CONTENTS

1  Diffusion and Fick’s First Law.

2  Diffusion and Fick’s Second Law.
   2.1 Fick’s Second Law and Diffusion from a Constant Surface Source.
   2.2 Fick’s Second Law and Diffusion from a Limited Supply.

3  Silicon Dioxide (SiO₂)
   3.1 The Oxide Growth.
   3.2 Growth Kinetics
      A. Initial Linear Growth.
      B. Parabolic Growth.
      C. Overall Growth Kinetics.

ADVANCED TOPICS

4  Derivation of Fick’s First and Second Laws

5  Derivation of the Maximum in Gaussian diffusion and concentration: The Gaussian wind.

NOTATION
1 Diffusion and Fick's First Law

It is well known that, by virtue of their random motions, gas particles diffuse from high to low concentration regions. When a perfume bottle is opened at one end of a room, the molecules diffuse out from the bottle and, after a while, can be smelled at the other end of the room. Whenever there is a concentration gradient of particles, whether electrons, atoms or molecules, there is a net diffusional motion of particles in the direction of decreasing concentration. The origin of diffusion lies in the random motions of the particles. To quantify particle flow, we define a particle flux, \( \Gamma \), just like current, as the number of particles crossing unit area per unit time. Thus if \( \Delta N \) particles cross an area \( A \) in time \( \Delta t \) then, by definition, the particle flux is

\[
\Gamma = \frac{\Delta N}{A \Delta t} \tag{1}
\]

In describing flux, \( \Gamma \), we must distinguish between species of particles. For example, the flux of He atoms refers only the number of He atoms crossing unit area in unit time. Suppose that the concentration of a given species of particles at some time \( t \) in a material decreases in the \( x \)-direction and has the concentration profile, \( C(x) \) vs. \( x \), shown in Figure 1-1.

Due to their thermal energy, particles will be moving around randomly and colliding with each other. Because of these random motions we can assume that, at any instant, at any one location such as at \( A \) about half the particles would be moving towards the left and half towards the right. There are more particles at position \( A \) than at \( B \). This means that at any instant there are more particles at \( A \) moving towards the right.
than there are particles at $B$ moving towards the left. There is therefore a net flow or net flux of particles from $A$ to $B$. If the concentrations were equal then there would be no net flux. At both $A$ and $B$ half the particles would be moving left and half moving right. It is apparent that the net flux depends on the concentration difference which in turn depends on the concentration gradient. All experiments show that the magnitude of the net flux of particles in the $x$–direction, from $A$ to $B$, depends on the concentration gradient. If $\Gamma$ is the net particle flux at point $x$ and $C$ is the concentration at that point then according to Fick's first law, the net particle flux $\Gamma$ at a position $x$ is proportional to the concentration gradient

$$\Gamma = -D \frac{dC}{dx}$$

\textit{Fick's first law} \hspace{1cm} (2)

where the negative sign ensures that the flux direction is from high to low concentration regions (negative gradient) and $D$ is a proportionality constant called the \textbf{diffusion coefficient}. $D$ depends on the species of diffusing particles and the medium in which they are diffusing. Equation (2), \textit{Fick's first law}, represents the relationship between the net particle flux and the driving force which is the concentration gradient. The diffusion coefficient $D$ for atomic diffusions in solids is thermally activated,

$$D = D_o \exp \left( -\frac{E_A}{kT} \right)$$

\textit{The diffusion coefficient and thermal activation} \hspace{1cm} (3)

where $D_o$ is a constant and $E_A$ is the activation energy, $k$ is the Boltzmann constant and $T$ is the temperature. Both $D_o$ and $E_A$ depend on the nature of the diffusion process in the crystal. Although, strictly, $E_A$ in eqn. (3) is in joules per atom it is typically quoted in many tables as electronvolts per atom. In materials engineering we frequently write eqn. (3) as

$$D = D_o \exp \left( -\frac{Q_{\text{diff}}}{RT} \right)$$

\textit{(3–a)}

where $Q_{\text{diff}}$ is the activation energy in joules per mole and $R$ is the gas constant, $N_A k$, where $N_A$ is Avogadro's number. To understand the origin of eqn. (3) we consider an impurity atom at an interstitial site $A$ in a crystal as shown in Figure 1-2.

Diffusion of an interstitial impurity atom in a crystal from one void to a neighboring void. The impurity atom at position $A$ must possess an energy $E_A$ to push the host atoms away and move into the neighboring void at $B$.

This figure also shows the variation of the potential energy ($PE$) of this atom from its present interstitial site at $A$ to a neighboring one at $B$. The $PE$ at $A$ is a minimum as this is an interstitial site, an equilibrium position for the impurity. If the impurity atom tries to move to $B$, it has to first break its bonds and then push apart the two neighboring atoms to "squeeze" its way through. This requires work. The $PE$ of the impurity atom is maximum when the repulsive forces are maximum at $A^*$. It is again at minimum at the interstitial site $B$. There is therefore a $PE$ barrier, $E_A$, against the motion of the impurity atom from $A$ to $B$. 

![Figure 1-2](image-url)
The $PE$ barrier $E_A$ represents the work needed to move the impurity from $A$ to $A^*$ from where the impurity will "roll" down to $B$. Due to thermal energy all the atoms in the crystal would be vibrating with a distribution of energies.

The probability that the impurity atom has just sufficient energy to overcome the $PE$ barrier, $E_A$, and go over $A^*$ to $B$ is proportional to the Boltzmann factor $\exp(-E_A/kT)$. The higher the temperature, the wider the energy distribution of vibrating atoms and greater the probability of surmounting the $PE$ barrier $E_A$. Thus the diffusion coefficient is thermally activated. It is clear that the diffusion coefficient depends sensitively on the $PE$ barrier $E_A$ which in turn depends on the diffusion process.

The example in Figure 1-2 is called interstitial impurity diffusion. There are a number of diffusion mechanisms for the migration of atoms in a crystal structure as depicted in Figure 1-3. In interstitial impurity diffusion an interstitial impurity atom in the bulk can jump from one interstitial site to a neighboring one as shown in (a).

Some diffusion processes in crystals. (a) An interstitial impurity atom diffuses from one interstitial to another in the volume. (b) An impurity atom diffuses along a grain boundary. (c) An impurity atom diffuses on the surface. (d) A neighboring atom can jump, diffuse, into a vacancy and thereby displace the vacancy. This is vacancy diffusion. (e) A host atom can diffuse by jumping into an interstitial site.

Figure 1-3

If the impurity atom is in a grain boundary as shown in (b) then the diffusion process along the grain boundary becomes easier as there are many voids and broken bonds at grain boundaries and a smaller $PE$ barrier has to be overcome, or less work has to be done. Grain boundaries therefore provide easy diffusion paths. The diffusion of impurities on the crystal surface is even easier because even less work is needed to break bonds to jump from one site to another as indicated in (c). In other words, $E_A$ is smaller.

Vacancies aid atomic migrations because a neighboring atom can jump into a vacancy as shown in (d). The vacancy becomes displaced and the whole process can then repeat itself to take one more atom and further displace the vacancy. This is called vacancy diffusion because effectively the vacancy is diffusing through the bulk. Another possibility is the displacement of a substitutional (or a host) atom into an interstitial site as shown in (e).

Each process has its own characteristic diffusion coefficient. Typically over a given range of temperatures for a given species of diffusing atoms one type of diffusion process dominates over others. If we were to plot $\ln(D)$ against the reciprocal temperature, $1/T$, from eqn. (3), we would get a straight line with a slope $-E_A/k$.

Diffusion coefficients are invariably plotted this way because the plots yield straight lines. Figure 1-4 shows the expected behavior of the diffusion coefficient for impurity diffusion in the bulk, along grain boundaries and on the surface for (a), (b) and (c) respectively in Figure 1-3. We note that the activation energy, as indicated by the slope, decreases with increasing ease of diffusion from volume (bulk), grain boundaries to surface.
Example 1–1: Steady state diffusion of hydrogen through a palladium sheet

Palladium has the ability to absorb hydrogen gas 800–900 times its volume. The hydrogen molecules dissociate on the Pd surface and become dissolved as H–atoms in the Pd structure taking up interstitial sites. When there is pressurized hydrogen gas on one side, A, of a thin Pd sheet then, as a result of diffusion, hydrogen gas appears on the other side, B, of the sheet as shown in Figure 1-5.

A schematic illustration of hydrogen gas diffusion through a thin Pd sheet under steady state conditions. There is a concentration gradient of H atoms from surface A to surface B. The pressure $P_A$ is greater than the pressure $P_B$.

If the pressures $P_A$ and $P_B$ at A and B are kept constant with $P_A > P_B$ then eventually steady state is reached in which there is a constant hydrogen concentration gradient and there is a constant net flux of hydrogen atoms from A to B. Hydrogen can be purified by diffusing the gas through a palladium sheet. In a certain purification process, the concentration of hydrogen, during a steady state operation, at the high and low concentration sides of the Pd sheet have been found to be 1 at.% and 0.2 at.%. How many kilograms of hydrogen can pass through per day through a 5 mm thick Pd membrane having an area of $0.5 \times 0.5$ m$^2$ at 500°C? The diffusion coefficient of hydrogen in Pd has $Q_{\text{diff}} = 24.07$ kJ mole$^{-1}$ and $D_o = 4.5 \times 10^{-7}$ m$^2$ s$^{-1}$. Atomic mass and density of Pd are 106.4 and 12.0 g cm$^{-3}$.

**Solution:**

The concentration of hydrogen at A and B is given as atomic percentage. To obtain the number of H–atoms per unit volume we first need to find the atomic concentration in Pd. If $N_A$ is Avogadro's number,
$M_A$ is the atomic mass and $\rho$ is the density of Pd then the number of Pd atoms per unit volume is

$$n_{\text{Pd}} = \frac{\rho N_A}{M_A} = \frac{(12.0 \times 10^3 \text{kg m}^{-3})(6.022 \times 10^{23} \text{mol}^{-1})}{(106.4 \times 10^{-3} \text{kg mol}^{-1})} = 6.8 \times 10^{28} \text{ atoms m}^{-3}.$$ 

The concentrations of H–atoms at $A$ and $B$ are,

$$C_A = 0.01 n_{\text{Pd}} = 0.01(6.8 \times 10^{28}) = 6.79 \times 10^{26} \text{ H–atoms m}^{-3}.$$ 

$$C_B = 0.002 n_{\text{Pd}} = 0.002(6.8 \times 10^{28}) = 1.36 \times 10^{25} \text{ H–atoms m}^{-3}.$$ 

As the pressures at $A$ and $B$ are constant we can take the concentrations at $A$ and $B$ to be constant as well. The concentration gradient of H–atoms through the Pd sheet is then

$$\frac{dC}{dx} = \frac{C_A - C_B}{d} = \frac{(6.79 \times 10^{26}) - (1.36 \times 10^{26})}{(5 \times 10^{-3})} = 1.087 \times 10^{29} \text{ m}^{-4}.$$ 

We can find the flux from Fick's first law but we need the diffusion coefficient $D$ at 500° C,

$$D = D_o \exp \left(-\frac{Q_d}{RT}\right) = (4.5 \times 10^{-7})\exp \left(-\frac{24070}{(8.3145)(773)}\right) = 1.063 \times 10^{-8} \text{ m}^2 \text{s}^{-1}.$$ 

and

$$\Gamma = D \frac{dC}{dx} = (1.063 \times 10^{-8} \text{ m}^2 \text{s}^{-1})(1.087 \times 10^{29} \text{ m}^{-4}) = 1.155 \times 10^{21} \text{ H–atoms m}^{-2} \text{s}^{-1}.$$ 

If $m_h$ is the mass of the hydrogen atom, then the mass of hydrogen flowing through an area $A$ in time $t$ is

$$\text{Mass} = \Gamma \cdot A \cdot t \cdot m_h.$$ 

Mass of hydrogen flowing per day is:

$$\text{Mass per day} = (1.155 \times 10^{21} \text{ H–atoms m}^{-2} \text{s}^{-1})(0.25 \text{ m}^2)(1 \times 60 \times 60 \times 24 \text{ s})(1.67 \times 10^{-27} \text{ kg}) = 0.042 \text{ kg}.$$ 

It takes about 24 days for 1 kg of hydrogen to flow.

2 Diffusion and Fick's Second Law:

2.1 Fick's Second Law and Diffusion from a Constant Surface Source

Fick's first law is useful under steady state conditions for calculating the flux from the concentration gradient but does not allow the concentration $C(x, t)$ at a point $x$ at time $t$ to be determined when the conditions are such that the concentration profile is changing with time as illustrated in Figure 2-1. At time $t = 0$, a long crystal such as Pd is suddenly exposed on one surface at $x = 0$ to a gas of atoms, say He atoms. We will assume that the Pd crystal has a small concentration $C_o$ of He atoms uniformly dissolved in the structure prior to exposure. Suppose that the gas pressure and hence the surface concentration, $C(0,t)$, of He atoms is maintained constant at $C_s$. At time $t = 0$, the concentration profile $C(x, t)$ of He–atoms is a step function, decreasing to zero just past the surface. As a result of diffusion He–atoms begin to migrate towards the far end of the crystal and with time the concentration $C(x, t)$ profile increases, or moves higher.

Two typical profiles at times $t = t_i$ and $t = t_j > t_i$ are shown in Figure 2-1. As the crystal is very long, the concentration at the far end ($x = \infty$) is still $C_o$. The conditions we have imposed, constant concentration, $C_o$, at the surface, and pre–existing concentration, $C_o$, in the bulk and a long crystal are only boundary conditions for this particular problem. The kinetics of the diffusion process, however, is described by Fick's second law, that relates the space and time derivatives of the concentration $C(x, t)$, through a partial differential equation:
\[
D \frac{\partial^2 C(x,t)}{\partial x^2} = \frac{\partial C(x,t)}{\partial t}
\]

*Fick's Second Law* (4)

\(\partial C/\partial t\) represents differentiation with respect to \(t\) only, keeping \(x\) constant, \(\partial C/\partial x\) represent differentiation with respect to \(x\) keeping \(t\) constant.

\[C(x,t) = \begin{cases} C_s & \text{at the surface at any time} \\ C(x,0) = C_o & \text{at } t = 0 \text{ anywhere in the bulk} \\ C(\infty,t) = C_o & \text{at any time at far end} \end{cases}
\]

These boundary conditions define what is called **diffusion from an unlimited supply** (or constant surface source). Given the boundary conditions in eqn. (5), Fick’s second law in eqn. (4) can be solved to obtain the concentration \(C(x,t)\) at \(x\) at time \(t\),

\[
\frac{C(x,t) - C_o}{C_s - C_o} = 1 - \text{erf} \left( \frac{x}{2\sqrt{D}t} \right)
\]

constant surface source (6)

where erf(\(z\)), called the **error function of \(z\)**, is a function in which \(z\) is the argument and is tabulated like
many other functions such as \(\log(z)\), \(\sin(z)\) etc. It is clear that in this particular diffusion case, \(z = x^2/(2\sqrt{Dt})\). The error function is defined in terms of an integration as

\[
\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-y^2) dy
\]

\(\text{Error function of } z \) (7)

Notice that the argument \(z\) of the function is in the upper limit of the integration and \(y\) is simply a variable of integration which can just as well be any symbol. \(1 - \text{erf}(z)\) is called the **complimentary error function** and is denoted by \(\text{erfc}(z)\). Many mathematical software packages these days have \(\text{erf}(z)\) function as one of their standard functions (e.g. Matlab and Mathcad). The concentration profiles, \(C(x, t)\), in Figure 2-1 therefore follow eqn. (6) which can be used to predict the concentration at any point at any time in the crystal as shown in Example 2-1.

---

**Example 2–1: Diffusion and Carburization of Steels**

The surfaces of steels for certain applications are hardened by diffusing carbon atoms into the steel at elevated temperatures. Typically a steel component is exposed to methane gas at high temperatures. The methane molecules dissociate at the surface and the carbon diffuses into the steel. Suppose that 0.15%C steel (iron with 0.15%C) is carburized at 750°C for 2 hours during which the surface concentration of carbon is 1.5%. What is the carbon concentration at 1 mm beneath the surface? What is the time required to achieve the same concentration by carburizing at 600°C. The diffusion coefficient of carbon in BCC iron has \(Q_{\text{diff}} = 80.3 \text{ kJ mole}^{-1}\) and \(D_o = 6.2 \times 10^{-7} \text{ m}^2 \text{s}^{-1}\).

**Solution:**

As the surface concentration is maintained fixed the diffusion process obeys eqn. (6). At the surface, \(C_s = 1.5\%.\) At time \(t = 0\), the concentration of carbon in the bulk is \(C_o = 0.15\%.\)

At \(T = 750 \text{ K} + 273 \text{ K} = 1023 \text{ K}\) we have

\[
D = D_o \exp\left(-\frac{Q_{\text{diff}}}{RT}\right) = (6.2 \times 10^{-7}) \exp\left(-\frac{80300}{(8.315)(1023)}\right) = 4.90 \times 10^{-11} \text{ m}^2 \text{s}^{-1}.
\]

At \(x = 10^{-3} \text{ m} (1 \text{ mm})\), at time \(t = 2 \times 60 \times 60 \text{ seconds} (2 \text{ hrs.})\). The argument of \(\text{erf}(z)\) in eqn. (6) is

\[
z = \frac{x}{2\sqrt{Dt}} = \frac{1.0 \times 10^{-3}}{2\sqrt{(4.90 \times 10^{-11})(2 \times 60 \times 60)}} = 0.8413
\]

To find \(C\) at \(x = 1 \text{ mm}\) and \(t = 2 \text{ hrs.},\) we need \(\text{erf}(z)\) in eqn. (6). We can carry out a numerical integration of eqn. (7) or use a mathematical software package (or a calculator that has the error function),

\[
\text{erf}(z) = \text{erf}(0.8413) = \frac{2}{\sqrt{\pi}} \int_0^{0.8413} \exp(-y^2) dy = 0.7659
\]

From eqn. (6),

\[
\frac{C(1 \text{mm}, 2 \text{hrs.}) - 0.15}{1.5 - 0.15} = 1 - 0.7659 \text{ or } C(1 \text{ mm}, 2 \text{ hrs.}) = 0.466\%. \text{ At the lower temperature of } 600^\circ \text{ C the diffusion coefficient will be smaller. At } T' = 600 \text{ K} + 273 \text{ K} = 873 \text{ K}\text{ we have,}
\]

\[
D' = D_o \exp\left(-\frac{Q_{\text{diff}}}{RT'}\right) = (6.2 \times 10^{-7}) \exp\left(-\frac{80300}{(8.315)(873)}\right) = 9.68 \times 10^{-12} \text{ m}^2 \text{s}^{-1}.
\]
We need the time $t'$ that it takes for the concentration at $x = 1$ mm to reach the same value of 0.466%. The concentrations $C(x, t)$, $C_s$ and $C_o$ are therefore the same, 0.466%, 1.5% and 0.15% respectively. The left hand side of eqn. (6) thus remains the same which means that we must have the same $z$ at time $t'$.

$$z = \frac{x}{2\sqrt{D't'}} = \frac{x}{2\sqrt{Dt'}} \quad \text{or} \quad D't' = D t$$

giving

$$t' = \frac{Dt}{D'} = \frac{\left(4.90 \times 10^{-11}\right)(2 \text{ hrs})}{\left(9.68 \times 10^{-12}\right)} = 10.1 \text{ hrs}$$

2.2 Fick’s Second Law and Diffusion from a Limited Supply

Fick’s second law in eqn. (4) is a general description of diffusion under non–steady state conditions. We will consider the case of diffusion involving a constant number of diffusing species of atoms all initially concentrated in a very thin surface region of the crystal as depicted in Figure 2-2. This type of diffusion is often called tracer diffusion because we can follow the time evolution of the concentration profile of diffusing atoms if the diffusing species are radioactive.

Diffusion with a limited supply of diffusing species. The total number of diffusing particles is constant.

The boundary conditions in this case are such that the total number of diffusing atoms per unit area remains constant and equal to $N_o$. Suppose that initially, at $t = 0$, all the diffusing atoms are contained within a very small thickness $x_o$. Then the concentration at the surface is, $C_s = N_o/x_o$. Assuming that the crystal is very long then the concentration profile at any point at any time obeys a Gaussian function of the form...
\[
C(x,t) = \frac{N_o}{\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) ; \quad x > x_o \quad \text{Gaussian diffusion from surface} \quad (8)
\]

The concentration profiles \( C(x,t) \) vs. \( x \) at times \( t = t_1 \) and \( t = t_2 \) in Figure 2-2 follow eqn. (8). If we were to substitute eqn. (8) into Fick's second law we would find that it satisfies it; it is a solution of eqn. (4). We generically call this type of diffusion \textbf{Gaussian diffusion}. It arises whenever there is a diffusional spread of a constant number of particles which are initially contained in a thin sheet. The number of diffusing species of particles remains constant. The area under the \( C(x,t) \) vs. \( x \) profiles in Figure 2-2 therefore represent the number of diffusing particles per unit area or \( N_o \). For an infinitely long sample the concentration profile across the specimen after infinite time becomes very small or \( C \to 0 \) as \( t \to \infty \), though the area under the curve is always \( N_o \). In practice, the sample length is finite and the eventual concentration after infinite time is not zero but \( C_o \) where \( C_o L = C_o x_o \) where \( L \) is the sample length (why?).

An interesting feature of Gaussian diffusion is that if we were to follow the time evolution of the concentration at one particular point in the sample such as that at \( x_1 \) in Figure 2-2 then the concentration \( C(x_1,t) \) first increases with time, reaches a maximum and then decreases towards its final value, \( C_o \). This is qualitatively apparent for \( C(x_1,t) \) for \( t = t_1 \), \( t = t_2 \) and \( t = \infty \) in Figure 2-2. Thus the time evolution of the concentration at a given point goes through a peak as depicted in Figure 2-3(a). Furthermore, the particle flux \( \Gamma \) at point \( x_1 \) is proportional to the concentration gradient at that time. It can be seen from Figure 2-2 that initially, just after \( t = 0 \), the concentration gradient is zero at \( x_1 \), but then it increases to finite values as shown at times \( t = t_1 \) and \( t = t_2 \) and becomes zero again as \( t \to \infty \).

\[ C(x_1,t) \quad (a) \quad \Gamma(x_1,t) \quad (b) \]

Concentration, \( C(x_1,t) \) and partical flux \( \Gamma(x_1,t) \) at position \( x_1 \) as a function of time.

\textbf{Figure 2-3}

The concentration gradient therefore also goes through a maximum with time. The behavior of the particle flux with time is illustrated in Figure 2-3(b). For an observer at \( x_1 \), there is as if it were a "passing particle wind" (or a" Gaussian wind") which comes and goes. The maximum particle flux occurs at a time earlier than that for maximum concentration, for example if the flux is maximum after 1 hr then the concentration reaches its peak about 3 hours. If \( t_c \) and \( t_f \) are the times when the concentration and flux reach their respective maxima at a given point as identified in Figure 2-3(a) and (b) then \( t_c/t_f = 3 \). The proof of this given in the Worked Example in section 5. If we were to introduce a fixed number of particles into the middle of the sample \( (x = 0) \) then diffusion will occur towards both ends of the sample as depicted in Figure 2-4. For a very long sample the concentration \( C(x,t) \) at point \( x \) will then be half of that predicted by eqn. (8) because the particles are now just as likely to diffuse left or right. Thus, in this case,

\[
C(x,t) = \frac{N_o}{2\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad \text{Gaussian diffusion from a central sheet of particles} \quad (9)
\]
Example 2–2: Gaussian diffusion

The solubility limit of boron atoms on the surface of the silicon crystal is of the order of \(10^{20} \text{ cm}^{-3}\) around 1000 °C. Suppose that at \(t = 0\) at 1100 °C the surface of a Si crystal has been suddenly saturated with B atoms and then the boron supply is turned off. The boron atoms with time will diffuse into the bulk of crystal. If we assume initially all the B atoms were within a lattice dimension \(a\) (0.543 nm) what will be concentration of B at 1 µm (micron) from the surface after 1 hour at 1100 °C? How many B atoms are there in parts per billion (ppb)? What would have been the concentration had the temperature been 800 °C?

The diffusion coefficient of B in Si has \(D_o = 0.76 \text{ cm}^2 \text{ s}^{-1}\), and \(E_A = 3.46 \text{ eV/atom}.\) Remember that \(D\) is thermally activated: \(D = D_o \exp(-E_A / kT)\).

Solution:

We are given the surface concentration of B as \(C_s = 1 \times 10^{20} \text{ cm}^{-3}\). Within a lattice parameter \(a\) there are \(N_o\) number of B atoms per cm\(^2\), where

\[
N_o = a \cdot C_s = (0.543 \times 10^{-7} \text{ cm})(1 \times 10^{20} \text{ cm}^{-3}) = 5.43 \times 10^{12} \text{ cm}^{-2}.
\]

The diffusion coefficient \(D\) at 1100° C or 1373° K is

\[
D = D_o \exp\left(-\frac{E_A}{kT}\right) = 0.76 \exp\left(-\frac{3.46 \times 1.6 \times 10^{-19}}{1.3807 \times 10^{-23}(1373)}\right) = 1.515 \times 10^{-13} \text{ cm}^2 \text{ s}^{-1}.
\]

Distance of interest \(x = 1 \mu\text{m}\) or \(10^{-4}\) cm and time of interest \(t = 1 \times 60 \times 60\) seconds. Then,

\[
\frac{x^2}{4DT} = \frac{(1.0 \times 10^{-1})^2}{4(1.515 \times 10^{-13})(1 \times 60 \times 60)} = 4.583
\]

and the concentration at \(x = 1 \mu\text{m}\) after 1 hr can be found from

\[
C(x,t) = \frac{N_o}{\sqrt{\pi DT}} \exp\left(-\frac{x^2}{4Dt}\right)
\]

\[
C(1 \mu\text{m}, 1\text{hr}) = \frac{5.43 \times 10^{12}}{\sqrt{\pi(1.515 \times 10^{-13})(1 \times 60 \times 60)}} \exp(-4.583) = 1.34 \times 10^{15} \text{ cm}^{-3}.
\]
The concentration of Si atoms in the crystal is \(5 \times 10^{22} \text{ cm}^{-3}\) so that the relative B concentration is \((1.34 \times 10^{15} \text{ cm}^{-3})/(5 \times 10^{22} \text{ cm}^{-3}) = 2.68 \times 10^{-8}\) or 26.8 ppb (parts per billion). For diffusion at 900° C, the new diffusion coefficient is \(D' = 1.035 \times 10^{-15} \text{ cm}^2 \text{s}^{-1}\). Then

\[
\frac{x^2}{4D'T} = 670.8 \text{ which means}
\]

\[
C(1 \mu\text{m}, 1\text{hr}) = \frac{5.43 \times 10^{12}}{\sqrt{\pi(1.035 \times 10^{-15})}(1 \times 60 \times 60)} \exp(-670.8) = 7.6 \times 10^{-274}
\]

It is apparent that when the temperature is decreased by 15%, the diffusion becomes almost totally insignificant over this time scale. Indeed, for the concentration at \(x = 1 \mu\text{m}\) to reach \(1.34 \times 10^{15} \text{ cm}^{-3}\) at 900°C we have to hold the crystal at this temperature for 150 hours (or 6 days). This can be shown quite simply by finding the new time \(t'\) which makes the concentration at \(x = 1 \mu\text{m}\) at 900°C equal to \(1.34 \times 10^{15} \text{ cm}^{-3}\), i.e.

\[
C(x, t') = \frac{N_o}{\sqrt{\pi D't'}} \exp\left(-\frac{x^2}{4D't'}\right) = 1.34 \times 10^{15} \text{ cm}^{-3}.
\]

The above expression can be solved by plotting \(C(x, t')\) against \(t'\) and finding that value of \(t'\) where \(C(x, t') = 1.34 \times 10^{15} \text{ cm}^{-3}\). This is left as an exercise for the student. The result is \(t' \approx 150\) hrs. or about 6 days. At lower temperatures this time scale becomes even longer going into years, decades and centuries. Indeed at 600 °C, the time scale is something like 20 centuries! Clearly the importance of temperature cannot be overemphasized.

### 3 Silicon Dioxide (SiO₂)

The major reason for the use of Si in integrated circuit fabrication is that its oxide, SiO₂, has a number of technologically important properties which are not exhibited to the same extent by the other semiconductors such as Ge and GaAs.

#### 3.1 The Oxide Growth

Semiconductor devices are fabricated by doping the silicon crystal by the diffusion of dopant atoms into the crystal. Diffusion process must be carried out selectively. Certain regions of the Si crystal surface must be masked. The success of the IC technology is due to the fact that silicon dioxide, SiO₂, grown on the surface of a Si crystal is an excellent barrier against diffusion. It can be easily etched to open windows to allow the diffusion process through these selective areas.

First a layer of SiO₂ is grown on the surface of the Si crystal as depicted in Figure 3-1(a). Then windows are etched in the SiO₂ by using an etchant. The etchant removes the oxide but does not attack the Si crystal surface so that over etching does not damage the crystal surface. This is an important technological advantage. Figure 3-1(b) shows an etched window in the SiO₂ layer to allow the diffusion of dopant atoms into the Si crystal. When the crystal is exposed to a gas of the dopant atoms, say B (boron) atoms, then B atoms diffuse into the Si crystal through these windows. As B atoms diffuse they dope the crystal \(p\)-type.

The SiO₂ layer adheres to the Si crystal without any cracks and pores and forms a continuous layer with an amorphous structure. In addition to acting as a diffusion mask, it also passivates the crystal surface. Without the passivation, the surface will have sufficient conductance to leak a current over the surface in device applications. Further, SiO₂ is a good insulator and serves as a dielectric layer in Metal Oxide Semiconductor (MOS) devices. The oxidation process can occur in various ambients. **Dry oxidation** involves the oxidation of Si in a pure oxygen gas atmosphere:
Si (solid) + O₂ (gas) → SiO₂ (solid)

Dry oxidation

An e-Booklet

The growth of SiO₂ on the Si crystal surface is an essential part of the IC fabrication process. It is readily grown as a thin layer on the crystal surface by thermal oxidation as shown in (a). The oxide passivates the crystal surface and acts as a mask against dopant atoms diffusing into the crystal during the diffusion stage for doping the crystal. It can be easily etched without harming the Si crystal surface to open a window through which the dopants can be diffused into the crystal as shown in (b).

Figure 3-1

Wet oxidation involves the oxidation of Si in an atmosphere containing water vapor (steam). Wet oxidation is faster than dry oxidation. An important process design factor is the control of the oxide thickness. Wet oxidation is expressed by

Si (Solids) + 2H₂O (vapor) → SiO₂ (solids) + 2H₂ (gas)  Wet oxidation

Thermal oxidation of Si wafers is carried out typically in an oxidation (or diffusion) tube which as shown schematically in Figure 3-2. Silicon wafers are placed in a quartz susceptor or boat which is then loaded into the quartz oxidation tube. The tube is heated in an electrical furnace to the desired temperature (typically 900 – 1250 °C) which is maintained constant and uniform within the tube during oxidation. The temperature and time for the oxidation process as well as the partial pressure of the oxidant determine the final oxide thickness. For rapid oxide growth at lower temperatures it is desirable to use oxidants at high pressures (e.g. up to 25 atm.). For example high pressure O₂ gas increases the concentration of O₂ molecules available for the oxidation process and hence increases the rate of the reaction in eqn. (1).

Figure 3-2
3.2 Growth Kinetics

As an example we consider dry oxidation but the concepts apply equally to wet oxidation. The amorphous SiO₂ has a relatively open structure so that the volume change on oxidation is large. For every 0.45 µm of Si consumed, 1 µm of oxide is formed. In other words, we remove 0.45 µm of Si and replace it with 1 µm of SiO₂. This means that the actual Si–SiO₂ interface shifts into the Si crystal as the oxidation proceeds as depicted in Figure 3-3. As the oxide gets thicker, the oxidant has to diffuse through the oxide layer to reach the Si crystal surface (Si–SiO₂ interface) where the oxidation process takes place. This picture of the oxidation process that involves the diffusion of the oxidant through the oxide layer to the interface has been shown to be true by radioactive tracer diffusion experiments. The oxidation kinetics can be understood in terms of two stages: **Initial linear growth, Parabolic growth, and Overall growth kinetics.**

![Figure 3-3](image-url)

The oxidation reaction occurs at the Si-SiO₂ interface. The oxidant first has to diffuse through the oxide layer to the interface.

**A. Initial Linear Growth**

Initially either the oxide layer is absent or it is very thin, as depicted in Figure 3-3, so that the oxidant molecules arrive very rapidly at the Si crystal surface. When this is the case, the reaction rate is not limited by the supply of the oxidant molecules but rather by the rate at which they are consumed in the oxidation reaction in eqn. (1). The concentration of oxidant molecules \( C_o \) on the surface depends on the partial pressure of O₂. From the ideal gas law

\[
P V = \left( \frac{N}{N_A} \right) R T \quad \text{(Ideal gas law)}
\]

where \( P \) is the partial pressure of the oxidant gas, \( V \) is the volume, \( N \) is the number of oxidant molecules, \( N_A \) is the Avogadro’s number and \( R \) is the gas constant. Thus,

\[
C_o = \frac{N}{V} = \frac{PN_A}{RT} \quad \text{(Concentration of oxidant molecules)}
\]

At a given temperature, \( C_o \) depends on the partial pressure \( P \) of the oxidant gas only. The rate of oxide growth is \( dx/dt \) which is as shown in Figure 3-3. The surface area stays the same so that \( dx/dt \) is also the rate of SiO₂ volume growth. The rate of SiO₂ volume growth is given by the rate of reaction in eqn. (1) which is proportional the concentration of the oxidant, \( C_o \), available (the concentration of Si atoms is fixed on the Si crystal surface)\(^1\)

---

\(^1\)Mass Action Law in first year chemistry states that the rate of the reaction \( A+B\rightarrow C \) is determined by \( d[C]/dt = k[A][B] \) where \([A]\) is the concentration of atoms of species \( A \) and so on, and \( k \) is a constant. Typically \( k \) exhibits an Arrhenius temperature dependence as the reaction is generally thermally activated.
\[
\frac{dx}{dt} = k_1 C_o
\]

Rate of volume growth (4)

where \(k_1\) is a constant. As the reaction involves breaking bonds to form SiO\(_2\) it is clear that the reaction in eqn. (4) must be thermally activated which means that \(k_1\) has an Arrhenius temperature dependence. Thus \(k_1 = k_{10} \exp(E_1 / kT)\) where \(E_1\) is the activation energy for the reaction in eqn. (1). When we integrate eqn. (4) it is apparent that \(x\) depends linearly on \(t\) as \(k_1\) and \(C_o\) are independent of \(t\). Thus we can integrate eqn. (4) to obtain,

\[
x = \frac{B}{A} (t + \tau)
\]

Linear growth equation (5)

where \(B/A\) is a constant\(^2\) which is simply \(k_1 C_o \tau\), and \(\tau\) is a constant that depends on the initial conditions on the Si crystal surface. For example, if the initial oxide thickness is \(x_i\) then obviously \(\tau = (Ax_i) / B\). The constant \(B/A\) is called the linear rate constant and is thermally activated by virtue of \(k_1\) being thermally activated.

\[
\frac{B}{A} = K_1 \exp\left(-\frac{E_1}{kT}\right)
\]

Linear rate constant (6)

where \(K_1\) is a constant related to \(k_1\). The value of the activation energy, \(E_1\), is about 2 eV/atom which is comparable to the energy, 1.8 eV/atom, required to break a Si–Si bond. After all, a Si–Si bond has to be broken to form the oxide, SiO\(_2\).

**B. Parabolic Growth**

When, after some time, the oxide becomes thicker the oxidant molecules have to diffuse through the oxide layer to reach the Si–SiO\(_2\) interface for the oxidation reaction in eqn. (1) to occur. The supply of the oxidant molecules is now limited by the rate of their diffusion through the oxide. As the oxide becomes thicker, it takes longer for the oxidant molecules to diffuse through the oxide and the rate of arrival at the interface becomes smaller. Thus, the rate of growth \(dx/dt\) is limited by the flux, \(\Gamma\), of O\(_2\) molecules. Using Fick's first law,

\[
\frac{dx}{dt} = k_2 \Gamma (\text{O}_2 \text{molecules}) = k_2 D \frac{C_o - C_I}{x}
\]

Fick's law and the rate of growth (7)

where \(k_2\) is a proportionality constant (temperature independent), \(D\) is the diffusion coefficient of the oxidant (O\(_2\)) in SiO\(_2\), and \(C_o\) and \(C_I\) are the concentrations of the oxidant at the surface and at the interface respectively. \(C_o\) is determined by the partial pressure of O\(_2\) on the surface and \(C_I\) is determined by the concentration of O\(_2\) in the oxide (number of oxygen molecules in SiO\(_2\) per unit volume) since just at the interface we have the oxide forming. We can view the oxidation at the interface as the immediate removal and incorporation of the arriving O\(_2\) molecules into the formation of an oxide so that the O\(_2\) concentration here, \(C_o\), cannot exceed that in the oxide. To obtain the thickness, \(x\), at time \(t\) we have to integrate eqn. (7).

To simplify the result we can assume that we are interested in the thickness after a long time so that

\[
\int_0^t \int_0^x x dx = k_2 D (C_o - C_I) \int_0^t dt
\]

After integration, this expression will produce the parabolic growth equation which is expressed as

\[
x^2 = Bt
\]

Parabolic growth equation (8)

Equation (8) actually applies after a long time, \(t \gg A^2/4B\). \(B\) in eqn. (8) is a constant that incorporates \(k_2 D (C_o - C_I)\). It depends on the pressure by virtue of \(C_o\) (eqn. (3)). Increasing the pressure increases \(C_o\) and thus \(B\) and hence the growth rate. It is thermally activated because the diffusion process, represented by

\(^2\)The reason for writing the constant as \(B/A\) instead of simply as a single constant, say \(K\), becomes apparent later where \(B\) and \(A\) appear in the overall growth equation.
D, is thermally activated. The diffusion of O₂ through the oxide involves bond breaking and overcoming a potential energy barrier at each jump. Thus,

\[ B = K_2 \exp\left(-\frac{E_2}{kT}\right) \]  

(9)

The activation energy \( E_2 \) depends on the oxidant species diffusing through the oxide. For O₂, \( E_2 \) is 1.24 eV/atom and for H₂O it is 0.79 eV/atom.

C. Overall Growth Kinetics

In the above treatment we considered two limiting cases. Initially, the growth is linear and later it is parabolic. It is not difficult to derive the general expression for the growth kinetics by considering the effects of oxidant flux by diffusion through the oxide and the reaction kinetics at the interface at the same time. The mathematics is beyond the scope of this book but the result is perfectly understandable as it predicts both the initial linear and final parabolic rates:

\[ x = \frac{A}{2} \left[ \left(1 + \frac{t + \tau}{A^2/4B}\right)^{1/2} - 1 \right] \]  

(10)

Growth kinetics

In eqn. (10) \( A \) and \( B \) are as defined in eqns. (6) and (9) and the constant \( \tau \) is related to the initial oxide thickness \( x_i \) and is given by

\[ \tau = \frac{x_i^2 + Ax_i}{B} \]  

(11)

It is left to the student to show that for short times, when \( t + \tau \ll A^2/4B \), eqn. (10) reduces to the linear growth rate in eqn. (5) and for long times, when \( t + \tau \gg A^2/4B \), it leads to the parabolic growth equation (8). Table 3–1 shows the oxide–thickness growth as a function of time for wet and dry oxidation on the (100) surface which is the surface used the fabrication of MOS devices. The effects of temperature and pressure are also shown. Both increase the growth rate. The thickness vs. time behavior closely follows eqn. (10).

![Figure 3-4](image-url)
Table 3-1 provides the values for the various constants for calculating $A$, $B$ and $\tau$ in the oxide thickness expression in eqn. (10). Using these values, one can readily calculate and predict the thickness given the pressure, temperature and time.

**Table 3–1**
Numerical values of various parameters in SiO$_2$ oxidation kinetics on the (100) surface

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Dry (O$_2$)</th>
<th>Wet (H$_2$O)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_1$</td>
<td>$\mu$m hr$^{-1}$</td>
<td>$3.708 \times 10^6$</td>
<td>$9.72 \times 10^7$</td>
<td>B/A</td>
</tr>
<tr>
<td>$E_1$</td>
<td>eV atom$^{-1}$</td>
<td>2.0</td>
<td>2.05</td>
<td>B/A</td>
</tr>
<tr>
<td>$K_2$</td>
<td>$\mu$m$^2$ hr$^{-1}$</td>
<td>770.4</td>
<td>385.2</td>
<td>B</td>
</tr>
<tr>
<td>$E_2$</td>
<td>eV atom$^{-1}$</td>
<td>1.23</td>
<td>0.79</td>
<td>B</td>
</tr>
<tr>
<td>$x_i$</td>
<td>$\mu$m</td>
<td>0.02</td>
<td>0</td>
<td>$\tau$</td>
</tr>
</tbody>
</table>

**Example 3–1: Oxidation time at 900°C for a 0.1 $\mu$m thick oxide**
A SiO$_2$ of thickness 0.1 $\mu$m is required to be grown on the (100) surface of a Si crystal by thermal oxidation at 900 °C. What are the oxidation times for wet and dry oxidation?

**Solution:**

a *Wet oxidation*: We can calculate $A$ and $B$ using the Arrhenius constants given in Table 3-1.

\[
B = K_2 \exp\left(\frac{-E_2}{kT}\right) = 385.2 \exp\left[-\frac{(0.79 \text{ eV})(1.6 \times 10^{-19} \text{ C})}{(1.38 \times 10^{-23} \text{ J K}^{-1})(900 \text{ K} + 273 \text{ K})}\right] = 0.1554 \mu$m hr$^{-1}$.
\]

\[
B \frac{A}{K_1} = K_1 \exp\left(\frac{-E_1}{kT}\right) = 9.72 \times 10^7 \exp\left[-\frac{(2.05 \text{ eV})(1.6 \times 10^{-19} \text{ C})}{(1.38 \times 10^{-23} \text{ J K}^{-1})(900 \text{ K} + 273 \text{ K})}\right] = 0.1513 \mu$m hr$^{-1}$.
\]

Thus A = 1.027 $\mu$m. For wet oxidation $x_i = 0$ so that $t = 0$. Rearranging eqn. (10) for $x$ in terms $t$, we obtain

\[
t = \frac{1}{B} x^2 + \frac{A}{B} x - t\tau
\]

Substituting $x = 0.1 \mu$m and calculating $t$ we find, $t = 0.725$ hr in agreement with that from the oxide thickness curve in Figure 3-4 at 900 °C.

b *Dry oxidation*: We can again calculate $A$ and $B$ using the Arrhenius constants given in Table 3-1 for dry oxidation.

\[
B = K_2 \exp\left(\frac{-E_2}{kT}\right) = 770.4 \exp\left[-\frac{(1.23 \text{ eV})(1.6 \times 10^{-19} \text{ C})}{(1.38 \times 10^{-23} \text{ J K}^{-1})(900 \text{ K} + 273 \text{ K})}\right] = 0.004 \mu$m hr$^{-1}$.
\]

\[
B \frac{A}{K_1} = K_1 \exp\left(\frac{-E_1}{kT}\right) = 3.708 \times 10^6 \exp\left[-\frac{(2.0 \text{ eV})(1.6 \times 10^{-19} \text{ C})}{(1.38 \times 10^{-23} \text{ J K}^{-1})(900 \text{ K} + 273 \text{ K})}\right] = 0.00947 \mu$m hr$^{-1}$.
\]

Thus $A = 0.423 \mu$m. For dry oxidation $x_i = 0.02 \mu$m so that

\[
\tau = \frac{x_i^2 + Ax_i}{B} = \frac{(0.02)^2 + (0.423)(0.02)}{0.004} = 2.213 \text{ hrs.}
\]
Rearranging eqn. (10) for \( x \) in terms \( t \),
\[
 t = \frac{1}{B} x^2 + \frac{A}{B} x - \tau
\]
Substituting \( x = 0.1 \) \( \mu \)m, \( t = 2.213 \) hr. and calculating \( t \) we find, \( t = 10.8 \) hrs. close to that from the oxide thickness curve in Figure 3-4 at 900 °C.

### ADVANCED TOPICS

#### 4 Derivation of Fick's First and Second Laws

Suppose that the concentration of a given species of particles at some time \( t \) in a tank of gas decreases in the \( x \)-direction and has the profile \( C(x, t) \) shown in Figure 4-1.

![Figure 4-1](image)

Net diffusion (flux) of particles from higher to lower concentrations. There are more particles crossing \( x_o \) coming from left, \( (x_o - \ell) \), than coming from right, \( (x_o + \ell) \).

We know that particles are moving around randomly due to their thermal energy. We will assume for simplicity that the we have unit cross sectional area and we will consider the one dimensional example case depicted in Figure 4-1. Collisions of the particles with each other randomize particle motions. On average, a particle moves a distance \( \ell \) in the \( x \)-direction between collisions which is called the mean free path. Suppose that \( \tau \) is the mean free time between the collisions which randomize the particle motions.

A particle moves a mean distance \( \ell \) in the \( +x \) or \( -x \) direction and then it is scattered and changes direction. Its mean speed along \( x \) is \( v_x = 1/\tau \). Let us evaluate the flow of particles in the \( +x \) and \( -x \) directions through the plane at \( x_o \) and hence find the net flow in the \( +x \) direction. We can divide the \( x \)-axis into hypothetical segments of length \( \ell \) so that each segment corresponds to a mean free path. Going across a segment, a particle suffers one scattering event.

Consider what happens during one mean free time, time it takes for a particle to move across a segment towards left or right. Half of the particles in \( (x_o - \ell/2) \) would be moving toward \( x_o \) and the other half away from \( x_o \) and in time \( \tau \) half of them will reach \( x_o \) and cross. The number of particles moving towards the right to cross \( x_o \) is \( \frac{1}{2} C_1 \ell \) where \( C_1 \) is the concentration at \( (x_o - \ell/2) \) and \( \ell \) is the volume of the segment (cross sectional area is unity). Similarly half of the particles in \( (x_o + \ell/2) \) would be moving towards left and in time \( \tau \) would reach \( x_o \). Their number is \( \frac{1}{2} C_2 \ell \) where \( C_2 \) is the concentration at \( (x_o + \ell/2) \). The net number of particles crossing \( x_o \) per unit time in \( +x \) direction is the electron flux,

\[
\Gamma = \frac{\frac{1}{2} C_1 \ell - \frac{1}{2} C_2 \ell}{\tau}
\]
Diffusion and Oxidation (© S.O. Kasap, 1990 - 2001: v.1.1)
An e-Booklet

\[ \Gamma = \frac{1}{2\tau}(C_1 - C_2) \]  (1)

As far as calculus of variations is concerned, the mean free path \( \lambda \) is small so that we can calculate \( C_1 - C_2 \) from the concentration gradient using\(^3\)

\[ (C_2 - C_1) = \left( \frac{dC}{dx} \right) \Delta x = \left( \frac{dC}{dx} \right) \lambda \]

We can now write the flux in eqn. (1) in terms of the concentration gradient as

\[ \Gamma = -\frac{1}{2\tau} \left( \frac{dC}{dx} \right) \]

or

\[ \Gamma = -D \left( \frac{dC}{dx} \right) \]  \hspace{1cm} Fick's first law  \hspace{1cm} (2)

where the quantity \( (\lambda^2/2\tau) \) has been defined as the diffusion coefficient of the particles and denoted by \( D \). Thus, the net particle flux \( \Gamma \) at a position \( x \) is proportional to the concentration gradient and the diffusion coefficient. The steeper is this gradient, the larger is the flux \( \Gamma \). In fact, we can view the concentration gradient \( dC/\lambda \) as the driving force for the diffusion flux, just like the electric field, \( -dV/\lambda \), is the driving force for the electric current; \( J = \sigma E = -\sigma dV/\lambda \).

Equation (2) is called **Fick's First Law** and represents the relationship between the net particle flux and the driving force which is the concentration gradient. The diffusion coefficient, \( D = \lambda^2/2\tau = \lambda v_x/2 \), is a measure of the ease with which the diffusing particles move in the medium because it is proportional to the mean free path between collisions and the mean speed \( v_x \) along \( x \). Equation (2) is useful for calculating fluxes from concentration gradients but does not allow the concentration \( C(x, t) \) at a point \( x \) at time \( t \) to be determined. This requires Fick's second law. Consider an elemental volume \( \lambda \) as shown in Figure 4–2. As the width \( \lambda \) is small we can assume that the concentration in this elemental volume is roughly \( C(x, t) \).

![Diagram of diffusion flux](image)

Consider a small elemental volume of width \( \lambda \). Average concentration in this volume is \( C(x) \) and the incoming and outgoing fluxes are \( \Gamma(x) \) and \( \Gamma(x+\lambda) \).

\hspace{1cm} Figure 4-2

\(^3\)We use a partial derivative because \( n \) may be a function of position and time, \( n = n(x,t) \).
Suppose that the particle fluxes at $x$ and $x+dx$ are $\Gamma_x$ and $\Gamma_{x+dx}$. If the incoming flux is larger than the outgoing flux then there is a net accumulation of particles in the volume. Then this can be expressed as follows

$$\Gamma_x - \Gamma_{x+dx} = \frac{dC}{dt} \delta x = \text{Net rate of accumulation of particles in the volume}$$

We can of course write the gradient of the flux across this small elemental volume as

$$\frac{d\Gamma}{dx} = \frac{\Gamma_{x+dx} - \Gamma_x}{\delta x}$$

and then eliminate $\Gamma_x - \Gamma_{x+dx}$ in previous equation to obtain,

$$\frac{d\Gamma}{dx} = \frac{dC}{dt}$$

Substituting Fick's first law from eqn. (2) for $\Gamma$ we obtain,

$$D \frac{d^2 C}{dx^2} = \frac{dC}{dt}$$

This is Fick’s second law. Of course the equation should have partial derivatives because we are differentiating $C(x,t)$, a function of two variable, with respect to $x$ and with respect to $t$ separately. We should really have written eqn.(3) as

$$D \frac{\partial^2 C(x,t)}{\partial x^2} = \frac{\partial C(x,t)}{\partial t}$$

### 5 Derivation of the Maximum in Gaussian diffusion and concentration: The Gaussian wind.

Consider the Gaussian diffusion of constant number $N_o$ of particles from the surface of a crystal into the bulk. Mathematical proof showed that the particle concentration $C(x,t)$ at a location reaches their maxima at time $t_c$ and $t_f$

$$t_c = \frac{x_1^2}{2D} \quad \text{and} \quad t_f = \frac{x_1^2}{6D}$$

so that $t_c/t_f = 3$, that is, the flux reaches a maximum first and then, three times later, the concentration reaches a maximum.

The concentration $C(x,t)$ at position $x$ and $t$ is given by

$$C(x,t) = \frac{N_o}{\sqrt{\pi D}} t^{-\frac{1}{2}} \exp\left(-\frac{x^2}{4Dt}\right)$$

The flux $\Gamma$ at position $x$ and time $t$ can be found from the concentration gradient because $\Gamma = -D(dC/dx)$. To find the gradient of $C(x,t)$ at a given instant, we differentiate $C(x,t)$ with respect to $x$ but keeping $t$ constant. This is called partial differentiation of $C(x,t)$ and is written as $\partial C/\partial x$,

$$\Gamma(x,t) = -D \frac{\partial C(x,t)}{\partial x} = \left(\frac{2x}{4Dt}\right) \frac{N_o}{\sqrt{\pi D}} t^{-\frac{1}{2}} \exp\left(-\frac{x^2}{4Dt}\right)$$

i.e.

$$\Gamma(x,t) = \frac{N_o}{2\sqrt{\pi D}} \frac{x}{t^{\frac{1}{2}}} \exp\left(-\frac{x^2}{4Dt}\right)$$

To find the maximum in $C(x,t)$ at $x = x_1$, occurring over time at $t = t_c$, we use
\[
\left[ \frac{\partial C(x_1, t)}{\partial t} \right]_{t=t_c} = 0
\]
\[
\left[ \frac{N_o}{\sqrt{\pi}D} \left( -\frac{1}{2} t_c^{-\frac{3}{2}} \right) \right] \exp\left( -\frac{x_1^2}{4Dt_c} \right) + \left[ \frac{N_o}{\sqrt{\pi}D} t_c^{-\frac{1}{2}} \right] \left[ \left( \frac{x_1^2}{4D} t_c^{-2} \right) \exp\left( -\frac{x_1^2}{4Dt_c} \right) \right] = 0
\]

Cancelling out terms and simplifying, we find that,
\[
-\frac{1}{2} t_c^{-\frac{3}{2}} + t_c^{-\frac{1}{2}} \left( \frac{x_1^2}{4D} t_c^{-2} \right) = 0 \quad \text{i.e.} \quad t_c = \frac{x_1^2}{2D}
\]

To find the maximum in \( \Gamma(x, t) \) at \( x = x_1 \), occurring over time at \( t = t_f \), we use
\[
\left[ \frac{\partial \Gamma(x_1, t)}{\partial t} \right]_{t=t_f} = 0
\]
\[
\left( \frac{N_o}{2\sqrt{\pi}D} x_1 \left( -\frac{3}{2} t_f^{-\frac{3}{2}} \right) \right] \left[ \exp\left( -\frac{x_1^2}{4Dt_f} \right) \right] + \left( \frac{N_o}{2\sqrt{\pi}D} x_1 t_f^{-\frac{1}{2}} \right) \left[ \left( \frac{x_1^2}{4D} t_f^{-2} \right) \exp\left( -\frac{x_1^2}{4Dt_f} \right) \right] = 0
\]

Cancelling out terms and simplifying we find that,
\[
t_f = \frac{x_1^2}{6D}
\]

Thus, the ratio \( t_c/t_f = 3 \) and the concentration reaches its peak after the flux. The expressions for \( C(x_1, t) \) and \( \Gamma(x_1, t) \) are schematically shown in Figure 2-3 (it is left as an exercise for the student to verify the sketches).

**Notation**

- \( A \) area; constant in linear growth rate
- \( B \) constant in parabolic growth rate
- \( C(x) \) concentration profile
- \( C_I \) concentration of the oxidant at the interface
- \( C_o \) concentration anywhere in the bulk
- \( C_s \) concentration at the surface
- \( D \) diffusion coefficient
- \( \mathcal{E} \) electric field
- \( E_A \) activation energy
- \( E_1 \) activation energy for the oxidation reaction
- \( E_2 \) activation energy for \( O_2 \) through the oxide.
- \( \text{erf}(z) \) error function of \( z \)
- \( \Gamma \) particle flux
- \( J \) current density
- \( k \) Boltzmann’s constant
- \( K_i \) constant associated with \( B/A \)
- \( K_2 \) proportionality constant at the rate of parabolic growth
- \( l \) mean free path
- \( M_A \) atomic mass
- \( N \) number of particles
- \( N_A \) Avogadro’s number
- \( P \) pressure
- \( PE \) potential energy
- \( Q_{\text{act}} \) activation energy in joules per mole
- \( R \) gas constant
- \( \rho \) density
- \( \sigma \) conductivity
- \( T \) temperature
- \( t \) time; oxidation time
- \( t_c \) time at which concentration reaches its maxima
- \( t_f \) time at which flux reaches its maxima
1. **Bipolar Junction Transistor** The bipolar junction transistor operation is based on the diffusion of charge carriers in the base region of the device. Consider an npn transistor as shown in Figure Q.1 (a). The emitter is heavily doped and has plenty of electrons. When a forward bias voltage is applied across the emitter-base junction as in Figure Q.1 (b), the emitter injects electrons into the base. The injected electron concentration \( n(0) \) at the edge of the base (near the emitter-base junction) is determined by the voltage \( V_{EB} \) and is given by the law of the junction:

\[
\begin{align*}
  n(0) &= n_0 \exp(eV_{EB}/kT)
\end{align*}
\]

where \( n_0 \) is the equilibrium electron concentration in the base, \( k \) is the Boltzmann constant and \( T \) is the temperature (K).

The reverse bias voltage \( V_{CB} \) determines the electron concentration at the edge of the base at the base-emitter junction at \( x = W_B \). With a few volts applied across the base-emitter, \( n(W_B) = 0 \). Under steady-state conditions (dc voltages and currents), the electron concentration profile decreases linearly from \( n(0) \) at \( x = 0 \) to 0 at \( x = W_B \) across the base width \( W_B \). Electrons diffuse across the base to reach the collector. There is no electric field in the base to drift the electrons. They are driven by Fick’s first law (concentration gradient). Use the definitions of the current density \( J \) and particle flux \( \Gamma \), to relate \( J \) to \( \Gamma \) and hence show that the current density due to the diffusion of electrons through the base is given by

\[
J = \frac{eD_n n_0}{W_B} \exp\left(\frac{eV_{EB}}{kT}\right)
\]

These electrons diffuse and reach the base-collector junction. There is an electric field in this junction region that sweeps the electrons into the collector. When the electrons reach the collector they become collected by the battery and constitute the collector current \( I_C \). The transistor action is
the fact that the collector current $I_C$ in the output circuit is controlled by the emitter-base voltage $V_{EB}$ in the input circuit.

The reason for the law of the junction is that electrons in the emitter and base have different potential energies: $PE = \text{Charge} \times \text{Voltage}$. The voltage $V_{EB}$ thus modulates the electron PE across the emitter-base junction and this change in the PE, $eV_{EB}$, controls the amount of injected electrons into the base via the Boltzman factor, $\exp(eV_{EB}/kT)$.

A particular $nnp$ transistor has the following properties:

Cross sectional area = 1 mm$^2$
Base width = 4 µm,
Equilibrium electron concentration in the base, $n_o = 1 \times 10^5$ cm$^{-3}$
Electron diffusion coefficient in Si = 37 cm$^2$ s$^{-1}$.
Emitter-base forward bias, $V_{EB} = 0.6$ V.
Calculate the collector current at room temperature.

2. Semiconductor fabrication Consider a wafer of Si crystal which has a uniform boron (B) doping of $1 \times 10^{17}$ cm$^{-3}$ in the bulk. Suppose that the Si wafer is exposed to a phosphorus (P) gas at 1250 °C for 10 minutes. During the diffusion process, the surface P concentration remains saturates at about $1 \times 10^{21}$ cm$^{-3}$. Where is the junction from the surface? How long does it take to have the same junction depth from the surface if the diffusion temperature is 1100 °C? State the assumptions used in your calculations. The diffusion coefficient of P in Si has $D_o = 3.85$ cm$^2$ s$^{-1}$, and $E_d = 3.66$ eV/atom.

Note: B doping makes Si, $p$-type and P-doping makes it $n$-type. The junction is where B and P concentrations are equal.

(Ans: 2.33 µm; 3.5 hrs)

3. Semiconductor fabrication Consider a wafer of Si crystal which has a uniform boron (B) doping of $1 \times 10^{16}$ cm$^{-3}$ in the bulk. The surface region of the Si crystal has been predoped with P so that the P concentration is approximately $1 \times 10^{21}$ cm$^{-3}$ over a small distance 0.01 µm. The Si wafer is then placed in a furnace at 1100 °C for about 45 minutes. Where is the junction from the surface? How long does it take to have the same junction depth from the surface if the diffusion temperature is 1000 °C? State the assumptions used in your calculations. The diffusion coefficient of P in Si has $D_o = 3.85$ cm$^2$ s$^{-1}$, and $E_d = 3.66$ eV/atom.

Note: B doping makes Si, $p$-type and P-doping makes it $n$-type. The junction is where B and P concentrations are equal.

(Ans: At 1100 °C, $z = 7.97$; $x_{\text{junction}} = 1.1$ µm; at 1000 °C, diffusion time $\approx 8.5$ hours)

4. Nitriding of steels Steels may be strengthened by nitriding, which involves diffusion of nitrogen into the surface at high temperatures. Suppose that the nitriding atmosphere produces 0.1 wt.% N at the surface of a steel initially containing 0.001 wt.%. Determine the nitrogen concentration as weight percentage at a distance 1 mm beneath the surface after 5 hours of exposure at 700 °C.

Note: For N diffusion in BCC Fe, $Q_d = 76.57$ kJ mol$^{-1}$ and $D_o = 4.7 \times 10^{-7}$ m$^2$ s$^{-1}$.

(Ans: 0.039 wt.%)
If the nitriding above is done at 650 °C, what is the time required to achieve the same nitrogen concentration at the same depth?

(Ans: 8.34 hrs = 8h 20m 41s)
Table 1
Error and complementary error function tables (from Matlab).

<table>
<thead>
<tr>
<th>z</th>
<th>erf(z)</th>
<th>erfc(z)</th>
<th>z</th>
<th>erf(z)</th>
<th>erfc(z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0.995322</td>
<td>0.00468</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0113</td>
<td>0.988716</td>
<td>2.25</td>
<td>0.998537</td>
<td>0.00146</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0564</td>
<td>0.943628</td>
<td>2.5</td>
<td>0.999593</td>
<td>0.000407</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1125</td>
<td>0.887537</td>
<td>2.75</td>
<td>0.999899 378</td>
<td>0.000101</td>
</tr>
<tr>
<td>0.25</td>
<td>0.2763</td>
<td>0.723673</td>
<td>3</td>
<td>0.999977909</td>
<td>0.0000221</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5205</td>
<td>0.4795</td>
<td>3.25</td>
<td>0.99999697</td>
<td>0.0000043</td>
</tr>
<tr>
<td>0.75</td>
<td>0.7112</td>
<td>0.288844</td>
<td>3.5</td>
<td>0.999999256901</td>
<td>0.00000743</td>
</tr>
<tr>
<td>1</td>
<td>0.8427</td>
<td>0.157299</td>
<td>3.75</td>
<td>0.99999886272</td>
<td>0.00000114</td>
</tr>
<tr>
<td>1.25</td>
<td>0.9229</td>
<td>0.077099</td>
<td>4</td>
<td>0.999999984582</td>
<td>0.000000154</td>
</tr>
<tr>
<td>1.5</td>
<td>0.9661</td>
<td>0.033894</td>
<td>4.25</td>
<td>0.999999998149</td>
<td>0.0000000185</td>
</tr>
<tr>
<td>1.75</td>
<td>0.98667</td>
<td>0.00468</td>
<td>4.5</td>
<td>0.999999999803</td>
<td>0.0000000197</td>
</tr>
</tbody>
</table>

![Graph of erf(z) and erfc(z) vs. z]

Error and complementary error functions vs. their arguments

Figure 3