

Diffusion

Contents

- Definition
- Mechanism
- Fick' Law
- Constant source diffusion
- Limited Source Diffusion
- Two-steps diffusion
- Junction depth measurement
- Resistivity Measurement
- Diffusion system

Diffusion

- Diffusion is the smoothing out that occurs in virtually all materials.
It is the natural tendency of particles to move from regions of high concentrations towards lower concentration regions.

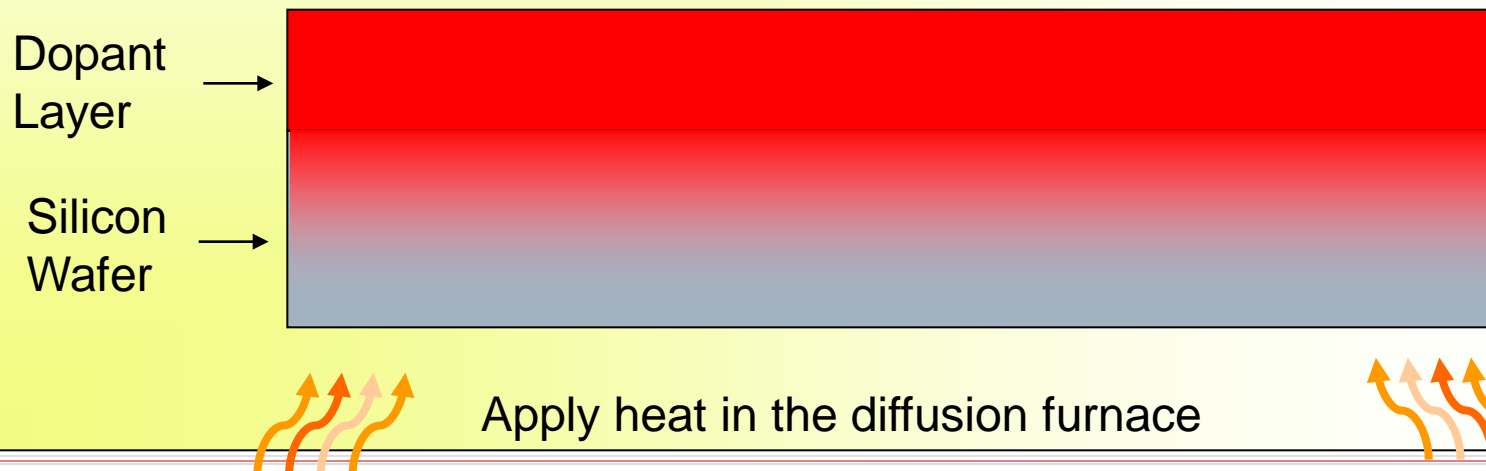
- Examples:
 - Diffusion of perfume in air
 - Tea in hot water
 - Ink in paper
 - Boron and impurity atoms in silicon lattice.

- Typical applications:
 - Form diffused resistors
 - Form sources/drains in MOS devices
 - Form bases/emitters in bipolar transistors

Diffusion process

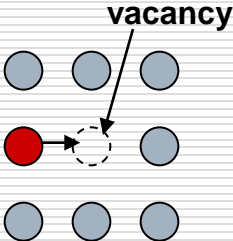

- The diffusion process is done by depositing a shallow– high concentration layer of impurities on top of the silicon wafer then expose the wafer to high temperatures (900–1200 °C)

Dopant layer
diffuses into the
silicon wafer



Two diffusion mechanisms.

- Impurities diffuse through silicon lattice can follow two mechanisms:

| Substitutional diffusion | Interstitial diffusion |
|---|--|
|  <p>Hops from one crystal lattice site to another, and hence substitutes for a lattice atom.</p> <p>Statistically, a number of vacancies exist in the silicon lattice.</p> |  <p>At high temperatures, vacancies can be created by displacing the silicon atoms (left drawing)</p> <p>Some impurities atoms can diffuse through the crystal by jumping from one interstitial site to another</p> |

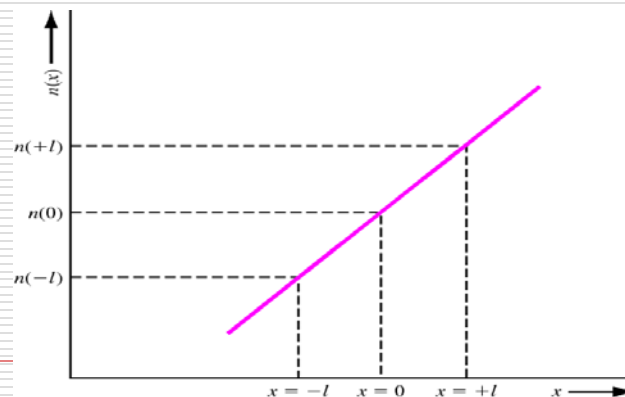
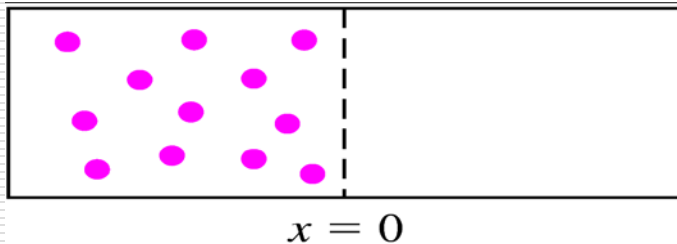
Diffusion

Fick's first law

- If there is a spatial variation of carrier concentration in the semiconductor material, the carriers tend to move from a region of high concentration to a region of low concentration. The current component is called diffusion current.
- Diffusion model:
The basic one-dimensional diffusion process follows Fick's first law of diffusion.

$$J = -D \frac{\partial N(x,t)}{\partial x}, \quad D: \text{diffusion coefficient}$$

Basically, Fick's law states that if particles in some collection ($D \neq 0$), they will move in a way to try to even out the distribution.



• Electron concentration versus distance

Diffusion

Fick's second law

□ Continuity relation

A continuity equation is a differential equation that describes *the conservative transport of some kind of quantity*. Since mass, energy, momentum, and other natural quantities are conserved, a vast variety of physics may be described with continuity equations. The total number of particles inside the cube would change in time due to the difference between the input and output flux.

$$\Delta N \Delta x A = J(x + \Delta x) A \Delta t - J(x) A \Delta t$$

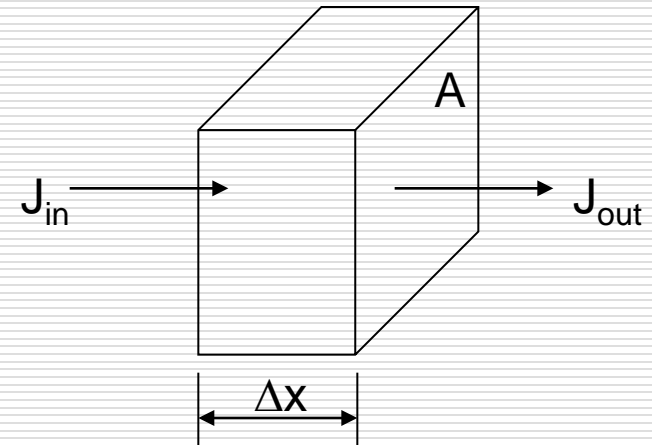
$$\frac{\Delta N}{\Delta t} = \frac{J(x + \Delta x) - J(x)}{\Delta x}$$

when Δx and $\Delta t \rightarrow 0$

$$\frac{\partial N}{\partial t} = - \frac{\partial J}{\partial x}$$

□ That is defined as Fick's second law of diffusion. Combing both laws

$$\frac{\partial N(x,t)}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial N(x,t)}{\partial x} \right)$$



Constant Source Diffusion

Complementary Error Function Profiles

- If The diffusivity, D , does not change with space, then

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

Solving this differential equation requires Boundary conditions.

- Constant-Source diffusion

The impurity concentration is held constant at the surface of the silicon wafer

– The boundary conditions:

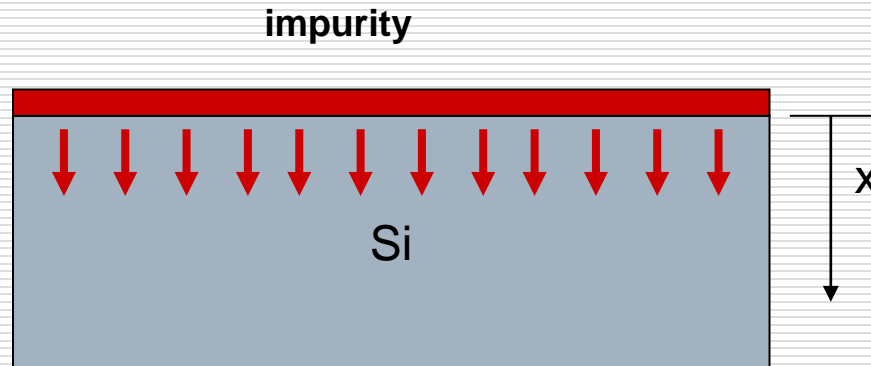
$N(x=0,t) = N_0$; Fixed surface density

$N(x>0,t=0) = 0$; No impurities distribution in silicon before heating

$N(x \rightarrow \infty, t) = 0$; No impurities distribution at the back of the silicon wafer

Constant Source Diffusion

Complementary Error Function Profiles



- The solution of the diffusion equation under these boundary conditions is

$$N(x,t) = N_o \operatorname{erfc}\left(x / \left(2\sqrt{Dt}\right)\right)$$

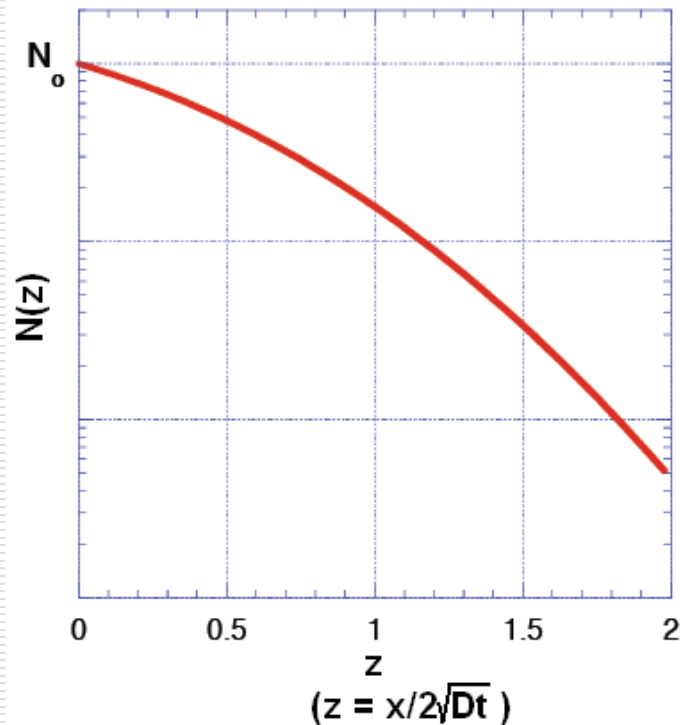
erfc is referred to as the complementary error function

$$\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$$

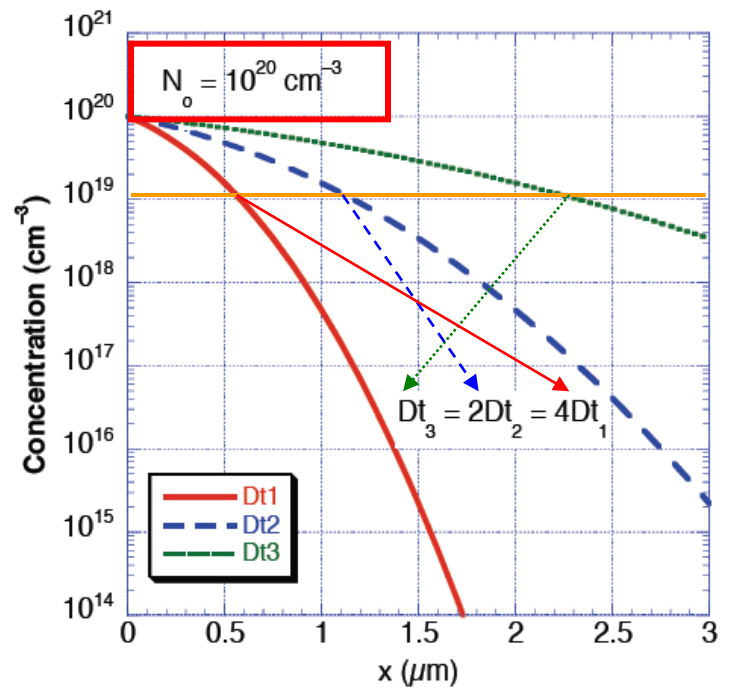
$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du$$

Constant Source Diffusion

Complementary Error Function Profiles



Complementary error function distribution



The effect of changing time and diffusivity on the diffusion penetration depth and rate

Constant Source Diffusion

Complementary Error Function Profiles

- **Dose**: The total amount of dopant (impurities) diffused in the wafer per unit area.
- The dose is found by integrating the diffusion impurities concentration throughout the silicon wafer.

$$Q = \int_0^{\infty} N(x,t) dx = \int_0^{\infty} N_o \operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) dx$$

$$= 2N_o \sqrt{Dt/\pi}$$

- Notice that the dose increases with time and impurities diffusivity.
- More typically, N_o , is determined by the **solid solubility limit** of the impurity in silicon.
 - Solid solubility limit : The amount of dopant that can be held per unit volume of the silicon at a certain temperature.

| Solid-solubility (cm ⁻³) | 900 | 1000 | 1100 | 1200 |
|--------------------------------------|----------------------|----------------------|----------------------|----------------------|
| Boron | 1.2x10 ²⁰ | 2.0x10 ²⁰ | 2.7x10 ²⁰ | 2.7x10 ²⁰ |
| Phosphorous | 7.0x10 ¹⁹ | 1.2x10 ²⁰ | 1.3x10 ²⁰ | 1.5x10 ²⁰ |

– The typical values of the solid solubility limit for boron and phosphorous –

Limited Source Diffusion Gaussian Profiles

- Limited-source diffusion (constant-dose diffusion)
A constant dose of impurity is introduced to the semiconductor from some source, then the source is removed and the diffusion proceeds with fixed amount of impurities.

– The boundary conditions :

$$Q = \int_0^{\infty} N(x,t) dx = \text{constant}$$

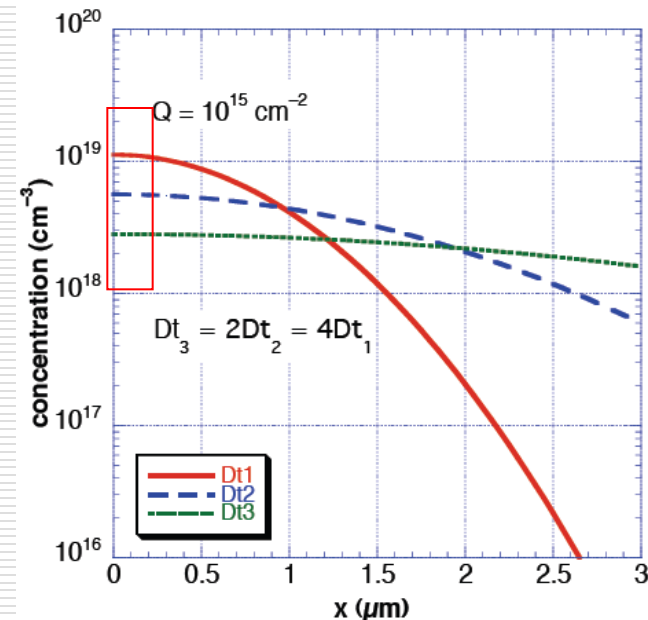
$$N(x > 0, t = 0) = 0, \text{ No impurities in silicon prior}$$

$$N(x \rightarrow \infty, t) = 0, \text{ No impurities at the back of Si}$$

- The solution of the diffusion equation :

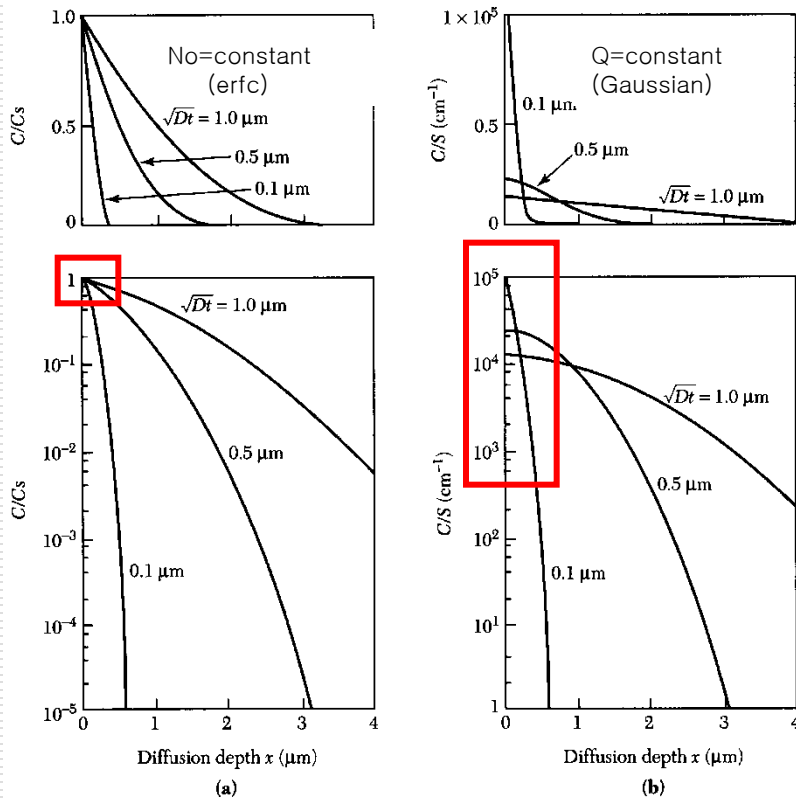
$$N(x,t) = \left(Q / \sqrt{\pi Dt} \right) \exp(-x^2 / 4Dt)$$

The concentration near at the surface is : $N(0,t) = \left(Q / \sqrt{\pi Dt} \right)$



- We notice that the surface concentration of dopant decreases for longer diffusion due to the finite number of impurities

Diffusion Profile Comparison



Complementary Error Function and Gaussian Profiles are Similar in Shape

$$\text{erfc}(z) = 1 - \text{erf}(z)$$

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp[-x^2] dx$$

Constant-source Limited-source

Two-steps diffusion

The predeposition step.

- The goal of the dopant predeposition diffusion is to move dopant atoms from a source to the wafer, and then allow the dopants to diffuse into the wafer.
- At low temperature and short time (small Dt product), the impurities do not diffuse far from the surface. Hence, the impurities are assumed to be just at the surface.

$$Q = 2N_o \sqrt{Dt} / \pi$$

This step is referred to as the predeposition step.

- The source of dopant can be in several forms – solid (boron nitride and phosphorus oxide ceramic discs), liquid (boron tribromide and POCl_3), or gas (diborane or phosphine).
- In order for the dopants to move into the silicon, they must be given energy, usually in the form of heat. In order for the diffusion to occur in a reasonable time, the temperature must be very high ($900^\circ\text{C} < T < 1200^\circ$).
- At this temperature the dopant (in the form of an oxide) reacts with the exposed silicon surface to form a highly doped glass. It is from this glass that the dopants can then diffuse into the wafer.

Two-steps diffusion

The *drive-in* step.

- we can remove the impurities source and follow up with a limited-source diffusion step. After the predeposition diffusion, the dopants are situated close to the surface of the wafer. However, they must diffuse even farther to lower the overall concentration in order for some of the devices to work properly.
- The first diffusion (predeposition) introduces dopants into the wafer.
- The second diffusion (drive-in) redistributes the dopants and allow the dopants to diffuse into the wafer more deeply (up to ~3 micrometers)
- The concentration near the surface after the drive-in step is

$$N(0,t) = \frac{2N_o}{\pi} \sqrt{\frac{(Dt)_{pr}}{(Dt)_{dr}}}, \text{ the suffix } pr \text{ referres to predeposition and } dr \text{ is drive - in}$$

- $Dt_{dr} \gg Dt_{pr}$: The impurity profile is closely represented by Gaussian.
- $Dt_{dr} \ll Dt_{pr}$: The complementary error function (this case is approximated as a constant-source)

Diffusion coefficient

- The rate of diffusion depends on time and the diffusion coefficient.
- The diffusion coefficient D depends exponentially on the temperature as follows

$$D = D_o \exp(-E_A/kT): \text{Arrhenius equation}$$

- Some of the most common dopant elements and their diffusion coefficients are presented below

| Element | Symbol | D_o (cm ² /sec) | E_A (eV) | Comment |
|-------------|--------|------------------------------|------------|-----------------------------------|
| Boron | B | 10.5 | 3.69 | Acceptor (to get p-type material) |
| Aluminum | Al | 8.0 | 3.47 | Acceptor (to get p-type material) |
| Gallium | Ga | 3.6 | 3.51 | Acceptor (to get p-type material) |
| Indium | In | 16.5 | 3.9 | Acceptor (to get p-type material) |
| Phosphorous | P | 10.5 | 3.69 | Donor (to get n-type material) |
| Arsenic | As | 0.32 | 3.56 | Donor (to get n-type material) |
| Antimony | Sb | 5.60 | 3.95 | Donor (to get n-type material) |

Junction formation using diffusion

- The goal of most diffusions is to form a pn junction by converting an p -type material to n -type or vice versa. That means when diffusing a certain impurity, the particles diffuse in a region where opposite type of dopant exist.
- The point at which the diffused impurity profile intersects the background concentration is referred to as **the metallurgical depth, x_j** .

$N(x_j) = N_B$ (The concentration of the background dopant in the wafer).

- For the constant-source diffusion,
$$N(x_j) = N_B = N_o \operatorname{erfc}(x_j / 2\sqrt{Dt}) \rightarrow x_j = 2\sqrt{Dt} \operatorname{erfc}^{-1}(N_B / N_o)$$

We can re - write the above relation as

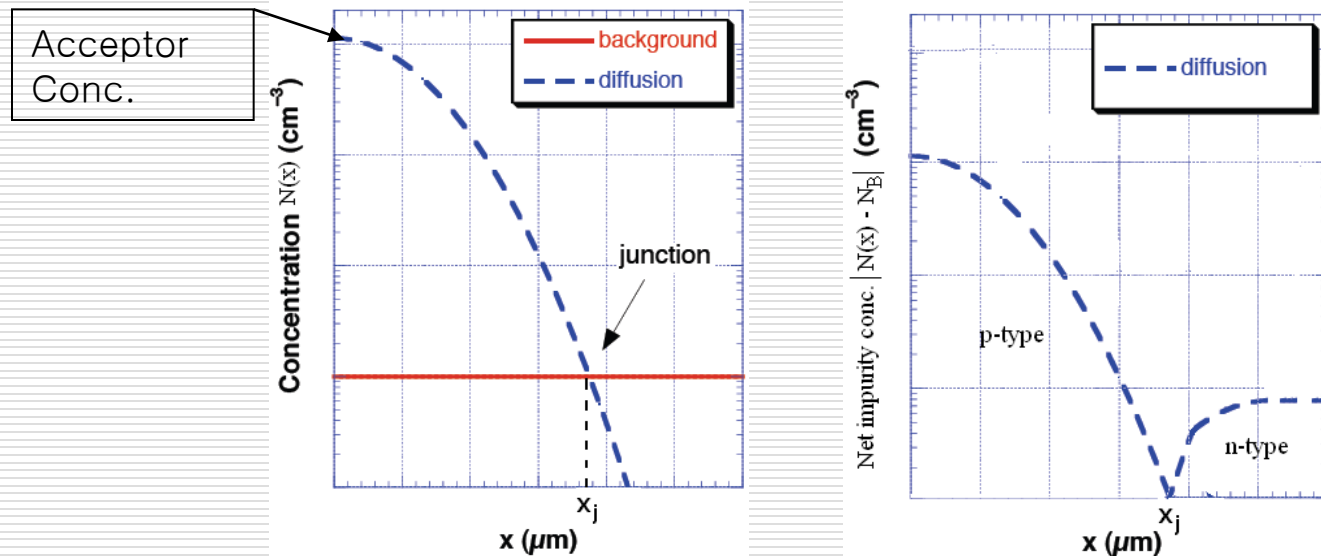
$$x_j = 2\sqrt{Dt} \operatorname{erf}^{-1}(1 - N_B / N_o)$$

- For the limited-source diffusion,
$$N(x_j) = N_B = \frac{Q}{\sqrt{\pi Dt}} \exp(-x^2 / 4Dt) \rightarrow x_j = 2 \left(Dt \cdot \ln \left(\frac{N_B}{N_o} \right) \right)^{1/2}$$

where $N_o = \frac{Q}{2\sqrt{\pi Dt}}$

Junction formation using diffusion

- At the junction, the net impurities is zero.
- If Diffusing Boron, the region before the junction is p -type and after the junction is n -type.



Junction depth measurement

Bevel and stain (groove and stain)

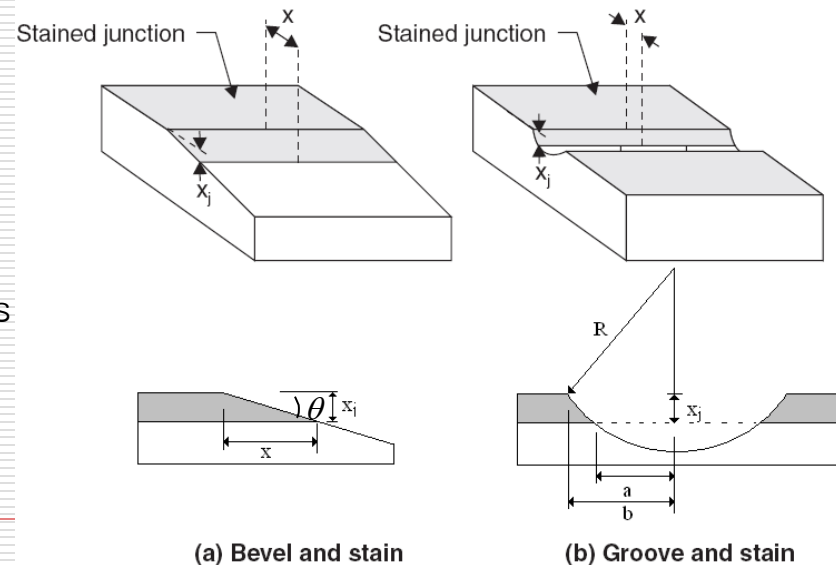
- Bevel and stain, and groove and stain, are easy to implement technique used to determine junction depth.
- The junction is then stained with chemicals (Concentrated HF acid with 0.1 to .5% nitric acid through exposure to high-intensity light).
- The distance x can be measured with an optical microscope.

- For the Bevel and stain, knowing the bevel angle, θ ($1 \sim 5^\circ$), the junction depth can be calculated as

$$x_j = x \tan \theta$$

- For the groove and stain the depth is calculated as

$$x_j = \sqrt{R^2 - b^2} - \sqrt{R^2 - a^2} \approx \frac{a^2 - b^2}{2R}, \quad \text{for } R \gg a$$



Sheet resistance

- R_s (ohms per square) is the sheet resistance. The **sheet resistance** is a measure of resistance of thin films that have a uniform thickness. It is commonly used to characterize materials made by semiconductor doping

- For a rectangular block, the resistance R is

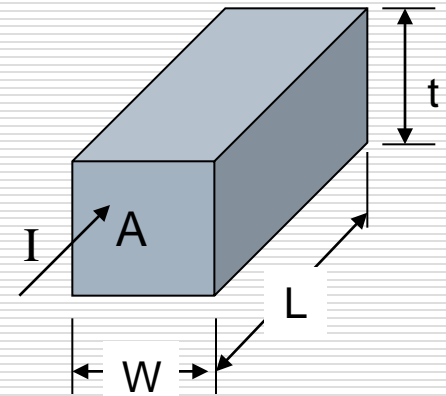
$$R = \rho L / A, \quad \text{where } \rho = 1/\sigma, \text{ and } \sigma \text{ is the conductivity}$$

$$\sigma = q(\mu_n n + \mu_p p), \quad \text{where } q \text{ is the effective electron charge} = 1.602 \times 10^{-19} \text{ C}$$

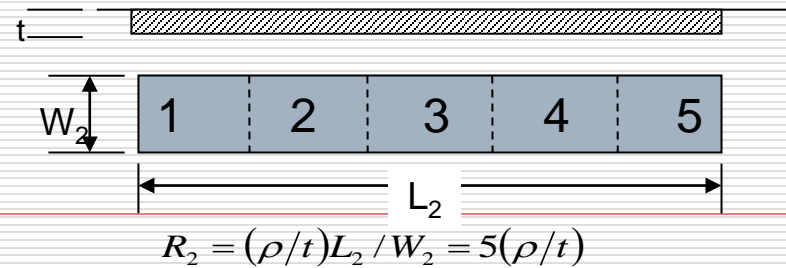
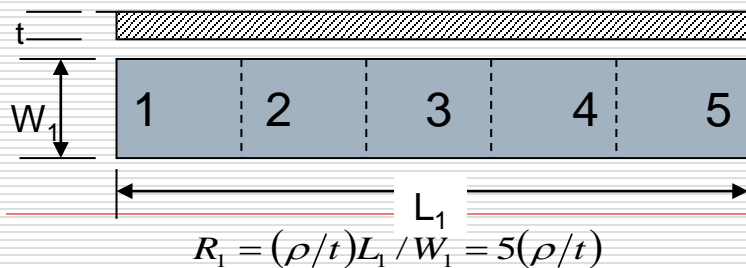
- L and A are the block's length and cross section area.

$$R = \rho L / (Wt) = (\rho/t)(L/W)$$

$$R = R_s (L/W)$$



- For a square, $L = W$. Therefore, $R = R_s$ for any size square



Sheet resistance

- The impurities diffusion is function of depth and hence the resistivity ρ is function of depth as well.
- The sheet resistance is defined in terms of the average ρ .

$$\bar{\rho} = \frac{1}{\frac{1}{x_j} \int_0^{x_j} \sigma(x) dx} \rightarrow R_s = \frac{\bar{\rho}}{x_j} = \left[\int_0^{x_j} \sigma(x) dx \right]^{-1} = \left[q \int_0^{x_j} \mu_n (n(x) + \mu_p p(x)) dx \right]^{-1}$$

- In extrinsic material, where impurities dominate the carriers

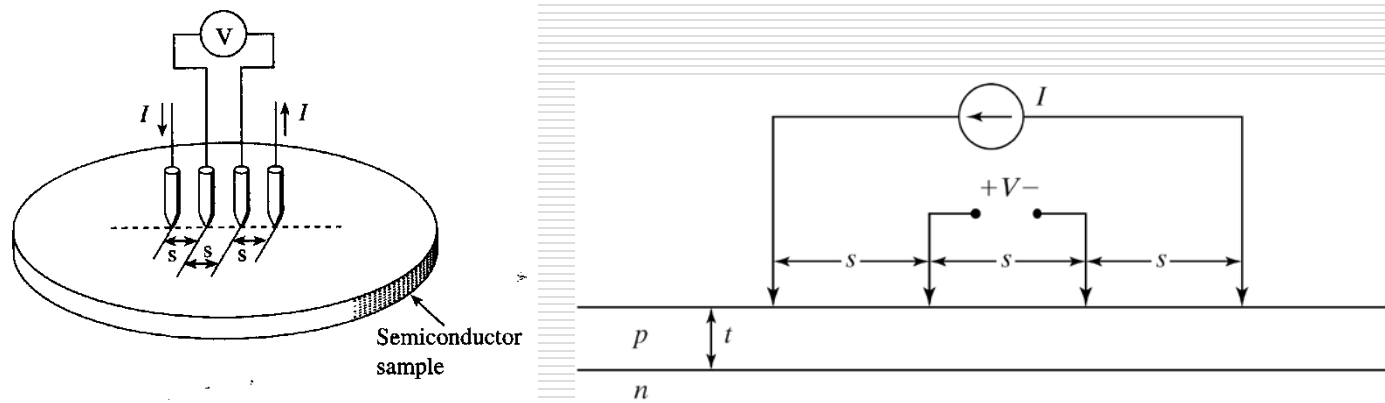
$$\bar{\rho} = \frac{1}{\frac{1}{x_j} \int_0^{x_j} \sigma(x) dx} \rightarrow R_s = \left[q\mu \int_0^{x_j} N(x) dx \right]^{-1}$$

Here μ is the carrier mobility and $N(x)$ is the net impurity concentration.

Resistivity Measurement

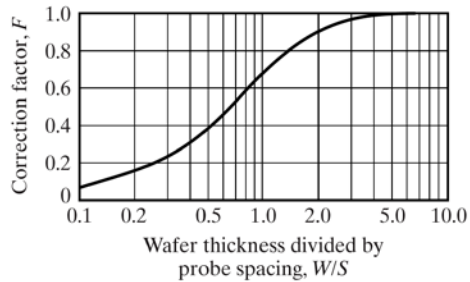
Four-Point Probe

- The 4-point probe with probe spacing s used for direct measurement of bulk wafer resistivity and the sheet resistance of thin diffused layers. A known current is forced throughout the outer probes, and the voltage developed is measured across the inner probes.

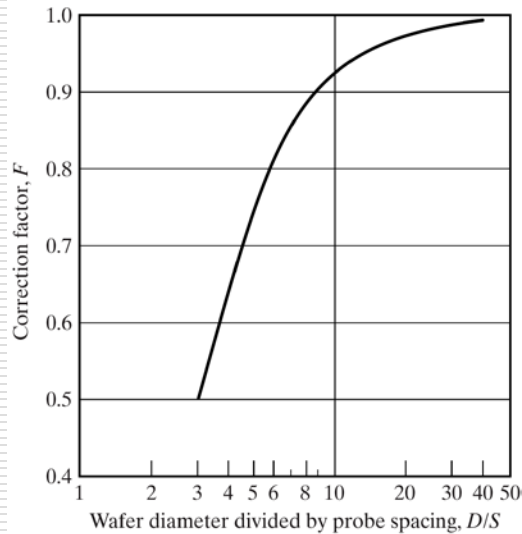


- The probe is assumed to be infinitesimally small and the wafer is semi-infinite.

Four-Point Probe Correction Factors



(a)



(b)

Correction Factors

(a) Wafers Thick Relative to the Probe Spacing

(b) Wafers of Finite Diameter

$$\rho = F \rho_{measured}$$

Resistivity Measurement Four-Point Probe

- In the case of very thick dopant layer, the current is emitted in a spherical form.
 - The differential resistance is defined as

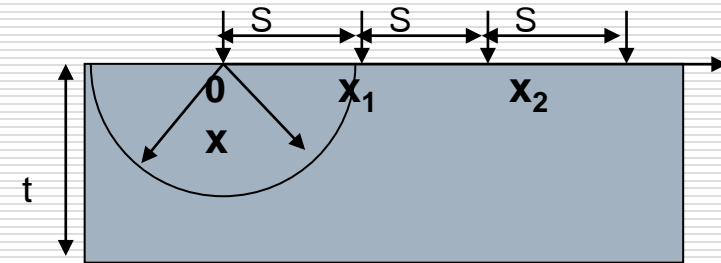
$$\Delta R = \rho \left(\frac{dx}{dA} \right)$$

- The total resistance between the inner probes is

$$R = \int_{x=s}^{x=2s} \rho \left(\frac{1}{2\pi x^2} \right) dx = \rho \left(\frac{-1}{2\pi x} \right) \Big|_s^{2s} = \frac{\rho}{4\pi s}$$

- Due to the summation of current from both probes and counting the symmetry, $R=V/2I$,

$$R = V/2I \rightarrow \rho = 2\pi s V/I$$



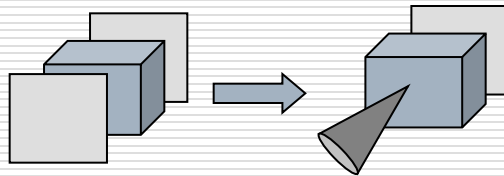
- In the case of very thin dopant layer, the current is emitted in a ring form.

$$R = \int_{x=s}^{x=2s} \rho \left(\frac{1}{2\pi x t} \right) dx = \rho \left(\frac{\ln x}{2\pi t} \right) \Big|_s^{2s} = \frac{\rho \ln(2)}{2\pi t}$$

$$R = V/2I \rightarrow \rho = \pi t V/I, \quad R_s = \rho/t = \frac{\pi}{\ln 2} V/I$$

Two-Probe Measurement

- ❑ Spreading Resistance Profiling (SRP), also known as Spreading Resistance Analysis (SRA), is a technique used *to analyze resistivity vs. depth* in silicon and germanium semiconductors
- ❑ *Spreading resistance* can be understood by thinking of a current passing through a unit cube of a conductor with two uniform electrodes. The current then flows uniformly and the measured resistance is that of the bulk material.



If one electrode is replaced by a probe with very small tip. The resistance is now greater and it is due to electrons crowding in the immediate vicinity of the probe. This resistance is referred to as the spreading-resistance , R_{sp} .

- ❑ If the probe area is made extremely small, then the measured resistance across the cub is dominated by the spreading resistance.

Two-Probe Measurement

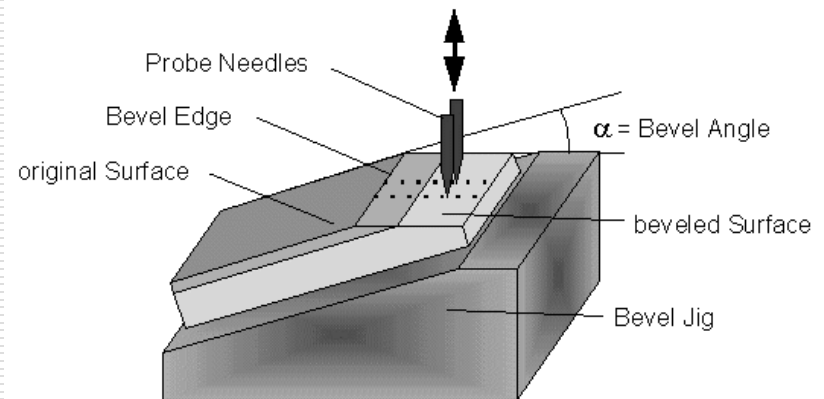
- Two probe tips, usually made of osmium and/or tungsten carbide, are employed. Each one is mounted on the end of a separate arm.
- The probe tips are shaped such that they can be positioned very close together, often with less than a 20 μm separation
- The probe tips are lowered onto the sample, a voltage (~ 5 mV) is applied across the probes and the resistance (spreading-resistance) is measured.

$$\rho = 2aR_{sp}$$

a : The probe diameter

R_{sp} : The spreading resistance

ρ : The resistivity of the slab in ohm-cm



Two-Probe Measurement

- The structure is angle-lapped and the probes are scanned across the junction.
- Approximating the structure as layers of uniform carrier distributions and that the probe measures a very small volume underneath it, the Measured resistance can be approximated by the spreading resistance and hence ρ is directly calculated as a function of depth.

- The resistivity is function of the carrier (impurity) concentration:

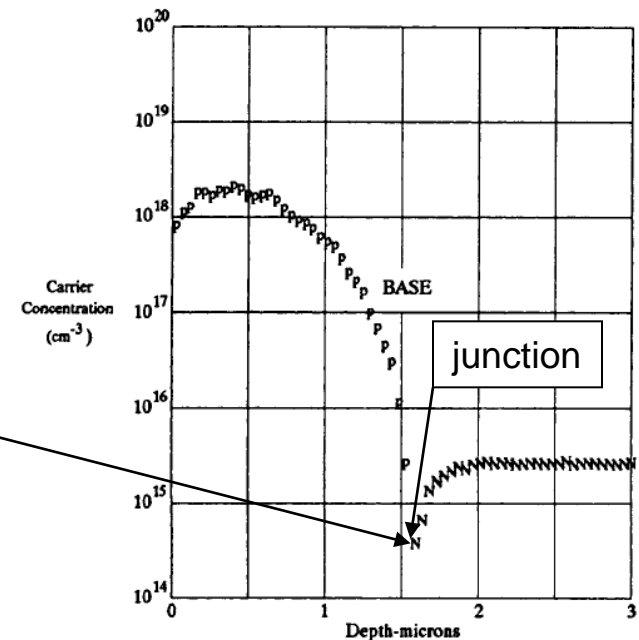
$$\rho(x) = \frac{x_j}{q\mu N(x)}$$

- From the relation of the resistivity and the dopant concentration, the carrier concentration can be calculated.

$$R = \frac{\bar{\rho}}{x_j} = \frac{1}{q\mu N(x)}$$

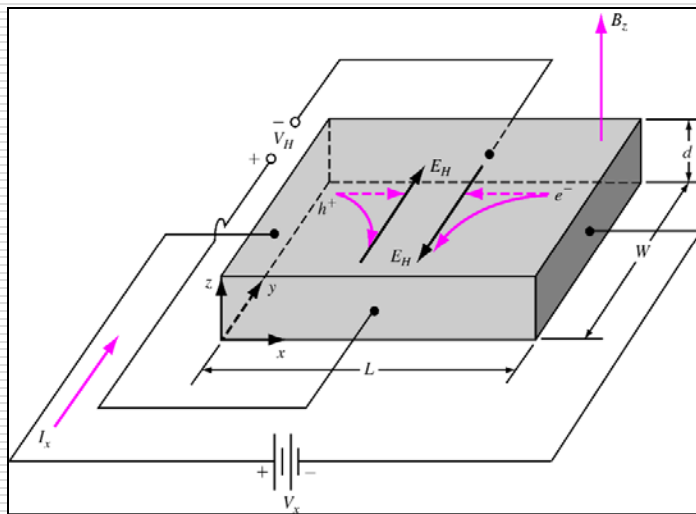
- If the probe motion is controlled and the angle is known, a profile versus depth can be calculated and the junction depth can easily be extracted.

- *The draw-backs of this technique:*
 - Destructive.
 - Measures active dopants only.
 - Requires large area



Hall Effect

- The hall effect is a consequence of the forces that are exerted on moving charges by electric and magnetic fields.
- We can use Hall measurement to
 - Distinguish whether a semiconductor is n or p type
 - To measure the majority carrier concentration
 - To measure the majority carrier mobility



Geometry for measuring the Hall effect.

Secondary Ion Mass Spectrometry

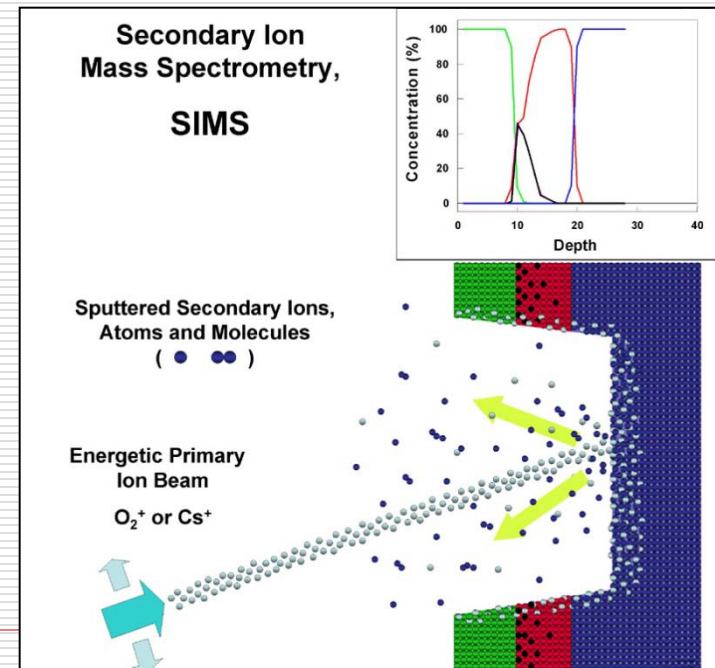
- ❑ SIMS bombards the silicon wafer with low energy ions ($\sim 1\text{--}20\text{ KeV}$) and sputters atoms from the surface.
- ❑ SIMS works by analyzing material removed from the sample by sputtering, and is therefore a locally destructive technique.
- ❑ Continuous bombarding of the surface with the ion beam results in etching and hence **a depth profile can be determined**.

- ❑ *Advantages of the SIMS method*

- Provides atomic distribution vs. depth.
- Identifies all element species.

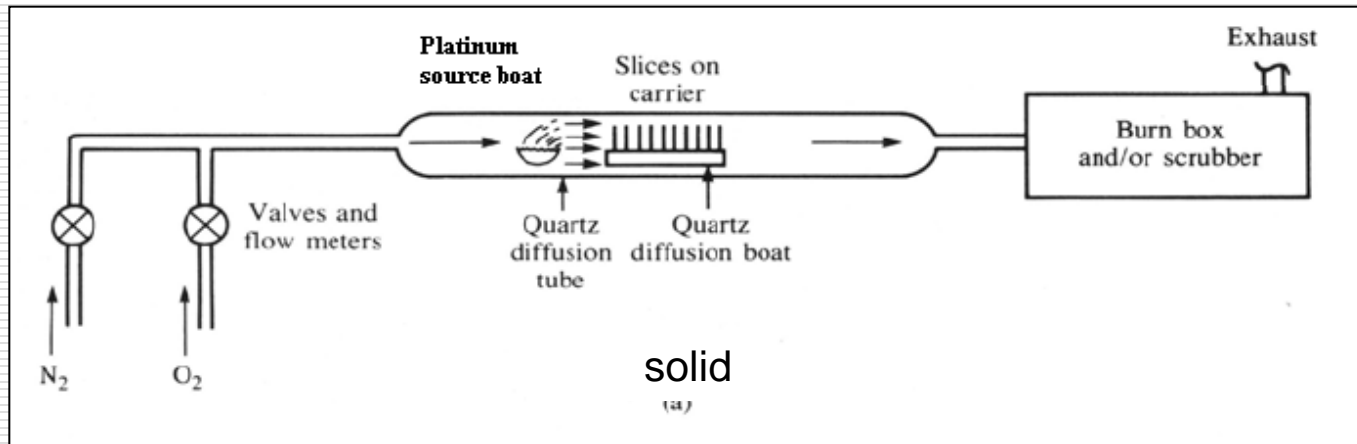
- Draw-backs of the SIMS method*

- Destructive.
- Low sensitivity, specially when decreasing the beam diameter as less ions are collected.
- Expensive equipments.



Diffusion Systems

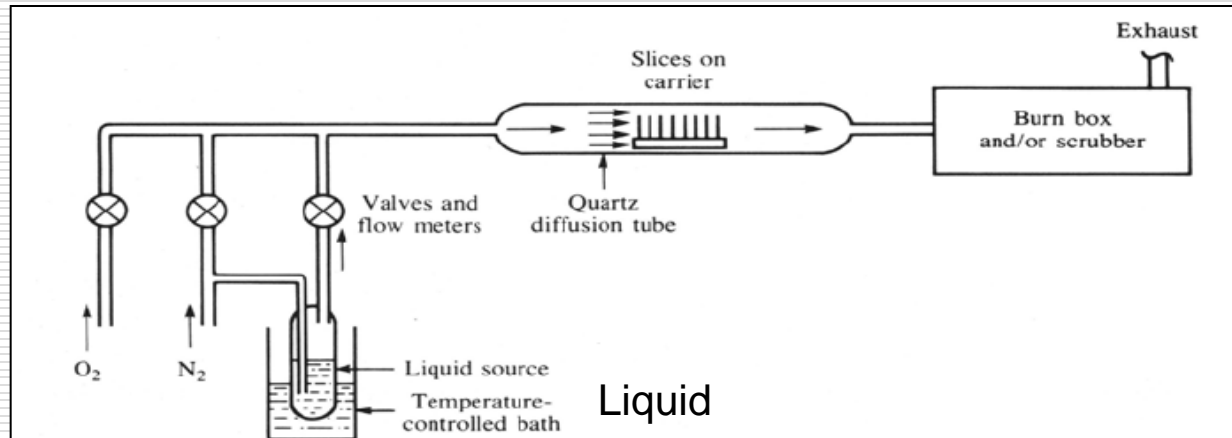
Open-furnace-tube systems



- **Solid-source system** : Carrier gases (N₂ or O₂) flow at a controlled rate over a source boat placed in the furnace tube. Solid boron and phosphorus impurity sources are also available in wafer form and are placed in the boat between adjacent pairs of silicon wafers.
- **Example reaction**: $2\text{As}_2\text{O}_3(\text{Arsenic trioxide}) + 3\text{Si} \rightarrow 4\text{As} + 3\text{SiO}_2$
(forms an oxide layer on the surface)

Diffusion Systems

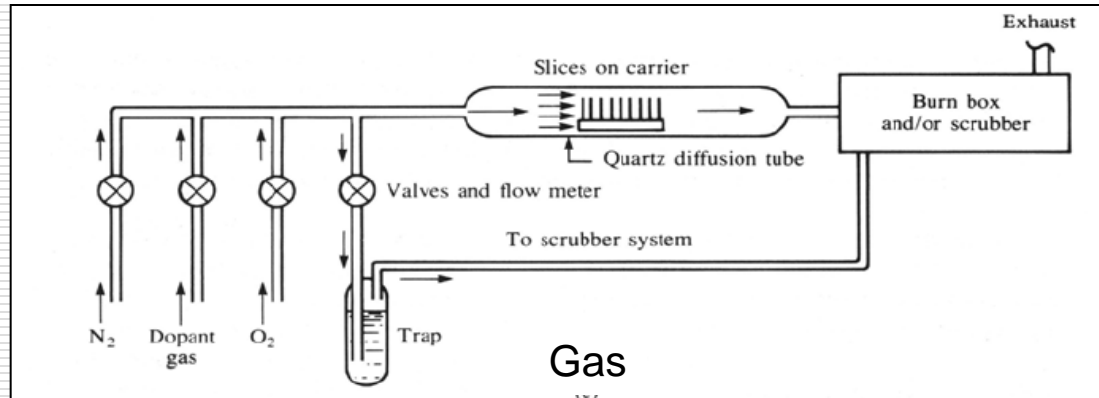
Open-furnace-tube systems



- Liquid-source system: A carrier gas passes through a bubble where it picks up the vapor of the liquid source.
- Example: $4\text{BBr}_3(\text{Boron tribromide}) + 3\text{O}_2 \rightarrow 2\text{B}_2\text{O}_3(\text{Boron trioxide}) + 6\text{Br}(\text{Bromine})$
=> preliminary reaction forms B₂O₃, which is deposited on the surface
; forms a glassy layer

Diffusion Systems

Open-furnace-tube systems



- **Gas-source system** : Dopant species apply directly to the furnace tube in the gaseous state. The common gas source are extremely toxic, and additional input purging and trapping systems are required to ensure that all the source gas is removed from the system. The output of diffusion systems should be processed by burning or by chemical and/or water scrubbing before being exhausted into the atmosphere
- **Examples:**
 - a) B_2H_6 (Diborane)+ $3O_2 \rightarrow B_2O_3$ (Boron oxide)+ $3H_2O$ (at 300 °C)
 - b) i) $4POCl_3$ (Phosphoryl chloride)+ $3O_2 \rightarrow 2P_2O_5$ (Phosphorus pentoxide)+ $6Cl_2$ (Chlorine)
(oxygen is carrier gas that initiates preliminary reaction)
 - ii) $2P_2O_5 + 5Si \rightarrow 4P$ (Phosphorus) + $5SiO_2$ (Silicon dioxide)