III.1 Semiconductors

Semiconductors are materials that lie between metals and insulators in terms of electrical conductivity, a property that grants them a fundamental role in electronics and materials science. In the periodic table, these materials belong to the column of elements with a valency of four, meaning they have four valence electrons in their outer electronic shell. This configuration allows semiconductors to have adjustable conductivity depending on environmental conditions, such as temperature and the presence of impurities. The most common semiconductors are silicon (Si) and germanium (Ge), with silicon being particularly favored due to its abundance and its favorable properties for manufacturing electronic devices.

Silicon, first isolated in 1823 by chemist Jöns Jacob Berzelius, is the main constituent of sand. This material has a density of 2.33 g/cm³ and melts at 1410 °C. It crystallizes in a cubic structure similar to that of diamond, forming a solid network where each cubic centimeter contains about 5×10^{22} silicon atoms. This crystalline structure allows silicon to exhibit electronic properties that are particularly useful in semiconductors.

III.1.2 Pure (or intrinsic) semiconductors

Pure, or intrinsic, silicon is a semiconductor with unique properties that distinguish it from conductors and insulators. At absolute zero (0 K), pure silicon behaves as an insulator, with its valence band fully occupied and its conduction band empty, resulting in no free charge carriers to support electrical conductivity. However, as temperature increases, thermal energy enables some electrons to gain sufficient energy to cross the band gap (approximately 1.1 eV at room temperature) from the valence band to the conduction band. This transition creates free electrons in the conduction band and leaves behind positively charged "holes" in the valence band, allowing both electrons and holes to contribute to electrical conductivity.

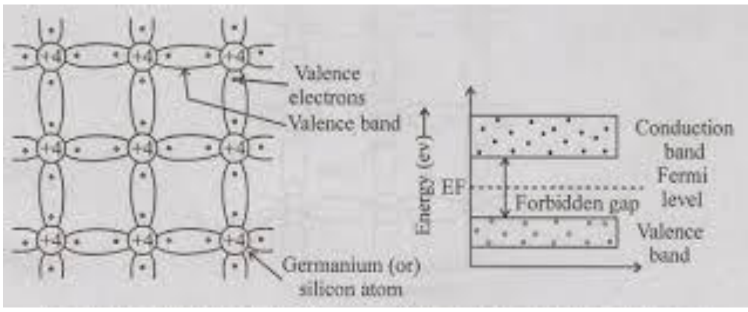

Figure 1 Illustration of band diagram.

In intrinsic silicon, the Fermi level, which represents the energy level at which the probability of electron occupancy is 50% at a given temperature, lies approximately midway between the valence and conduction bands. This mid-gap positioning reflects the equal probability of electron presence in either band in pure silicon at equilibrium, giving intrinsic silicon a relatively low electrical conductivity. To enable practical use in electronics, however, silicon must undergo doping — a process that introduces controlled amounts of impurities into the silicon lattice to either increase electron (N-type) or hole (P-type) concentration, thereby enhancing its conductivity. Doped silicon is thus essential for applications requiring tailored electrical properties, such as transistors, diodes, and integrated circuits, where controlled charge carrier behavior is key to functionality.

III.1.2 Doped silicon

Doping is a process involving the introduction of impurity atoms into the silicon crystal to increase the number of charge carriers: free electrons or holes. This process enables the creation of two types of semiconductors: N-type (negative) or P-type (positive), depending on the properties of the introduced impurity.

P-type silicon

P-type silicon is a type of doped semiconductor created by introducing trivalent impurity atoms, such as boron or aluminum, into the silicon crystal lattice. These dopant atoms have only three valence electrons, one less than silicon's four, which results in the formation of "holes" — essentially vacancies where an electron is missing — in the silicon's valence band. These holes act as positive charge carriers, as they can attract electrons from neighboring atoms, creating a movement of positive charges within the crystal structure. This movement allows for electrical conductivity, albeit at a lower efficiency than electron movement in N-type silicon.

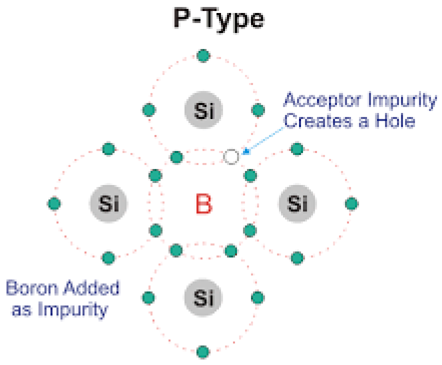

Figure 2 P-type doping of silicon

In P-type silicon, the Fermi level shifts closer to the valence band, reflecting a higher concentration of holes compared to electrons. Since the majority charge carriers in P-type material are holes, it is often used in conjunction with N-type silicon to form P-N junctions, which are the foundation of many electronic devices such as diodes, transistors, and photovoltaic cells. The controlled use of P-type and N-type materials enables precise management of electrical current, essential for the operation of semiconductors in electronics.

N-type silicon

N-type silicon is a semiconductor formed by doping pure silicon with elements from Group V of the periodic table, such as phosphorus (P), arsenic (As), or antimony (Sb). These dopants possess five valence electrons, which means that when they replace silicon atoms (which have four valence electrons), an extra electron is introduced into the silicon lattice. This additional electron becomes a free carrier, contributing to electrical conductivity. In N-type silicon, free electrons are the majority charge carriers, while holes (the absence of electrons) serve as the minority carriers. As a result, N-type silicon exhibits increased conductivity compared to intrinsic silicon, making it suitable for various electronic applications, such as diodes, transistors, and photovoltaic cells.

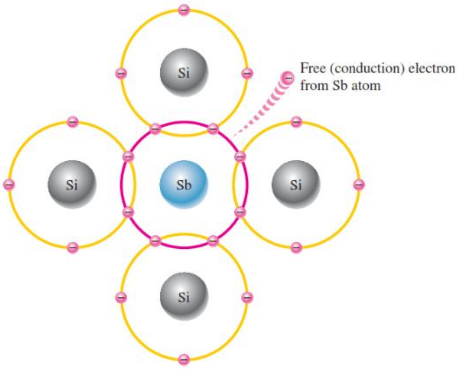

Figure 3 P-type doping of silicon

In the band diagram of N-type silicon, the Fermi level—the energy level at which the probability of finding an electron is 50%—shifts closer to the conduction band due to the abundance of electrons provided by the dopants. This shift indicates that a higher concentration of electrons is available for conduction compared to intrinsic silicon. The conduction band is the range of energy levels where electrons can move freely, facilitating electrical current. The increased density of electrons enhances the material's ability to conduct electricity, making N-type silicon a fundamental component in modern electronics. The interaction between N-type and P-type silicon, creating PN junctions, is critical for the operation of various semiconductor devices, enabling functionalities like rectification and amplification.

III.2 Notion of doping

Doping is the intentional introduction of impurities into a pure semiconductor material to alter its electrical properties. This process is crucial in semiconductor technology, as it enables the creation of materials with specific conductive behaviors, tailored for various electronic applications. The impurities, known as dopants, are typically added in very small amounts—usually in the range of parts per million (ppm)—to achieve the desired effect without significantly altering the semiconductor's overall structure.

There are two primary types of doping: **N-type doping** and **P-type doping**. In N-type doping, elements from Group V of the periodic table, such as phosphorus, arsenic, or antimony, are added to silicon. These dopants have five valence electrons, resulting in an excess of free electrons in the material, which serve as majority charge carriers. Conversely, P-type doping involves the addition of elements from Group III, such as boron, aluminum, or gallium, which have only three valence electrons. This creates "holes" or electron deficiencies, where the majority charge carriers are the holes, and the electrons become the minority carriers. The ability to control the type and concentration of dopants allows engineers to design semiconductors with tailored electrical characteristics, enabling the development of various electronic devices like diodes, transistors, and integrated circuits.

III.2.1 Diffusion mechanism

There are several mechanisms through which atoms can diffuse in a solid, such as a silicon wafer during thermal diffusion. Each mechanism varies based on the type of dopant atoms, the temperature, and the characteristics of the crystal lattice. Here are the main diffusion mechanisms in solids:

1. **Interstitial Diffusion**: In this mechanism, atoms move through the spaces between the atoms in the crystal lattice, known as interstitial sites. This type of diffusion typically involves smaller atoms (like hydrogen or carbon) that are able to fit within these spaces without disturbing the primary lattice structure. Interstitial diffusion is usually faster than other mechanisms since it doesn’t require the movement of atoms from their original lattice positions.
2. **Vacancy Diffusion**: Here, atoms move by hopping into vacant lattice sites or "vacancies." This requires that vacancies are present in the crystal lattice, which is more likely at higher temperatures when atomic vibrations cause some atoms to leave their original positions. Larger dopant atoms, like boron and phosphorus, often diffuse by the vacancy mechanism in silicon, and this is a common mechanism for substitutional atoms in semiconductors.
3. **Substitutional Diffusion**: In substitutional diffusion, the dopant atom replaces a silicon atom in the lattice. This mechanism often involves vacancy diffusion, as a vacancy allows the dopant to replace a silicon atom in the lattice structure. This is typical for larger dopant atoms, like arsenic, that require an available site to incorporate into the silicon lattice.
4. **Interstitialcy Diffusion**: This mechanism is a combination of interstitial and substitutional diffusion. A dopant atom enters an interstitial site, pushes a lattice atom out of its position, and then replaces that atom, which moves into another interstitial position. This complex process is typically slower and less common but can occur in certain materials under specific conditions.
5. **Grain Boundary Diffusion**: In polycrystalline materials, atoms can diffuse along grain boundaries, which are regions between individual crystal grains. These boundaries act as high-diffusivity paths because atoms are less tightly bound than in the crystal lattice, making diffusion much faster along these boundaries. This mechanism is not significant in single-crystal silicon but is important in materials like metals and ceramics.
6. **Surface Diffusion**: Atoms can also diffuse along the surface of a material, where atoms are less tightly bonded than within the bulk material. Surface diffusion is generally faster than bulk diffusion and becomes more significant in thin films or nanostructures, where surface effects dominate due to the high surface-to-volume ratio.

These diffusion mechanisms can operate simultaneously, but their relative significance depends on factors like temperature, lattice structure, and the specific dopants used.

III.2.2 Diffusion Theory

The theory of diffusion describes how particles spread from regions of higher concentration to regions of lower concentration over time, driven by thermal motion. In solids, diffusion involves the movement of atoms or molecules through a crystal lattice and is essential in materials science, especially in processes like semiconductor doping. The theoretical framework for diffusion is based on **Fick's laws**, which mathematically describe how the concentration of diffusing species changes over time and space.

III.2.2.1 Fick's Laws of Diffusion

1. **Fick's First Law**: This law applies to **steady-state diffusion**, where the concentration gradient remains constant over time. It states that the flux of particles, J, is proportional to the concentration gradient, dC/dx :

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

   where:

   - J is the diffusion flux (amount of substance per unit area per unit time),
   - D is the diffusion coefficient (a measure of how easily particles can diffuse in a given material),
   - $\frac{dC}{dx}$ is the concentration gradient.

   The negative sign indicates that diffusion occurs from areas of high concentration to low concentration.

2. **Fick's Second Law**: This law applies to **non-steady-state diffusion**, where the concentration gradient changes over time. Fick's second law describes how the concentration, C, changes with both position x and time t:

   $$\partial C/\partial t = D\partial^2 C/\partial x^2$$

   This equation is used when studying diffusion profiles, such as those found in semiconductor doping, where the concentration of dopants changes dynamically.

- **Doping profile for case of Limited Source**

The case is applied when an initial concentration spike (delta function) diffuses out over time. For example, consider a scenario where a very thin layer of dopant atoms (initial concentration is M) is introduced at $x = 0$ at $t = 0$. The solution to Fick's second law in this case is:

$$C(x,t)=\frac{M}{\sqrt{4\pi Dt}}e^{-\frac{x^2}{4Dt}}$$

where:

- M is the total amount of dopant per unit area introduced initially,
- $\sqrt{4\pi Dt}$ represents the spread of the concentration over time.

This solution describes a Gaussian distribution centered around $x=0$, with the width of the distribution (standard deviation) increasing over time, as $\sigma=\sqrt{2Dt}$. This spread reflects the diffusion of the dopant atoms away from the origin.

Since the dopant will move into the semiconductor as time increases, in order to keep the total dopant S constant, the surface concentration must decrease. This is the case, since the surface concentration with x = 0:

$$C(x,t)=\frac{M}{\sqrt{4\pi Dt}}$$

- **Doping profile of case of constant Source**

In this cases the surface concentration of a dopant is maintained constant over time, such as in a semi-infinite medium with a constant concentration applied at the surface. Suppose that for $t=0s$, the dopant concentration $C(x,0)=0$ for all $x>0$ and a constant surface concentration C_s is applied at $x=0$ for $t>0$. The solution to Fick's second law in this scenario is:

$$C(x,t)=C_s\,erc(x/\sqrt{2Dt})$$

where:

- C_s is the constant surface concentration,
- $erfc$ is the complimentary error function, which describes how the concentration changes with distance and time and $\sqrt{2Dt}$ is the diffusion length.

Figure 5, prentes a typical doping profile, on both linear (upper) and logarithmic (lower) scales, the normalized concentration as a function of depth for three values of the diffusion length corresponding to three consecutive diffusion times and a fixed D for a given diffusion temperature are plotted. Note that as the time progresses, the dopant penetrates deeper into the semiconductor.

In this case, the concentration gradually decreases with depth x from the surface concentration C_s, creating a diffusion front that progresses deeper into the material over time. The error function profile is commonly seen in diffusion processes like doping in semiconductors, where a constant dopant source is applied to the surface.

The total number of dopant atoms per unit area of the semiconductor is given by

$$Q(t) = \int_{0}^{\infty} C(x,t)dx$$

Substituting the profil equation $C(x,t)$ in $Q(t)$ yields

$$Q(t) = \frac{2}{\sqrt{\pi}} C_s \sqrt{Dt} \sim 1.13 C_s \sqrt{Dt}$$

3. Diffusion Coefficient and Arrhenius Relationship

The diffusion coefficient, D, is temperature-dependent and follows the **Arrhenius relationship**:

$D = D_0 e^{-Q/RT}$

where:

- D_0 is the pre-exponential factor (a constant related to the material and diffusing species),
- Q is the activation energy for diffusion,
- R is the gas constant,
- T is the absolute temperature.

This equation shows that higher temperatures increase the diffusion coefficient, allowing particles to move more freely through the lattice.

It has several solutions depending on the boundary and initial conditions. Below are two common cases: the Gaussian profile and the error function (erf) profile.

III.3 Thermal doping of silicon

In silicon doping, two primary diffusion system types are utilized: (i) open-tube and (ii) closed-tube systems. Each diffusion system is designed to ensure specific characteristics essential to effective and reproducible doping in silicon wafers, including:

1. **Control of Surface Concentration**: The dopant surface concentration should be adjustable across a wide range, reaching up to the solid solubility limit.
2. **Surface Integrity**: The doping process must preserve the integrity of the silicon surface.
3. **Removal of Residual Dopants**: Any excess dopant remaining after diffusion should be removable without difficulty.
4. **System Reproducibility and Capacity**: Diffusion systems should reliably handle multiple wafers simultaneously, maintaining uniformity across batches.
5. **Temperature Precision**: Temperature control within the diffusion furnace must maintain a stable zone with minimal variation (±0.5°C).

A diffusion furnace is designed to achieve uniform temperature profiles ranging from 600°C to 1200°C, typically using a feedback controller to adjust the heating rate. The diffusion tube, usually high-purity fused silica, is carefully managed to avoid contamination—one tube is used per dopant type. In larger tubes, a carrier is mechanically inserted to transport wafers, while gas and dopant flow are controlled from the opposite end. The furnace ramps up from 600°C at a controlled rate (3-10°C/min), minimizing thermal shock to both wafers and tube components. In practice, the tube temperature is rarely allowed to fall below 600°C to prevent devitrification of the silica.

III.3.1 Boron Diffusion Systems for P-Type Doping

Boron is the most commonly used dopant for producing P-type silicon due to its high solubility limit (~6×10^{20} atoms/cm^3). However, boron's large lattice misfit factor (0.254) can induce strain and lead to defect formation, so the practical concentration limit is approximately 5×10^{19} atoms/cm^3. Multiple diffusion sources are available for boron, including solid, liquid, and gaseous sources, each selected for specific applications and controlled diffusion profiles.

Solid Sources: Solid boron nitride (BN) wafers, slightly larger than silicon wafers, can serve as boron diffusion sources. These BN wafers are oxidized at 750–1100°C, forming a boron oxide (B_2O_3) layer on the surface, which then serves as the dopant source. This reaction is expressed as:

$$4\,BN\,(s) + 7\,O_2\,(g) \rightarrow 2\,B_2O_3\,(s) + 4\,NO_2$$

With minimal carrier gas flow, often dry nitrogen, back diffusion of contaminants is prevented, and excellent dopant uniformity is achieved.

Liquid Sources (Bubbler Systems): Liquid-phase boron diffusion commonly uses **boric oxide (B_2O_3)**, which reacts with silicon as follows:

$$2\,B_2O_3 + 3\,Si\,(s) \rightarrow 4\,B\,(s) + 3\,SiO_2\,(s)$$

This reaction is conducted in a nitrogen atmosphere with a small oxygen percentage to control B_2O_3 levels on the wafer surface. Liquid sources like BCl_3 are introduced through a bubbler system, where an inert gas, such as argon, is used to vaporize BCl_3 and transport it to the diffusion chamber, facilitating precise control over boron surface concentration.

Gaseous Sources: Diborane (B_2H_6) and boron trichloride (BCl_3) are commonly used gaseous dopants. The reaction for B_2H_6 decomposition is:

$$\boldsymbol{B_2H_6\,(s) + 3\,O_2\,(g) \rightarrow B_2O_3\,(s) + 3\,H_2O}\,(l)$$
$$\boldsymbol{4\,BCl_3\,(g) + 3\,O_2(g) \rightarrow 2\,B_2O_3}(s) + \boldsymbol{6\,Cl_2(g)}$$

Using gaseous dopants allows precise control of concentration through gas flow adjustment, which is essential for achieving high uniformity across wafers.

III.3.2 Phosphorus Diffusion Systems for N-Type Doping

Liquid Sources (Bubbler Systems):Phosphorus diffusion uses liquid sources like phosphorus oxychloride ($POCl_3$), which is widely chosen for its ability to form N-type regions with high doping levels (up to 5×10^{20} atoms/cm^3). $POCl_3$ is introduced in a diffusion system through a bubbler, where nitrogen or argon gas transports vaporized $POCl_3$ to the furnace, enabling controlled deposition onto the wafer surface. The reaction proceeds as follows:

$$4POCl_3(l) + \mathbf{3}\ \boldsymbol{O_2}(g) \rightarrow 2P_2\boldsymbol{O}_5(s) + 6Cl_2(g)$$

P_2O_5 then reacts with the silicon substrate, depositing phosphorus atoms while producing a silicon dioxide (SiO_2) byproduct:

$$2P_2\boldsymbol{O_5} + 5Si \rightarrow 2P(dopant) + 5SiO_2$$

By adjusting the temperature of the $POCl_3$ bubbler, the concentration of phosphorus on the wafer surface can be finely controlled. Alternative gaseous phosphorus sources, such as phosphine (PH_3), are also used, offering high control via gas flow systems.

Solid Sources: A common solid source for phosphorus doping is phosphorus pentoxide (P_2O_5). Solid P_2O_5 wafers or pellets are positioned near the silicon wafers in the diffusion tube. At elevated temperatures, P_2O_5 vaporizes and reacts with the silicon substrate, depositing phosphorus atoms. The high temperature of the diffusion furnace (typically 900–1100°C) facilitates the reaction:

$$2\boldsymbol{P_2O_5} + 5Si \rightarrow 2P(dopant) + 5SiO_2$$

This reaction deposits phosphorus on the silicon surface and simultaneously forms a silicon dioxide layer, which acts as a protective coating to prevent contamination. The solid-source method provides a high phosphorus concentration, making it suitable for applications requiring heavily doped N-type regions.

Gaseous Sources: Gaseous sources for phosphorus doping, such as **phosphine (PH_3)**, are commonly used for precise concentration control and ease of handling in automated systems. PH_3 gas is introduced into the diffusion chamber along with an oxidizing agent, such as O_2, and undergoes the following reaction:

$$PH_3(g) + 4O_2(g) \rightarrow P_2O_5(s) + 3H_2O(l)$$

The resulting P_2O_5 forms on the silicon surface and diffuses into the wafer, creating N-type regions. By controlling the flow of PH_3, precise doping profiles and uniformity across wafers are achievable, which is especially beneficial in applications requiring tight tolerances. Additionally, PH_3 allows lower doping levels to be achieved by adjusting the gas flow, providing flexibility for varying device specifications.

The table below provides comparison of each source's suitability in different semiconductor doping contexts based on control, safety, and contamination factors.

Table 1 A summary of advanges and disadvantage of doping sources.

Doping Source	Advantages	Disadvantages
Gaseous	- Precise control over dopant concentration - Uniform doping across wafer - Flexible for high/low concentrations	- High toxicity, corrosive, and explosive risks - Requires costly, specialized equipment for safe handling
Liquid	- Enables high doping concentrations - Efficient delivery through inert gas (e.g., N_2) - Lower safety risks compared to gases	- Sensitive to system geometry (bubblers, temp controls) - Can form residues (e.g., Cl_2), requiring cleaning
Solid	- Simple setup with lower initial costs - Capable of high concentrations over long durations - Easier handling and safer than gases	- Less uniform doping, limiting precise applications - Higher contamination risk on silicon surface - Residue removal can be difficult

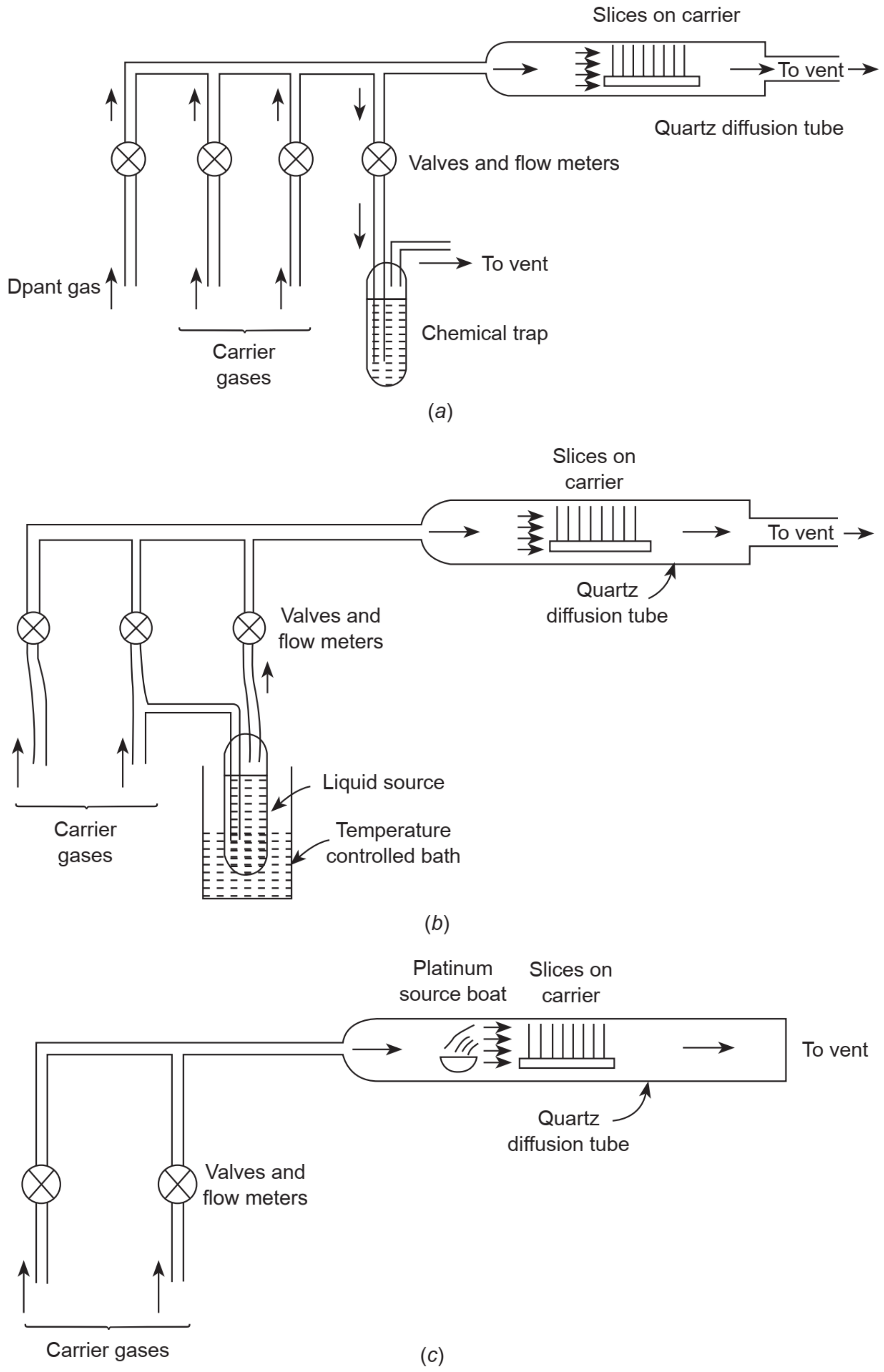

Figure. 4 (a) Gas source (b) liquid source (c) solid source diffusion system.

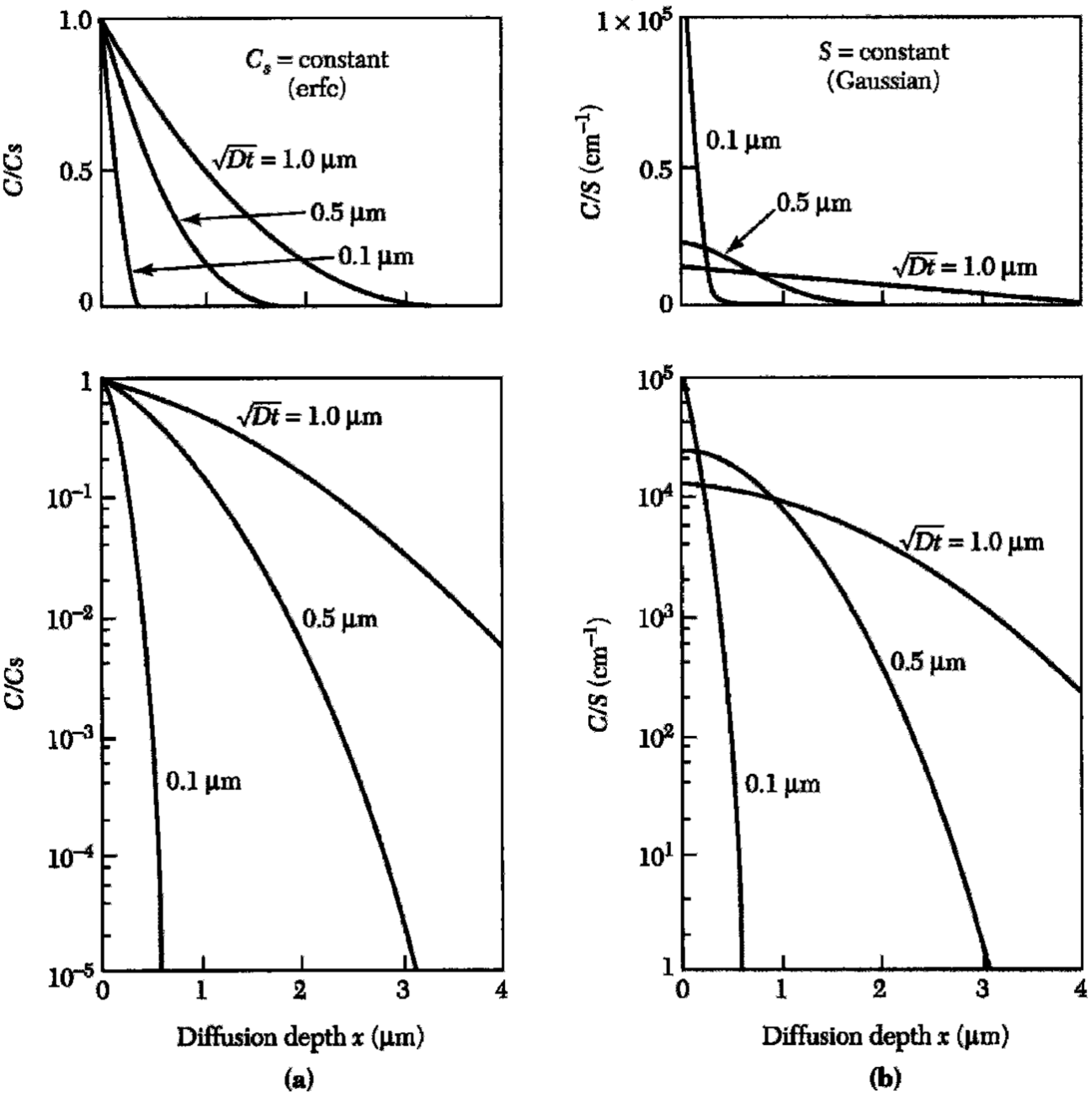

Figure 5. Diffusion profiles : (a) normalized erfc versus distance for successive diffusion times; (b) normalized Gaussian function versus distance.

III.3.4 Comparison between thermal diffusion and ion implantation doping:

A comparison between thermal diffusion doping and ion implantation for silicon doping, presented in a table 2.

Table2 comparison between thermal diffusion doping an ion implantation doping.

Aspect	Thermal Diffusion Doping	Ion Implantation Doping
Process	Dopants are introduced through high-temperature diffusion from a dopant source in the form of gas, liquid, or solid.	Dopant ions are accelerated in an electric field and implanted directly into the silicon lattice.
Doping Profile	Gradual dopant concentration with a decreasing concentration gradient as depth increases (Gaussian or error function profile).	Sharply defined and customizable concentration profile, often box-like or tailored.
Depth Control	Less precise, typically results in deeper diffusion profiles and is temperature-dependent.	Highly precise control over junction depth; adjustable by changing ion energy.
Uniformity of Doping	Depends on dopant source uniformity and diffusion conditions; generally less precise.	Very high uniformity and precision due to controlled implantation parameters.
Equipment Requirements	Uses a high-temperature furnace; simpler and less expensive equipment.	Requires specialized ion implantation equipment, which is more expensive and complex.
Surface Damage	Minimal damage as atoms diffuse naturally into the lattice at high temperature.	Can cause lattice damage due to high-energy ions; requires post-implantation annealing to repair damage.
Annealing Requirements	Annealing is inherent in the diffusion process at high temperatures.	Annealing is necessary to repair implantation damage and activate dopants.
Temperature Sensitivity	High temperatures (typically 800–1200°C) are essential for effective diffusion.	Can be done at lower temperatures initially, but annealing is required afterward.
Control Over Doping Concentration	Moderate control; depends on time and temperature settings.	Excellent control, allowing for precise adjustment of dopant concentration.
Typical Applications	Widely used in basic doping processes, solar cells, shallow junctions, and simpler microelectronic devices.	Preferred for advanced IC fabrication, shallow and precise junctions, and complex semiconductor devices.
Cost	Generally lower cost; ideal for less precise doping applications.	Higher cost due to sophisticated equipment and precise control.

In summary, **thermal diffusion doping** is simpler and cost-effective, suited for applications where deep, gradual doping profiles are acceptable, such as solar cells or certain basic semiconductor processes. **Ion implantation** provides higher precision and control, making it ideal for advanced microelectronics that require shallow, sharp doping profiles and highly uniform dopant distributions.