla C)} \quad (7)$$

## 4 Analytical Solutions

### 4.1 Guassian solution in an infinite medium

1. Consider a spike or delta function of dopant in the middle of a lightly doped region

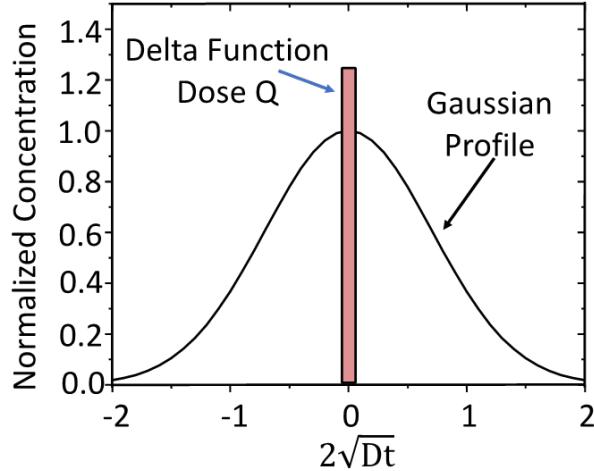


Figure 8: Delta function in the middle

2. Taking the origin to be at the delta function, the boundary conditions are

$$\boxed{C \rightarrow 0 \text{ as } t \rightarrow 0 \text{ for } |x| > 0} \quad (8)$$

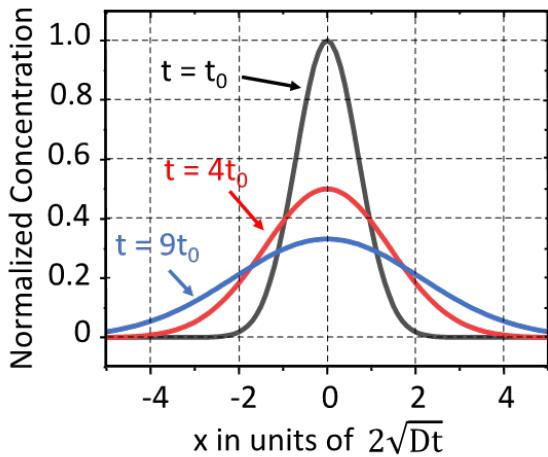
$$\boxed{C \rightarrow \infty \text{ as } t \rightarrow 0 \text{ for } |x| = 0} \quad (9)$$

$$\boxed{\int_{-\infty}^{+\infty} C(x, t) dx = Q} \quad (10)$$

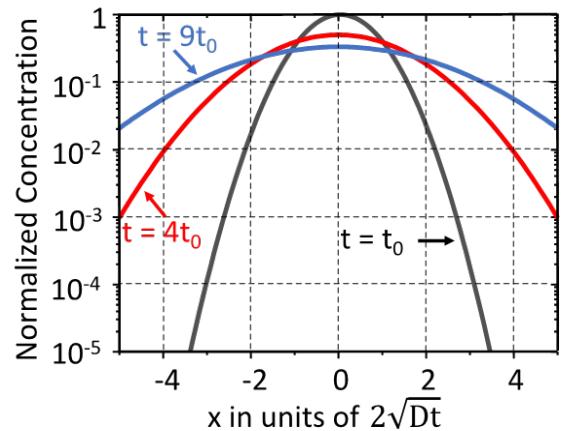
3. Q is the total quantity or dose of the dopant that is contained in the spike

4. The solution of Fick's second law that satisfies these boundary conditions is

$$\boxed{C(x, t) = \frac{Q}{2\sqrt{\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right) = C(0, t) \exp\left(-\frac{x^2}{4Dt}\right)} \quad (11)$$



(a) Gaussian Profile - Linear Scale



(b) Gaussian Profile - log Scale

Figure 9: Gaussian Profile

5. The Gaussian Profiles are shown in fig.12b
6. The Dose Q remains constant
7. The peak concentration  $C(0, t)$  decreases as  $1/\sqrt{t}$
8. when the distance from the origin is  $x = 2\sqrt{Dt}$ , the peak concentration falls by  $1/e$
9. An approximate measure of how far the dopant has diffused is given by  $2\sqrt{Dt}$

## 4.2 Gaussian solution near a surface

- (a) This is a simplification of previous case

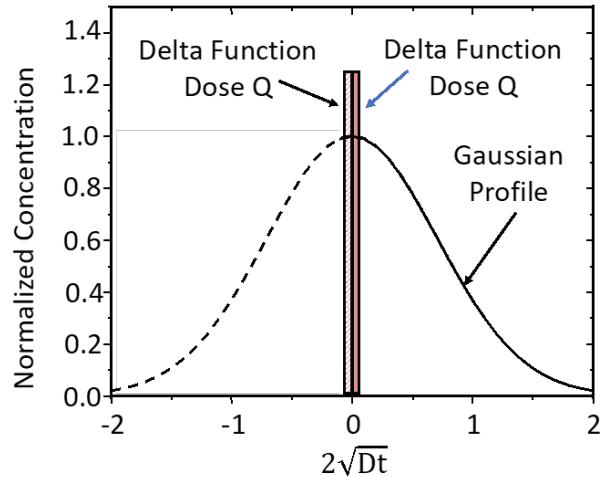


Figure 10: Delta function at the surface

(b) The diffusion of the dopant near a surface is given by

$$C(x, t) = \frac{Q}{\sqrt{\pi D t}} \exp\left(-\frac{x^2}{4 D t}\right) = C(0, t) \exp\left(-\frac{x^2}{4 D t}\right) \quad (12)$$

(c) The surface concentration is given by

$$C(0, t) = \frac{Q}{\sqrt{\pi D t}} \quad (13)$$

- (d) These surface boundary conditions may be unrealistic in practise because generally there is segregation into a deposited or growing oxide layer or evaporation into the ambient, requiring numerical simulation for accurate results
- (e) This is a useful analytical solution when a dose  $Q$  is introduced "near" the surface and annealed long enough that the initial distribution is reasonable approximated by a delta function.

Slide #5

### 4.3 Error-function solution in infinite medium

- (a) Consider the case when there is infinite source of dopants  
 (b) In this case the boundary conditions are

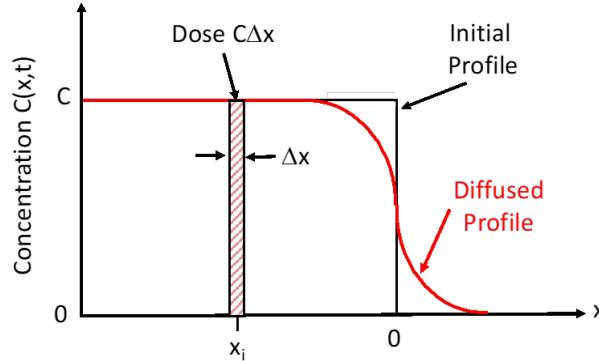


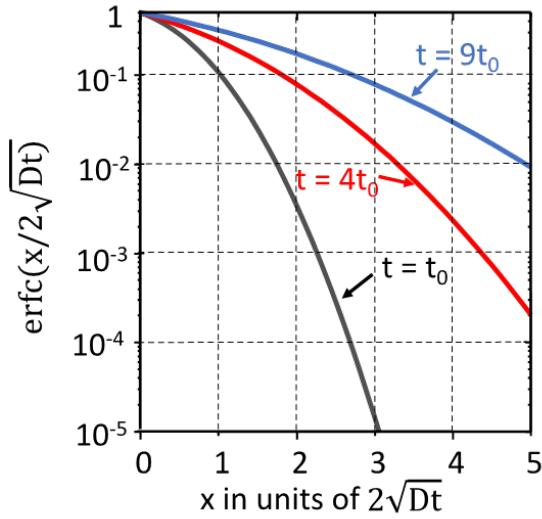
Figure 11: Error function in infinite medium

$$C(x, t) = 0 \text{ at } t = 0 \text{ for } x > 0 \quad (14)$$

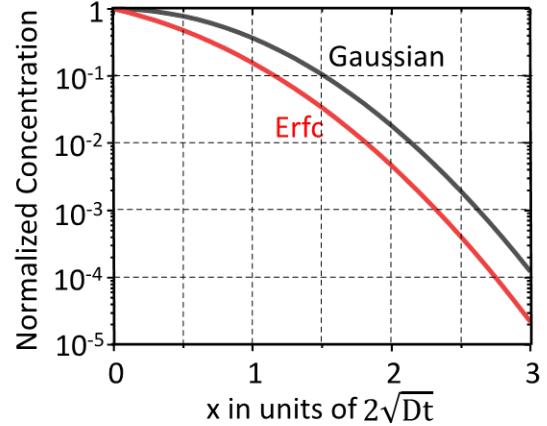
$$C(x, t) = 0 \text{ at } t = 0 \text{ for } x < 0 \quad (15)$$

(c) The solution of the diffusion equation from infinite source becomes

$$C(x, t) = \frac{C}{2} \left[ 1 - \operatorname{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \right] = \frac{C}{2} \operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) \quad (16)$$



(a) erfc Profile



(b) Comparison of erfc and Gaussian profiles

Figure 12: Diffusion Profiles

#### 4.4 Error function solution near a surface

- (a) The error function also describes the diffusion kinetics when the profile is characterized by a constant surface concentration at all times
- (b) Such a diffusion profile might occur if the diffusion occurred from a gas ambient with a concentration above the solid solubility of the dopant in the solid
- (c) If a source concentration is held constant at a value  $C_s$ ,

$$C(x, t) = C_s \left[ \operatorname{erfc} \left( \frac{x}{2\sqrt{Dt}} \right) \right] \quad (17)$$

- (d) The error-function solution is approximately triangular on a linear scale. The dose introduced can be approximated by the area of a triangle of height  $C_s$  and a base equal to the diffusion distance  $2\sqrt{Dt}$ . The dose  $Q$  is given by

$$Q = \frac{2C_s}{\pi} \sqrt{Dt} \quad (18)$$

- (e) The error-function solution applies when there is an infinite supply of dopant, which implies that an increasing dose of dopant is introduced into the diffusion process to maintain a constant surface concentration
- (f) The Gaussian solution applies when the initial dose of dopant is fixed. Consequently, the surface concentration must drop as the dopant diffuses deeper into the bulk
- (g) Modern structures employ doped regions in which concentration-dependent diffusion, electric field effects, dopant segregation and complicated point-defect driven diffusion processes take place

- (h) All of these effects generally require numerical methods to calculate the resulting dopant profiles

Slide #6

## 5 Intrinsic Diffusion Coefficients of Dopants in Si

- (a) Diffusion coefficients are a measure of how fast dopants move in a solid
- (b) The diffusion coefficients of common impurities in silicon are found to go exponentially faster at higher temperatures
- (c) This is because the native point defects increases exponentially with temperature and thus vacancies and interstitials increase exponentially
- (d) Diffusivity as a function of temperature is given by

$$D = D_0 \exp\left(-\frac{E_a}{kT}\right) \quad (19)$$

- (e) Plots of diffusion coefficients as a function of temperature are plotted in Fig.13

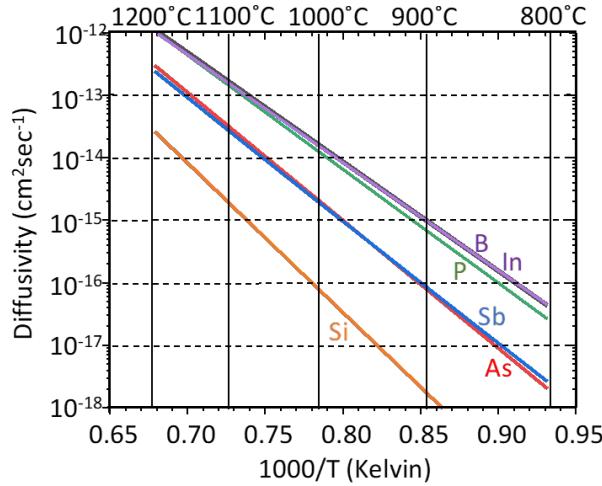


Figure 13: Diffusion Coefficient as a function of temperature

- (f) The intrinsic concentration in silicon is quite high at normal diffusion temperatures (ex: at 1000 °C,  $n_i = 7.14 \times 10^{18} \text{ cm}^{-3}$ ), the material behaves as intrinsic material
- (g) The analytical solutions and arrhenius plots are valid if the doping is intrinsic at the diffusion temperatures

- (h) Shallow junctions necessitates slow diffusers. For N-type regions, this has led to the dominance of As as a dopant because it has both a small D and high solid solubility. For P-type regions, B is unfortunately the only dopant with a high solid solubility, and its higher diffusivity means that fabricating shallow P-type regions is usually more difficult than forming shallow N-type regions

## 6 Effect of Successive Diffusion Steps

- (a) Since there are often multiple diffusion steps in full IC process, they must be added in some way before the final profile can be predicted
- (b) If all the diffusion steps occurred at a constant temperature where the diffusivity is the same, then the effective  $Dt$  product is given by

$$(Dt)_{eff} = D_1 t_1 + D_1 t_2 + \dots \quad (20)$$

- (c) In other words, doing a single step in a furnace for a total time of  $t_1 + t_2$  is same as doing two separate steps, one for time  $t_1$  and other for time  $t_2$
- (d)  $Dt$  product is considered to be a measure of thermal steps or thermal budget that is used in a process
- (e) Consider a dopant that is diffused at a temperature  $T_1$  with diffusivity  $D_1$  for time  $t_1$  and then at a temperature  $T_2$  with diffusivity  $D_2$  for time  $t_2$ , the effective  $Dt$  is given by

$$(Dt)_{eff} = D_1 t_1 + D_2 t_2 + \dots \quad (21)$$

- (f) The highest temperature steps in the process dominate the thermal budget since the diffusion coefficient is an exponential function of temperature

## 7 Design and Evaluation of Diffused Layers

- (a) The key parameters that are important in designing a diffused layer are the sheet resistance, the surface concentration and the junction depth
- (b) There are useful design curves known as Irvin's curves that capture the surface concentration of dopant as a function of background concentration and the sheet resistivity
- (c) A surface concentration at the solid solubility limit might imply that the profile is likely to be described by an error function type solution for a constant surface concentration
- (d) A low surface concentration might mean that the profile was driven in with a long heat cycle indicating a Gaussian type solution is applicable

## 8 The Physical Basis of Atomistic Diffusion

- (a) The atomic scale of understanding forms the physical basis for the modern day process tools.
- (b) Point defects (vacancies and interstitials) are key players in dopant diffusion at the atomic level. Understanding point defect behavior offers a unified explanation for anomalous diffusion effects.
- (c) Fig.14 illustrates the basic dopant diffusion mechanisms involving point defects.

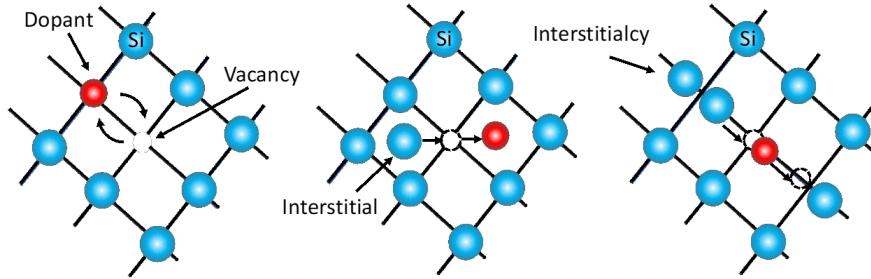


Figure 14: (a) Schematic of vacancy-assisted diffusion mechanism (left), (b) of interstitial-assisted kick-out diffusion (center) and (c) interstitialcy-assisted diffusion mechanisms (right)

- (d) A vacancy next to a dopant atom allows the dopant to hop to an adjacent site, enabling diffusion
- (e) A silicon interstitial can “kick out” a substitutional dopant, making the dopant mobile
- (f) The dopant diffuses through a random walk, alternating between mobile interstitial and stable substitutional states.
- (g) A dopant and interstitial can also migrate as a bound pair, known as the interstitialcy mechanism
- (h) In this mechanism, the dopant and interstitial share a lattice site and move together until they split.
- (i) Both kick-out and interstitialcy involve initiation and release by a silicon interstitial.