CHAPTER 1

Basic Semiconductor Physics and Technology

Power electronic circuits utilise power semiconductor switching devices which *ideally* present infinite resistance when off, zero resistance when on, and switch instantaneously between those two states. It is necessary for the power electronics engineer to have a general appreciation of the semiconductor physics aspects applicable to power switching devices so as to be able to understand the vocabulary and the non-ideal device electrical phenomena. To this end, it is only necessary to attempt a qualitative description of switching devices and the relation between their geometry, material parameters, and physical operating mechanisms.

Typical power switching devices such as diodes, thyristors, and transistors are based on a monocrystalline group IV silicon semiconductor structure or, increasingly, a group IV polytype, silicon carbide. These semiconductor materials are distinguished by having a specific electrical conductivity, σ, somewhere between that of good conductors (>10^{20} free electron density) and that of good insulators (<10^{3} free electron density). Silicon is less expensive, more widely used, and a more versatile processing material than silicon carbide, thus the electrical characteristics and processing properties of silicon are considered first, and in more detail.

In pure silicon at equilibrium, the number of *electrons* is equal to the number of *holes*. The silicon is called *intrinsic* and the electrons are considered as negative charge-carriers. Holes and electrons both contribute to conduction, although holes have less mobility due to the covalent bonding. Electron-hole pairs are continually being *generated* by thermal ionization and in order to preserve equilibrium previously generated pairs *recombine*. The intrinsic carrier concentrations n_i are equal, small (1.4×10^{10} /cc), and highly dependent on temperature. In order to fabricate a power-switching device, it is necessary to increase greatly the free hole or electron population. This is achieved by deliberately doping the silicon, by adding specific impurities called *dopants*. The doped silicon is subsequently called *extrinsic* and as the concentration of dopant N_c increases, the resistivity $\rho = 1/\sigma$ decreases.

n-type:- Silicon doped with group V elements, such as As, Sb or P, will be rich in electrons compared to holes. Four of the five valence electrons of the group V dopant will take part in the covalent bonding with the neighbouring silicon atoms, while the fifth electron is only weakly attached and is relatively 'free'. The semi-conductor is called *n-type* because of its free negative charge-carriers. A group V dopant is called a *donor,* having donated an electron for conduction. The resultant electron impurity concentration is denoted by N_D - the donor concentration.

p-type:- If silicon is doped with atoms from group III, such as B, Aℓ, Ga or In, which have three valence electrons, the covalent bonds in the silicon involving the dopant will have one covalent-bonded electron missing. The impurity atom can accept an electron because of the available thermal energy. The dopant is thus called an *acceptor*, which is ionised with a net positive charge. Silicon doped with acceptors is rich in holes and is therefore called *p-type.* The resultant hole impurity concentration is denoted by N_A - the acceptor concentration.

Electrons in n-type silicon and holes in p-type are called *majority carriers,* while holes in n-type and electrons in p-type are called *minority carriers*. In a given silicon material, at equilibrium, the product of the majority and minority carrier concentration is a constant:

$$p_o \times n_o = n_i^2 \tag{1.1}$$

where p_o and n_o are the hole and electron equilibrium carrier concentrations.

Therefore, the majority and minority concentrations are given by:

$$\text{for an } n\text{-type } n_o = N_D \text{ therefore } p_o = \frac{n_i^2}{N_D} \text{ and}$$
$$\text{for a } p\text{-type } p_o = N_A \text{ therefore } n_o = \frac{n_i^2}{N_A} \tag{1.2}$$

These equations show that the number of minority carriers decreases as the doping level increases. The resistivity, ρ, of doped silicon is

$$\rho = \frac{1}{\sigma} = \frac{1}{q\left(\mu_n n + \mu_p p\right)} = \frac{\mathrm{E}}{J} = \frac{V}{\frac{I}{A}L} \tag{1.3}$$

where: $\sigma = 1/\rho$ = conductivity, $\Omega^{-1}.cm^{-1}$
$\rho = 1/\sigma$ = resistivity, Ω.cm
μ_n = electron mobility, cm^2 / V-s
μ_p = hole mobility, cm^2 / V-s
q = electron charge, 1.602×10^{-19} C
n = electron concentration, cm^{-3}
p = hole concentration, cm^{-3}

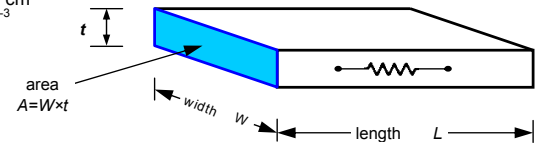


Figure 1.1. *Elemental doped silicon.*

Resistance of semiconductor materials is usually expressed in terms of sheet resistance $R_\square$, which is related to resistance as follows. At a given temperature, the impurity depth x_j, mobility μ, and impurity distribution $N(x)$ are related to sheet resistance by

$$R_\square = \frac{1}{q\int_0^{x_j} \mu N(x)dx} \quad \Omega/\text{square} \tag{1.4}$$

The average resistivity is $\bar{\rho} = R_\square x_j$ and given a length L and width w, as defined in figure 1.1, the resistance, R, is given by

$$R = \rho\frac{L}{A} = \frac{\rho}{t}\times\frac{L}{w} = R_\square\frac{L}{w} \quad \Omega \tag{1.5}$$

For parallel n-doped profiles, the resistance can be estimated by treating each layer independently:

$$R_{total}^{-1} = R_1^{-1} + R_2^{-1} + .. = \frac{wt_1}{L\rho_1} + \frac{wt_1}{L\rho_1} + .. = \frac{w}{L}\left(\frac{t_1}{\rho_1} + \frac{t_1}{\rho_1} + ..\right) = \frac{w}{L}\left(R_{\square 1}^{-1} + R_{\square 2}^{-1} + ..\right) = \sum_{i=1} q\mu_{ni}N_{Di}t_i \tag{1.6}$$

Example 1.1: ***Resistance of homogeneously doped silicon***

Silicon doped with phosphorous ($N_D = 10^{17}$ /cm^3) measures 100μm by 10μm by 1μm. Calculate the sheet resistance and resistance between opposite faces, assuming the electron mobility at this doping level is μ_n = 720cm^2/ V-s. Doping to produce a p-type material has a hole mobility of 40% that for electrons. Recalculate sheet resistance and resistance values.

Solution

From equation (1.3), the resistivity, ρ, of doped silicon is

$$\rho = \frac{1}{\sigma} = \frac{1}{q\left(\mu_n n + \mu_p p\right)}$$

Since $n >> p$ in the n-type silicon and assuming complete ionization

$$\rho \approx \frac{1}{q\mu_n n} \approx \frac{1}{q\mu_n N_D} = \frac{1}{1.6\times10^{-19}\times720\times10^{17}} = 0.086\,\Omega\text{cm}$$

For a length of 100μm, the resistance is

$$R = \rho \times \frac{Length}{Area} = \rho \times \frac{L}{w \times t} = 0.086 \times \frac{100 \times 10^{-4}}{10 \times 10^{-4} \times 1 \times 10^{-4}} = 8.6\,\text{k}\Omega$$

From equation (1.5) the sheet resistance is given by

$$R_s = R\frac{W}{L} = 8.6\,\text{k}\Omega \times \frac{10 \times 10^{-4}}{100 \times 10^{-4}} = 860\,\Omega/\text{square}$$

If the length is assumed to be one of the shorter dimensions, then for a length 10μm or 1μm, the resistance is 86Ω or 0.86Ω, respectively, while the sheet resistance possibilities, depending on the thickness reference axis, are 86 Ω/square and 8.6 Ω/square.

For a p-type material, the 40% decrease in mobility of holes μ_p increases resistivity by a factor of 1/0.4 = 2.5. Each aspect resistance therefore increases by a factor 2.5, viz., increases to 21.5kΩ, 215Ω, and 2.15Ω for lengths 100μm, 10μm, and 1μm, respectively. From equation (1.4) the sheet resistances are increased to 2.15kΩ/square, 215Ω/square, and 21.5Ω/square.

♣

The carrier concentration equilibrium can be significantly changed by irradiation with photons, the application of an electric field or by heat. Such carrier injection mechanisms create *excess carriers.*

If n-type silicon is irradiated by photons with enough energy to ionise the valence electrons, electron-hole pairs are generated. There is already an abundance of majority electrons in the n-type silicon, thus the photon-generated excess minority holes are of more relative and detectable importance. If the light source is removed, the time constant associated with recombination, or decay of excess minority carriers, is called the *minority carrier hole lifetime,* τ_h. For a p-type silicon, exposed to light, excess minority electrons are generated and after the source is removed, decay at a rate called the *minority carrier electron lifetime,* τ_e. The minority carrier lifetime is often termed the *recombination lifetime.*

A manufacturing processing difficulty with high-voltage, large-area semiconductor devices is that of obtaining uniformity of n-type phosphorus doping throughout the usual high-resistivity silicon starting material. Normal crystal growing (by liquid encapsulated, contactless, Czochralski crystal growth – see section 1.19.3i) and doping techniques give no better than ±10 per cent fluctuation around the wanted resistivity at the required low concentration levels (<10^{14} /cc). Final device electrical properties will therefore vary widely in all lattice directions. Tolerances better than ±1 per cent in resistivity and homogeneous distribution of phosphorus can be attained by neutron radiation, commonly called *neutron transmutation doping,* NTD. The neutron irradiation flux transmutes silicon atoms first into a silicon isotope with a short 2.62-hour half-lifetime, which then decays into phosphorus. Subsequent thermal annealing removes any crystal damage caused by the irradiation. Neutrons can penetrate over 100mm into silicon, thus large silicon crystals can be processed using the NTD technique.

A p-n junction is the location in a semiconductor where the doping changes from p to n while the monocrystalline lattice continues undisturbed. A bipolar diode is thus created, which forms the basis of any bipolar semiconductor device.

The donor-acceptor doping junction is formed by any one of a number of process techniques, namely alloying, diffusion, epitaxy, ion implantation or the metallization for ohmic contacts.

Power semiconductor device processing involves most of the following range of process possibilities.

- ***Deposition*** is any process that grows, coats or otherwise transfers a material onto the wafer. Available technologies consist of physical vapour deposition (PVD), chemical vapour deposition (CVD), electrochemical deposition, molecular beam epitaxy (MBE), and atomic layer deposition.
- ***Removal*** processes are any that remove material from the wafer either in bulk or selective form and consist primarily of etch processes, both wet etching and dry etching such as reactive ion etch.
- ***Patterning*** covers the series of processes that shape or alter the existing shape of the deposited materials and is generally referred to as lithography. In conventional lithography, the wafer is coated with a chemical called a *photoresist*. The photoresist is exposed by a 'stepper', a machine that focuses, aligns, and moves the mask, exposing select portions of the wafer to short wavelength UV light. The unexposed regions are washed away by a developer solution. After etching or other processing, the remaining photoresist is removed by oxygen plasma ashing or stripping.
- ***Modification*** of electrical properties consists of doping transistor sources and drains in diffusion furnaces and by ion implantation. These doping processes are followed by furnace annealing or in advanced devices, by rapid thermal annealing which serve to activate the implanted dopants. Modification of electrical properties extends to reduction of the dielectric constant in low-k insulating materials via exposure to ultraviolet light.

1.1 Processes forming and involved in forming semiconductor devices

1.1.1 Alloying

At the desired region on an n-type wafer, a small amount of p-type impurity is deposited. The wafer is then heated in an inert atmosphere and a thin film of melt forms on the interface. On gradual, slow cooling, a continuous crystalline structure results, having a step or abrupt pn junction as shown in figure 1.2. This process is not employed to form modern p-n junctions but can be used at the metallisation stage of wafer fabrication.

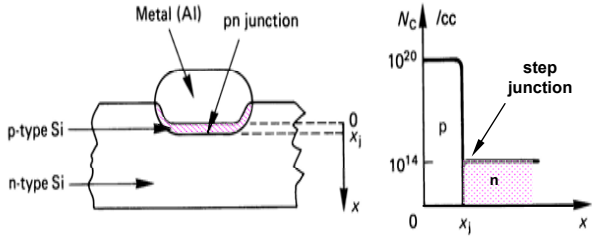


Figure 1.2. *n-Si to Aℓ metal alloy junction: (a) cross-section where x_j is the junction depth below the metal-semiconductor boundary and (b) doping profile of the formed step junction.*

1.1.2 Diffusion

Diffusion, the movement of a chemical species from a region of high concentration to a region of lower concentration, is one of the two major processes by which chemical dopants are introduced into a semiconductor (the other process being ion implantation). The controlled diffusion of dopants into silicon to alter the type and level of conductivity of semiconductor materials is the foundation of forming a p-n junction and formation of devices during wafer fabrication, as shown in figure 1.3. It is used to form bases, emitters, and resistors in bipolar devices, as well as drains and sources in MOS devices. It is also used to dope polysilicon layers. It is not applicable to SiC processing.

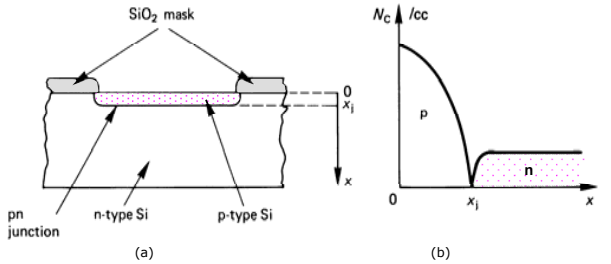


Figure 1.3. *Diffused pn junction: (a) cross-section where x_j is the junction depth below the silicon surface and (b) doping concentration profile.*

The mathematics that govern the mass transport phenomena of diffusion is based on two concepts.

First Concept

Whenever an impurity concentration gradient, ∂C/∂x, exists in a finite volume of a matrix substance (the silicon substrate in this context), the impurity material has a natural tendency to move in order to distribute itself more evenly within the matrix and decrease the concentration gradient.

Given time, this flow of impurities eventually results in homogeneity within the matrix, causing the net flow of impurities to stop. The mathematics of this transport mechanism is based on the flux, *J*, of material across a given plane is proportional to the concentration gradient across that plane. That is:

$$J = -D\frac{\partial N(x,t)}{\partial x} \tag{1.7}$$

where J is the flux,
$D = \mu kT$ is the diffusion constant or diffusivity for the material that is diffusing the solvent, m/s,
$\partial N(x,t)/\partial x$ is the concentration gradient. k is Boltzmann's constant, and μ is ionic mobility.
The diffusion constant, D, of a material is also referred to as *diffusion coefficient* or *diffusivity* and is related to mobility, μ, by $D=\mu kT$. It is expressed in units of length^2/time, such as μm^2/hour. The negative sign of the right side of the equation indicates that the impurities flow to the lower concentration.

Second Concept
Equation (1.7) does not account for the fact that the gradient and local concentration of the impurities in a finite volume of material decreases with an increase in time, an aspect that is important to diffusion processes.
The flux J_1 of impurities entering a section of a material with a concentration gradient is different from the flux J_2 of impurities leaving the same section. From the law of conservation of matter, the difference between J_1 and J_2 must result in a change in the concentration of impurities within the section, assuming that no impurities are formed or consumed in the section.
The second concept states that the change in impurity concentration over time is equal to the change in local diffusion flux, or

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

or, from the first concept, equation (1.7)

$$\frac{\partial N(x,t)}{\partial t} = \frac{\partial\left(D\frac{\partial N(x,t)}{\partial x}\right)}{\partial x} \tag{1.8}$$

If the diffusion coefficient is independent of position, such as when the impurity concentration is low, then the second concept may be simplified to:

$$\frac{\partial N(x,t)}{\partial t} = D\frac{\partial^2 N(x,t)}{\partial x^2} \tag{1.9}$$

There are two major ways by which to deposit impurities into a substance by thermal diffusion. In the first method, known as *predeposition*, a flux of impurities continuously arrives at the surface of the substrate such that the concentration gradient of the impurity remains constant at the surface of the substrate, as shown in figure 1.4b. In the second method, known as redistribution or *drive-in* diffusion, a thin layer of the impurity material is deposited on the substrate. In this case, the impurity gradient at the surface of the substrate decreases with time, as shown in figure 1.4c.

The semiconductor diffusion process is usually performed in two steps: predeposition and then drive-in.
During *predeposition*, the impurity dopant is added to the wafer n-type silicon substrate. Predeposition is done in a diffusion furnace at temperatures around 1000 to 1250ºC. The dopant is introduced into the furnace, and may be in the form of a gas, solid or liquid. Gaseous dopants are mixed with an inert carrier gas, such as nitrogen or argon, and introduced into the furnace. Solid dopants are often applied in a powder form. The solid is heated and a stream of carrier gas moves the dopant into the furnace. Liquid sources are used by bubbling an inert carrier gas through the liquid dopant, and the gas saturated with the liquid is added to the furnace. This compound breaks down as a result of the high temperature, and is slowly diffused into the substrate. The maximum impurity concentration occurs at the surface, tailing off towards the inside.
The wafers are then put into a second furnace at higher temperatures (about 1300ºC) to *drive-in* the dopant. The drive-in process usually occurs in an oxidizing atmosphere so that a protective layer of SiO_2 is grown over the diffused layer.

Table 1.1: Dopants and chemical reactions

Dopant state	Dopant type	dopant	chemistry
gas	p-type	diborane B_2H_6	$B_2H_6 + 3O_2 \rightarrow B_2O_3 + 3H_2O$
	n-type	arsine AsH_3	$2AsH_3 + 3O_2 \rightarrow As_2O_3 + 3H_2O$
		phosphine PH_3	$2PH_3 + 4O_2 \rightarrow B_2O_5 + 3H_2O$
liquid	p-type	BBr_3	$4BBr_3 + 3O_2 \rightarrow 2B_2O_3 + 6Br_2$
	n-type	$AsC\ell_3$, $P0C\ell_3$	$4P0C\ell_3 + 3O_2 \rightarrow 2P_2O_5 + 6C\ell_2$
solid	p-type	BN, B_2O_3	$2B_2O_3 + 3Si_4 \rightarrow 4B + 3SiO_2$
	n-type	As_2O_3, P_2O_5	$2As_2O_3 + 3Si \rightarrow 4As + 3Si0_2$
			$2P_2O_5 + 5Si_4 \rightarrow 4P + 5SiO_2$

Typical dopants and silicon chemical reactions are shown in Table 1.1, while common diffusion coefficients and activation energies, referenced to 0 degree Kelvin, are shown in Table 1.2.
The diffusion process is a junction forming technique that is not applicable to silicon carbide, and other wide band gap material, wafer processing.

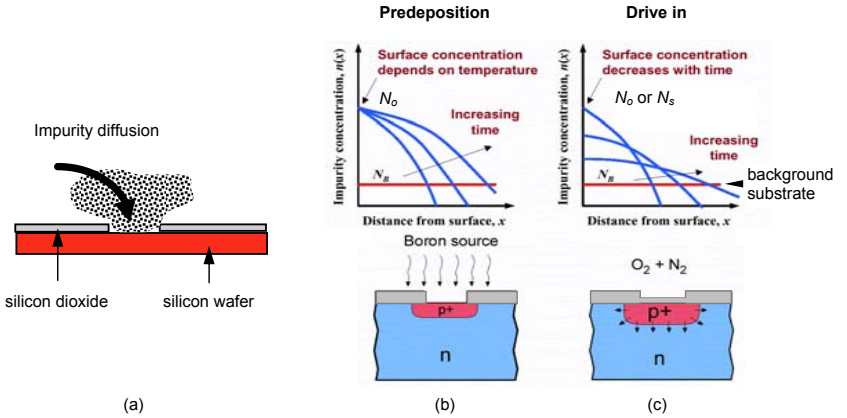


Figure 1.4. *Diffusion processes:*
(a) pictorial representation of mechanism; (b) predeposition diffusion; and (c) drive in diffusion.

The doping profile is mathematically defined and is varied by controlling the vapour mixture concentration, the furnace temperature, and time of diffusion.
If the source concentration is continuously replenished – ***predeposition dose***, thus maintained constant, the surface concentration is $N(0,t) = N_s$, ($N(\infty,t) = 0$), and the initial concentration is $N(x,0)=0$, then the doping profile is given by a complementary error function, *erfc*.

$$N(x,t) = N_s\, erfc\left(\frac{x}{2\sqrt{Dt}}\right)$$

$$= N_s\left(1 - erf\left(\frac{x}{2\sqrt{Dt}}\right)\right) = N_s\left(1 - erf\,(u)\right) = N_s\left(1 - \frac{2}{\sqrt{\pi}}\int_0^u e^{-v^2}dv\right) \tag{1.10}$$

where $u = \dfrac{x}{2\sqrt{Dt}} = \dfrac{x}{\text{diffusion length}}$

The area under the diffusion profile is the total amount of dopant diffused into the wafer:

$$Q(t) = \int_o^\infty N(x,t)\,dx = \frac{2}{\sqrt{\pi}}N_s\sqrt{Dt} = 1.13N_s\sqrt{Dt} \tag{1.11}$$

The junction depth is where the doping profile $N(x,t)$ equals the background doping N_B level, that is

$$N(x_j,t) = N_s\, erfc\left(\frac{x_j}{2\sqrt{Dt}}\right) = N_B$$

Rearranging, gives the junction depth x_j as

$$x_j = 2\sqrt{Dt} \times erfc^{-1}\frac{N_B}{N_s} \tag{1.12}$$

If natural dopant depletion occurs **– *drive in***, that is the initial dose S at the surface is not replenished, then the profile is an exponential function, which gives a Gaussian diffusion distribution.

$$N(x,t) = \frac{S}{\sqrt{\pi Dt}}e^{-\frac{x^2}{4Dt}} \tag{1.13}$$

where

$$\int_o^\infty N(x,t)\,dx = S = \text{non-replenished inital surface dose} \tag{1.14}$$

The diffusion length, $x = 2\sqrt{Dt}$, is an approximate measure of how far the dopant has diffused, which is the distance from the surface to where the concentration has fallen to 1/*e*, 36.8%.

The surface concentration, which is not replenished ($dN(0,t)/dx = 0$), diminishes with time, according to

$$N(0,t) = \frac{S}{\sqrt{\pi Dt}} \tag{1.15}$$

The junction depth is where the doping profile $N(x,t)$ equals the background doping N_B level, that is

$$N(x_j,t) = \frac{S}{\sqrt{\pi Dt}} e^{-\frac{x^2}{4Dt}} = N_B$$

Rearranging, gives the junction depth x_j as

$$x_j = 2\sqrt{Dt} \times \left(\ell n \frac{S}{\sqrt{\pi Dt}\, N_B} \right)^{½} = 2\sqrt{Dt} \times \left(\ell n \frac{N(0,t)}{N_B} \right)^{½} \tag{1.16}$$

In both processing cases, $N(\infty,t) = 0$.
The gradient of the diffusion profile is obtained by differentiating equation (1.13):

$$\frac{dN(x,t)}{dt} = -\frac{x}{2Dt} N(x,t) \tag{1.17}$$

The gradient is zero at x = 0 and x = ∞, and the maximum gradient occurs at $x=\sqrt{2Dt}$.
Diffusivity D varies with temperature according to

$$D = D_o \, e^{-\frac{E_a}{kT}} \tag{1.18}$$

where D_o = diffusion coefficient (in cm^2/s) extrapolated to infinite temperature
E_a = activation or threshold energy in eV, which is not particularly temperature dependant.
For multiple diffusions, the effective $Dt_{effective}$ equals the sum of each diffusion process Dt.

Table 1.2: Typical diffusion coefficients and activation energies at 0K

Element		D_o	E_a
	0 K	cm^2/s	eV
boron	B	1.0	3.50
phosphorous	P	4.7	3.68
antimony	At	4.58	3.88
arsenic	As	9.17	3.99
indium	In	1.20	3.50

Example 1.2: ***Constant Surface Concentration diffusion - predeposition***

For a constant-source boron diffusion into n-type 10^{15} cm^{-3} silicon at 1000°C, the surface concentration is maintained at 10^{19} cm^{-3} and the diffusion time is 1 hour. Find

i. Total amount of dopant diffused, $Q(t)$ and the gradient at $x = 0$ and
ii. The gradient and location (junction depth) where the dopant concentration reaches 10^{15} cm^{-3}.

Solution

Using data for boron in Table 1.2, equation (1.18) gives the diffusion coefficient of boron at 1000°C as

$$D = D_o e^{-\frac{E_a}{kT}} = 24e^{-\frac{3.05}{8.614\times10^{-5}\times1273}} = 1.39\times10^{-14}\,cm^2/s$$

so the diffusion length is

$$\sqrt{Dt} = \sqrt{1.39\times10^{-14}\times3600} = 7.07\times10^{-6}\,cm$$

i. The area under the diffusion profile from equation (1.11) is

$$Q(t) = 1.13N_s\sqrt{Dt} = 1.13\times10^{19}\times7.07\times10^{-6} = 8.0\times10^{13}\,cm^{-2}$$

$$\left.\frac{dN}{dx}\right|_{x=0} = -\frac{N_s}{\sqrt{\pi Dt}} = -\frac{10^{19}}{\sqrt{\pi}\times7.07\times10^{-6}} = -7.98\times10^{23}\,cm^{-4}$$

ii. From equation (1.10) rearranged, when $N_B = 10^{15}$ cm^{-3}, the junction depth x_j is given by

$$x_j = 2\sqrt{Dt}\times erfc^{-1}\left(\frac{N_B}{N_s}\right) = 2\sqrt{Dt}\times erfc^{-1}\left(\frac{10^{15}}{10^{19}}\right)$$

$$= 2\times7.07\times10^{-6}\times2.75 = 0.389\,\mu m$$

$$\left.\frac{dN}{dx}\right|_{x=0.389\mu m} = -\frac{N_s}{\sqrt{\pi Dt}} e^{-\frac{x^2}{4Dt}} = -4.0\times10^{20}\,cm^{-4}$$

♣

Example 1.3: ***Constant Total Dopant diffusion – drive in - 1***

Arsenic was pre-deposited by arsine gas, and the resulting dopant per unit area was 10^{14} cm^{-2}. How long would it take to drive the arsenic in to x_j = 1 μm? Assume a background doping of $N_{sub} = 10^{15}$ cm^{-3}, and a drive-in temperature of 1200°C. For As, assume D_o = 24 cm^2/s, and E_a = 4.08 eV at 1200°C.

Solution

From equation (1.18) the diffusion coefficient for arsenic at 1200°C is

$$D = D_o e^{-\frac{E_a}{kT}} = 24\times e^{-\frac{4.08}{8.614\times10^{-5}\times1473}} = 2.602\times10^{-13}$$

Rearranging equation (1.13) gives

$$x_j^2 = 10^{-8} = 4Dt\times\ell n\left(\frac{S}{N_B\sqrt{\pi Dt}}\right) = 1.04\times10^{-12}t\times\ell n\left(\frac{1.106\times10^5}{\sqrt{t}}\right)$$

That is

$$t\times\ell n\, t - 23.22t + 19230 = 0$$

An iterative solution gives t=1191.7s or approximately 19.9 minutes.

♣

Example 1.4: ***Constant Total Dopant diffusion – drive in - 2***

An arsenic constant-dose diffusion is performed with an initial dose of S=10^{14} cm^{-2}. The diffusion temperature is 1100°C for 2 hours. The starting wafer had a p-type substrate background doping of 10^{17} cm^{-3}. Find the concentration of the As at the surface and find the junction depth.

Solution

From Table 1.2

$$D = D_o e^{-\frac{E_a}{kT}} = 9.17e^{-\frac{3.99}{8.614\times10^{-5}\times1100+273}} = 2.07\times10^{-14}\,cm^2/s$$

Then the diffusion length is

$$\sqrt{Dt} = \sqrt{2.07\times10^{-14}\times7200} = 1.22\times10^{-5}\,cm$$

The surface concentration is

$$\left.\frac{dN}{dx}\right|_{x=0} = N_o = \frac{N_s}{\sqrt{\pi Dt}} = \frac{10^{18}}{\sqrt{\pi}\times1.22\times10^{-5}} = 4.6\times10^{18}\,cm^{-3}$$

From equation (1.13) rearranged, the junction depth for Gaussian diffusions is

$$x_j = 2\sqrt{Dt}\;\ell n\left(\frac{N_o}{N_B}\right)^{½}$$

$$= 2\times1.22\times10^{-5}\,cm\times\ell n\left(\frac{4.6\times10^{18}\,cm^{-3}}{10^{17}\,cm^{-3}}\right)^{½}$$

$$= 0.467\mu m$$

♣

1.1.3 Epitaxy growth - deposition

Epitaxy or epitaxial growth is the process of depositing a non-volatile, thin solid layer typically 0.5 to 100 μm, of single crystal material over a single crystal substrate.

Chemical Vapour Deposition (CVD)

Epitaxial growth is usually achieved using chemical vapour deposition (CVD). (Specifically, Metal-Organic Chemical Vapour Deposition, MOCVD or Metal-Organic Vapour Phase Epitaxy, MOVPE) The semiconductor deposited film is often the same material as the substrate, and the process is known as homoepitaxy, or simply, epi, as with silicon deposition on a silicon substrate. If the substrate is an

ordered semiconductor crystal (that is mono-silicon, gallium arsenide), the process continues building on the substrate with the same crystallographic orientation, with the substrate acting as a seed for the deposition. If an amorphous/polycrystalline substrate surface is used, the film will also be amorphous or polycrystalline. A key feature of epitaxy is that a lightly doped layer of epitaxial silicon can be grown upon a heavily doped silicon substrate, thus creating a layer of differing conductivity that can serve as an insulating layer or intrinsic buffer region.

Chemical vapour deposition CVD (see section 1.2.1) of silicon epitaxy occurs in an epitaxial reactor that consists of a quartz induction heated reaction chamber into which a susceptor is placed. The susceptor has two functions:

- mechanical support for the wafers and
- an environment with uniform thermal distribution.

The technological method of introducing reactant gases with only the substrates heated inside a reactor is called Vapour Phase Epitaxy, a schematic of which is shown in figure 1.5.

A possible fabrication process is as follows. A pre-cleaned, polished, near perfect silicon crystal surface acts as a substrate for subsequent deposition. Usually hydrogen chloride is used to etch the wafers. The pre-doped silicon is heated to about 1150°C in a quartz reactor tube at atmospheric pressure. A hydrogen gas flow carrying a compound of silicon such as silicon tetrachloride $SiC\ell_4$ or silane SiH_4 is passed over the hot substrate surface, and silicon atoms are deposited, growing a new continuous lattice. If phosphine (PH_3) arsine (AsH_3) or diborane (B_2H_6) is included in the silicon compound carrier gas flow of H_2 and N_2, a layer of the required doping type and resistivity occurs. Up to 100μm of doped silicon can be grown on substrates for power devices at a high growth rate of about 1 μm/min at 1200°C. A ultra low crystalline fault rate is essential if uniform electrical properties are to be attained. Selective deposition, depending on the substrate surface masking, is possible.

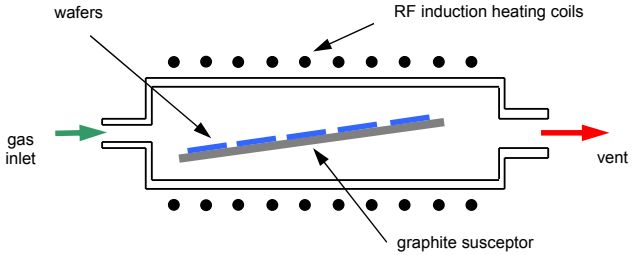


Figure 1.5. *Typical cold-wall vapour phase epitaxial reactor.*

There are four major chemical sources of silicon for epitaxial deposition:

- silane, SiH_4, as per equation (1.19)
- silicon tetrachloride, $SiC\ell_4$, as per equation (1.22);
- trichlorosilane, $SiHC\ell_3$, as per equation (1.20); and
- dichlorosilane, $SiH_2C\ell_2$, as per equation (1.21).

Chemical reaction equations can describe the growth of epitaxial layers. Each of the chemical sources mentioned can be described by an over-all reaction equation that shows how the vapour phase reactants form the silicon epitaxial film. For example, the over-all pyrolytic reaction for silicon epitaxy by silane decomposition reaction is:

$$SiH_4 \rightarrow Si + 2H_2 \qquad (\text{at } 1000°C \text{ to } 1100°C) \qquad (1.19)$$

Hydrogen reduction of trichlorosilane is

$$SiHC\ell_3 + H_2 \rightarrow Si + 3HC\ell \qquad (1.20)$$

Reduction of dichlorosilane is

$$SiH_2C\ell_2 \rightarrow Si + 2HC\ell \qquad (1.21)$$

However, such over-all reaction equations do not describe the complete CVD process as to how the gas phase reactants interact or how the epi species are adsorbed on the substrate surface. For instance, the over-all reaction for the hydrogen reduction of silicon tetrachloride $SiC\ell_4$ to form a silicon epitaxial layer is as follows:

$$SiC\ell_4 + 2H_2 \rightarrow Si + 4\,HC\ell \qquad (\text{at } 1150°C \text{ to } 1300°C) \qquad (1.22)$$

However, intermediate chemical species such as $SiHC\ell_3$ and $SiH_2C\ell_2$ are present during the silicon epitaxial growth:

$$SiC\ell_4 + H_2 \leftrightarrow SiHC\ell_3 + HC\ell$$
$$SiHC\ell_3 + H_2 \leftrightarrow SiH_2C\ell_2 + HC\ell$$

$$SiH_2C\ell_2 \leftrightarrow SiC\ell_2 + H_2$$
$$SiHC\ell_3 \leftrightarrow SiC\ell_2 + HC\ell$$
$$SiC\ell_2 + H_2 \leftrightarrow Si + 2HC\ell$$

These equations illustrate that even if a given process is described by a single over-all reaction, the process is actually a combination of many simultaneous chemical reactions.

The growth rate of an epitaxial layer depends on several factors:

- the chemical sources;
- the deposition temperature; and
- the mole fraction of reactants.

Silicon epitaxy improves the performance of bipolar devices. By growing a lightly doped epi-layer over a heavily-doped silicon substrate, a higher breakdown voltage across the collector-substrate junction is achieved while maintaining low collector resistance. Lower collector resistance allows a higher operating speed with the same current. Epitaxy is also used in IC fabrication. By fabricating a CMOS device on a thin (3 to 7 microns) lightly doped epi layer grown over a heavily-doped substrate, latch-up occurrence is minimized – a phenomena applicable to power devices such as the MOSFET and IGBT.

As well as improving device performance, epitaxy also allows better control of doping concentrations within devices. The layer can also be made oxygen and carbon free. The disadvantages of epitaxy include higher cost of wafer fabrication, additional process complexities, and problems associated with growth defects in the epi-layer.

Molecular Beam Epitaxy (MBE)

A schematic diagram of an MBE machine is shown in Figure 1.6. Generally, such machines consists of three vacuum sections, of which the growth chamber is the most important. The buffer section is involved in the preparation and storage of the wafers before entering the growth chamber. The load lock is used to insert and remove samples while retaining vacuum integrity. Samples are loaded onto a rotational magnetic holder known as, Continual Azimuthal Rotation (CAR). Cyropanels are used in conjunction with the vacuum system to keep the partial pressure of undesirable gases such as CO_2 and H_2O around 10^{-11} Torr. The principle of operation is that gaseous substances are bound to the cold surfaces within the pump by means of cryocondensation, cryosorption or cryotrapping. Epitaxial growth starts with the many heated cells, called effusion cells or *Knudsen* cells that contain a compound of the particular atomic species to be added into the vacuum chamber. Each source is independently heated until atoms of the source material are able escape by thermionic emission. An advancement of MBE, Gas Source MBE (GSMBE), uses room temperature gases for the source materials, thus avoiding contamination problems and high substrate temperatures that can cause segregation of dopant atoms.

Within the ultra-high vacuum, the free atoms have a long mean-free path and collisions with other atoms are infrequent. Atoms from the sources are able to travel in a straight line until they collide with the substrate material. A computer remotely operates the shutter controls, allowing the emission of different atom species to be directed at the substrate. The typical rate of growth with MBE is around a single mono-layer per second. Although slow, this allows for abrupt changes in material composition. Under proper conditions, the beam of atoms will attach to the substrate material and an epitaxial layer will begin to form.

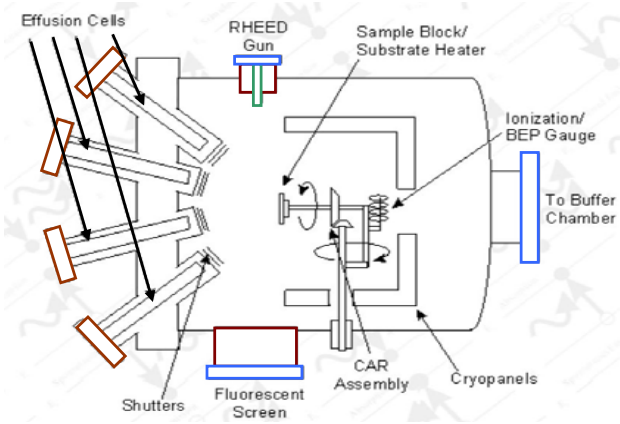


Figure 1.6. *Schematic diagram of MBE machine.*

Thickness control is determined using an ion gauge mounted facing the beam sources. The Beam Equivalent Pressure (BEP) of the material sources measures the rate of layer growth. Alternatively, Reflection High-Energy Electron Diffraction (RHEED) is another measuring technique. Electrons are emitted from an electron gun at a glancing incidence to the CAR. The reflected and diffracted electrons emit light in a distinctive pattern when striking a phosphor screen. Starting from a flat substrate, the electrons are not scattered greatly and are recorded as an intense beam. As material is deposited on the surface, the atoms create islands of epitaxial growth with an associated material surface reflectivity decrease as the electrons are scattered. As the deposition process continues, material builds up on the surface, the islands join and create new flat surfaces with the original substrate becoming voids in the newly created material. As further material forms, the voids begin to fill and the reflectivity increases once again, not reaching the initial value, since the process of deposition is random and the surface never regains the flat profile of the initial polished substrate. By monitoring the oscillations in the reflectivity, thickness and growth rate of the epitaxial material can be estimated.

1.1.4 Ion-implantation and damage annealing

Ion Implantation is the process of depositing chemical dopant species (atoms stripped of electrons) into a substrate by directly bombarding the substrate with high-energy ions of the chemical being deposited, as shown in figure 1.7.

Diffusion and ion implant are the two major processes by which chemical species or dopants are introduced into a semiconductor such as silicon to form electronic structures. The advantage of ion implant over diffusion is its more precise control for depositing dopant atoms into the substrate (10^{11} to 10^{18} cm^{-2}), giving excellent doping level uniformity and production repeatability.

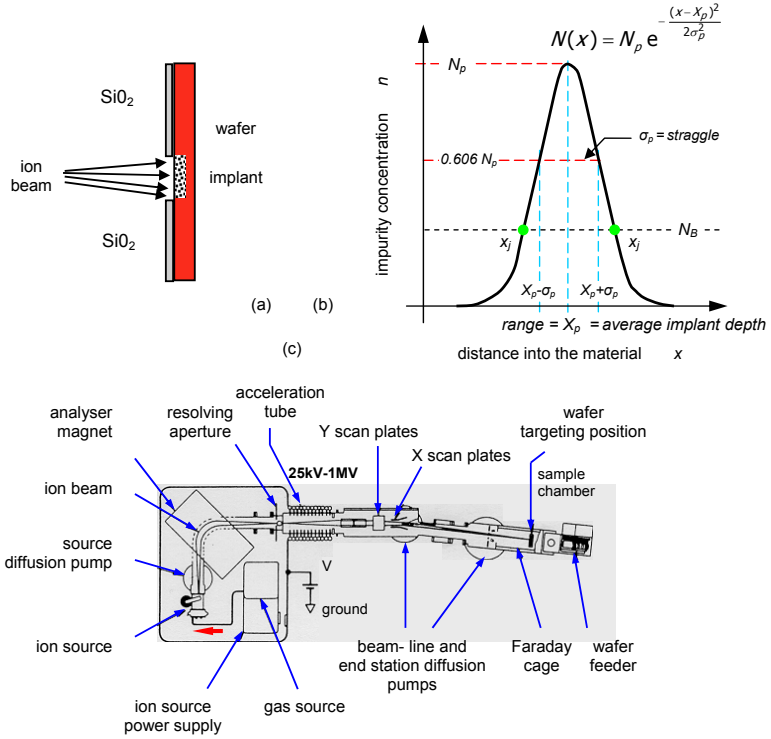


Figure 1.7. *Ion implantation:*
(a) pictorial representation of mechanism; (b) implanted ion distribution; and (c) implanting system.

The implanted profile shown in figure 1.7b, where two junctions may be formed, can be approximated by a Gaussian distribution function, which is similar to equation (1.13):

$$N(x) = \frac{S}{\sqrt{2\pi}\,\sigma_p} e^{-\frac{(x-X_p)^2}{2\sigma_p^2}} = N_p\, e^{-\frac{(x-X_p)^2}{2\sigma_p^2}} \qquad (1.23)$$

where S = ion dose per unit area, cm^{-2}

σ_p = symmetrical standard deviation, *straggle*, in the projected range of the implanted ions, cm

The depth of average or mean projected *range* (peak) is at X_p along the axis of incidence, where the maximum concentration occurs.

$$S = \frac{1}{q}\int_o^t I_{beam}(t')\,dt' = \frac{\frac{\text{ion beam current (A)}}{q} \times \text{implant time}}{\text{implant area}} \qquad (1.24)$$

The point where the diffused impurity profiles intersects the background concentration N_B is the metallurgical junction depth, x_j, where the net impurity concentration is zero. From equation (1.23)

$$N_B = N(x) = N_p\, e^{-\frac{(x_j-X_p)^2}{2\sigma_p^2}} \quad \text{re-arranged gives}$$

$$x_j = X_p \pm \sigma_p\sqrt{2\,\ell n\frac{N_p}{N_B}} \qquad (1.25)$$

where $N_p = S / \sqrt{2\pi}\,\sigma_p$ is the peak concentration.

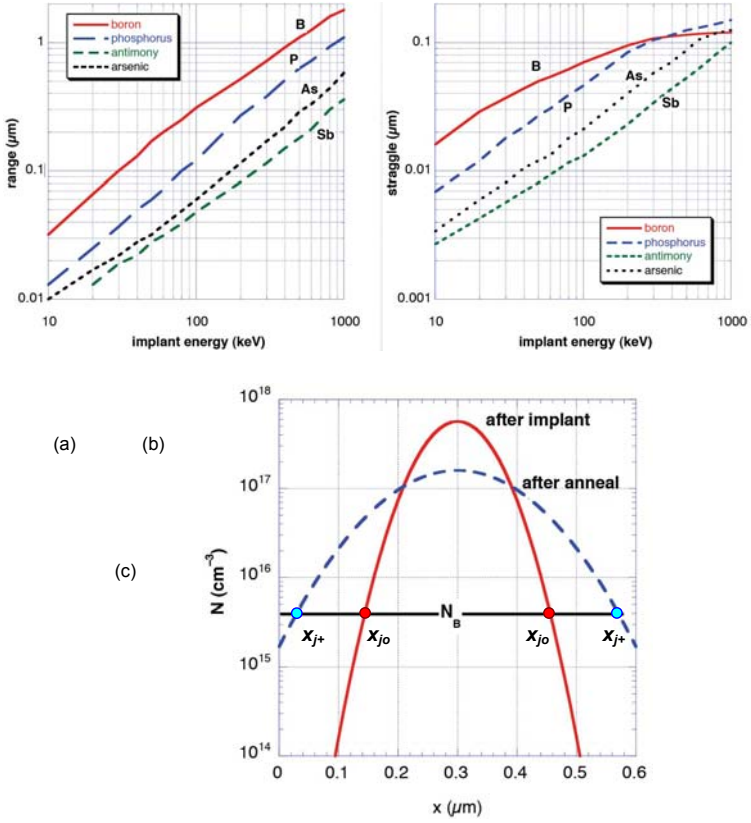


Figure 1.8. *Ion implanted silicon: (a) range X_p dependency on energy;*
(b) straggle σ_p dependency on energy; and (c) typical junction change after annealing.

Doping, which is the primary purpose of ion implanting, is used to alter the type and level of conductivity of semiconductor materials. It is used to form bases, emitters, and resistors in bipolar devices, as well as drains and sources in MOS devices. It is also used to dope polysilicon layers.
Typically, a gaseous dopant is ionized by electric discharge or by heat from a hot filament. The ions are separated using an electromagnetic field that bends the positively-charged particles to a selected band. This ion band is then passed through a high-current accelerator. The high-velocity beam of ions is focused on the wafer, causing the dopant ions to strike the wafer surface and penetrate. Sometimes a mask is used to implant a designated pattern on the wafer. As with diffusion, ion implantation allows the formation of junctions by changing the conductivity characteristics of precise regions in the wafer.
The basic procedure for ion implantation into silicon is as follows:

Ion impurities (B, P or As) are vaporised and accelerated by an electric field in a vacuum at high keV energies at the pre-doped silicon substrate, which is at room temperature. The ions penetrate the lattice to less than a few microns, typically 1μm at about ½ MeV. The resultant implanted doping profile is Gaussian, with the smaller ion like boron, penetrating deeper.

These high-energy atoms enter the crystal lattice and lose their energy by colliding with silicon atoms before finally coming to rest at some depth. Adjusting the acceleration energy controls the average deposition depth of the impurity atoms. Heat treatment is subsequently used to anneal or repair the crystal lattice disturbances caused by the atomic collisions.

Every implanted ion collides with several target atoms before it comes to rest. Such collisions may involve the nucleus of the target atom or one of its electrons. The total power of a target to stop an ion, or its total stopping power S_T, is the sum of the stopping power of the nucleus and the stopping power of the electron. Stopping power is described as the energy loss of the ion per unit path length of the ion.
Implantation energies are typically 10keV to 1MeV, giving ion distributions with depths of 10 nm to 10 μm from doses vary from 10^{12} ions/cm^2 for threshold voltage adjustment in MOSFETs to 10^{18} ions/cm^2 for the formation of buried insulating layers. Figures 1.8a and 1.8b show implant energy characteristics of typical dopants in silicon.
The damage caused by atomic collisions and bombardment during high-energy ion implantation changes the material structure therefore electrical characteristics of the target substrate. Many target atoms are displaced, creating deep electron and hole traps which capture mobile carriers and increase resistivity. *Annealing* is therefore needed to repair the lattice damage and put dopant atoms in substitutional sites where they can be electrically active again.
Annealing can be considered as apart of a drive-in Gaussian diffusion process, where the effective *Dt* of the implant process is $\frac{1}{2}\sigma_p^2$, which produces a profile following annealing given by

$$N(x) = \frac{S}{\sqrt{2\pi\left(\sigma_p^2 + 2Dt\right)}} e^{-\frac{(x-X_p)^2}{2\sigma_p^2+4Dt}} = N_p^* \times e^{-\frac{(x-X_p)^2}{2\sigma_p^2+4Dt}} \qquad (1.26)$$

The peak concentration N_p^* decreases but its depth is unchanged, and the deviation increases, as illustrated in figure 1.8c.

Silicon damage caused by ion implantation includes:

- the formation of crystal defects such as Frenkel defects, vacancies, di-vacancies, higher-order vacancies, and interstitials;
- the creation of local zones of amorphous material within the supposedly crystalline structure
- formation of continuous amorphous layers as the localized amorphous regions grow and overlap; and
- microscopic sputtering and directional non-linear ion channelling.

The first two damage types are categorized as 'primary crystalline damage'. Restoring the ion-implanted substrate to its pre-implant condition requires the substrate being subjected to a reparative thermal process known as *annealing*.
Ion implantation damage annealing has five major components:

- electrical activation of the implanted impurities;
- primary crystalline damage annealing;
- annealing of continuous amorphous layers;
- dynamic annealing; and
- diffusion of implanted impurities.

Annealing is conducted in a neutral environment, such as in Ar or a N_2 atmosphere in a stack furnace.

Electrical activation of the implanted impurities refers to the process of increasing the electrical activity of newly implanted impurity atoms during annealing, which usually do not occupy substitutional sites after being implanted. Temperatures up to 500°C remove trapping defects, releasing carriers to the valence or conduction bands in the process. Electrical activity decreases again at 500 to 600°C, because of the formation of dislocations. Beyond 600°C, electrical activation increases until a peak at 800 to 1000°C.

In summary, primary crystalline damage annealing consists of:

- recombination of vacancies and self-interstitials in the low temperature range, up to 500°C;
- formation of dislocations at 500 to 600°C which can capture impurity atoms; and
- dissolution of these dislocations at 900 to 1000°C.

Annealing of the continuous amorphous layers that extend to the surface occur by solid-phase epitaxy between 500 to 600°C. The crystalline substrate beneath the amorphous layers initiates the recrystallization of the amorphous layers, with the regrowth proceeding towards the substrate surface. Factors affecting the recrystallization rate include crystal orientation and the implanted impurities. Amorphous layers that do not extend to the surface anneal differently, with the solid-phase epitaxy occurring at both amorphous-single crystal interfaces and the regrowth interfaces meeting below the surface.

Dynamic annealing effects refers to the healing of implant damage while the implantation process is occurring. This takes place because the heat applied to the wafer during implantation makes the point defects more mobile.

Diffusion of implanted impurities relates to the mass transport of implanted species across a concentration gradient within an implanted layer during the annealing process. The presence of implant damage makes this diffusion process more complex than what occurs in an undamaged single-crystal substrate. Diffusion of implanted impurities during annealing degrades devices that have shallow junctions or narrow base and emitter regions if the thermal processing is not rapid enough, particularly in the case of boron ion implantation.

Example 1.5: Ion implantation

For a 100 keV boron implant with a dose of $S = 5\times10^{14}$ cm^{-2}, calculate

i. the peak concentration,
ii. the junction depth, if the substrate phosphorus background doping level is 10^{15} /cm^3, and
iii. the surface concentration.
iv. the junction depths after annealing for 30 minutes at 1000°C.

Solution

From figure 1.8 parta a and b, for a 100keV boron implant, the peak concentration occurs at a depth (range) of $X_p = 0.31$ μm and the ion implant standard deviation (straggle) is $\sigma_p = 0.07$ μm.

i. From equation (1.23)

$$N(x) = \frac{S}{\sqrt{2\pi}\,\sigma_p} e^{-\frac{(x-X_p)^2}{2\sigma_p^2}}$$

Differentiation gives

$$\frac{dn}{dx} = -\frac{S}{\sqrt{2\pi}\sigma_p}\frac{2(x-X_p)}{2\sigma_p^2} e^{-\frac{(x-X_p)^2}{2\sigma_p^2}} = 0$$

which confirms that the maximum concentration occurs when $x = X_p$.
Substitution into equation (1.23) gives the concentration $N(x = X_p = 0.31\ \mu m) = 2.85\times10^{18}$ cm^{-3}.

ii. The junction depth, with a background doping level of 10^{15} /cm^3, is given by equation (1.25), that is

$$x_j = X_p \pm \sigma_p\sqrt{2\ell n\frac{N_p}{N_B}}$$

$$= 0.31 \pm 0.07\sqrt{2\ell n\frac{2.85\times10^{18}}{10^{15}}} = 0.31 \pm 0.28\mu m$$

Two junctions are formed, at 0.03 μm and 0.59 μm below the incident surface.

iii. Since the ion implant has formed two junctions within the n-substrate, the surface concentration is dominated by the background doping level of the substrate, 10^{15} /cm^3. The surface ion implant doping is given by equation (1.23)

$$N(x=0) = \frac{S}{\sqrt{2\pi}\sigma_p} e^{-\frac{X_p^2}{2\sigma_p^2}} = 2.85\times10^{18} \times e^{-\frac{0.31^2}{2\times0.07^2}} = 1.57\times10^{14}/\text{cm}^3$$

The n-type surface concentration is 10^{15} /cm^3 – 1.57×10^{14} /cm^3 = 8.43×10^{14} /cm^3.

iv. For heat treatment, from example 1.2, $Dt = 1.39 \times 10^{-14}\,cm^2/s \times 30x60s = 2.5 \times 10^{-11}\,cm^2$. Equating equation (1.26) to the background concentration $10^{15}/cm^3$

$$N_B = \frac{S}{\sqrt{2\pi\left(\sigma_p^2 + 2Dt\right)}} e^{-\frac{(x-X_p)^2}{2\sigma_p^2+4Dt}}$$

$$10^{15}\,cm^{-3} = \frac{5\times10^{14}\,cm^{-2}}{\sqrt{2\pi\left(\left(7\times10^{-6}\,cm\right)^2 + 2\times2.5\times10^{-11}\,cm^2\right)}} e^{-\frac{(x-31\mu m)^2}{2\times\left(\left(7\times10^{-6}\,cm\right)^2+2\times2.5\times10^{-11}\,cm^2\right)}}$$

yields junction depths

$$x_j = 0.31\mu m \pm 0.44\mu m$$
$$= 0.75\mu m$$

The implant after annealing reaches the surface, resulting in one p-n junction 0.75um below the surface. The peak concentration N_p^* at 0.31µm is $2 \times 10^{18}\,cm^{-3}$.

♣

1.2 Thin Film Deposition

A thin film is a layer with a high surface area-to-volume ratio. Thin films are extensively used to apply dopants and sealants to wafers and microelectronic parts, and can be a resistor, a conductor, an insulator or a semiconductor. Thin films can be deposited with a thickness of between a few nanometres to about 100µm. The film can subsequently be locally etched using processes described in the Lithography and Etching sections of this chapter, sections 1.5 and 1.6, respectively.

Thins films behave differently from bulk materials of the same chemical composition in several ways. Thin films are sensitive to surface properties while bulk materials generally are not. Thin films are also more sensitive to thermo-mechanical stresses. Thin film integrity is influenced by the quality of its adhesion to and conformal coverage of the underlying layer, residual or intrinsic stresses after deposition, and the presence of surface imperfections such as pinholes.

The adhesion of a thin film to the substrate or underlying layer is paramount to ensuring thin film reliability. A thin film that is initially adhering to the underlying layer may lift off after the device is subjected to thermo-mechanical stresses. Reliable thin film adhesion depends on the cleanliness of the surface upon which it is deposited. Optimum substrate roughness also affects thin film adhesion. An ultra-smooth substrate decreases adhesion tendency. A rough substrate on the other hand can result in coating defects, which can also lead to thin film adhesion failures.

Regardless of the deposition process, thin films always have an intrinsic stress that can be either tensile or compressive. High residual stresses can lead to adhesion problems, corrosion, cracking, and deviations in electrical properties. Thus, proper deposition is critical to minimize intrinsic stresses in thin films.

Deposition technology is classified into two reaction types, viz. chemical and physical:

i. Depositions that result because of a chemical reaction:
- Chemical Vapour Deposition (CVD)
- Electrodeposition
- Epitaxy
- Thermal oxidation

These processes exploit the creation of solid materials directly from chemical reactions in gas and/or liquid compositions or with the substrate material. The solid material is usually not the only product formed by the reaction. By-products can include gases, liquids and other solids.

ii. Depositions that result because of a physical reaction:
- Physical Vapour Deposition (PVD)
- Casting

Common to these processes is that the material deposited is physically moved onto the substrate. In other words, there is no chemical reaction that forms the material on the substrate. This is not completely correct for casting processes, though it is more convenient to classify them as physical.

Whether the process is physical or chemical, the processing deposition reactor uses either:

- a *cold wall* system, where the heating process uses radio frequency or infra red heating, while
- a *hot wall* system uses a thermal heating resistive element or series of elements forming heating zones.

1.2.1 Chemical Vapour Deposition (CVD)

A fluid precursor undergoes a chemical change at a solid surface, leaving a solid layer.
In this process, the substrate is placed inside a reactor into which a number of gases are supplied, as shown in figure 1.9. The principle of the process is that a chemical reaction occurs between the source gases, with the solid material product of that reaction condensing on all surfaces inside the reactor.
CVD is capable of producing thick, dense, ductile, and good adhesive coatings on metals and non-metals such as glass and plastic. In contrast to PVD coating in the 'line of sight', CVD can simultaneously coat all surfaces of the substrate. The thin films from chemical deposition techniques tend to be conformal, rather than directional.

CVD processes are used to produce a thin film with good step coverage. A variety of materials can be deposited, however, some form hazardous by-products during processing. The quality of the material varies from process to process, however generally a higher process temperature yields a material with higher quality and fewer defects. They are generally not suitable for mixtures of materials. CVD processing is not possible for some materials; there is no suitable chemical reaction.

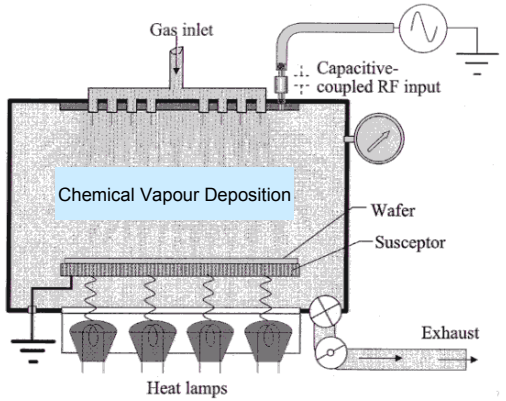


Figure 1.9. *Typical CVD processing reactor system.*

Chemical deposition is categorized by the phase of the precursor:

- Plating relies on liquid precursors, often a solution of water with a salt of the metal to be deposited. Some plating processes are driven only by reagents in the solution (usually for noble metals), but the most important process is electroplating. Although not commonly in semi-conductor processing, it has resurfaced with the use of chemical-mechanical polishing.

  1. Conventional CVD coating processing requires a metal compound that will volatilize at a low temperature and decompose to a metal when it contacts the substrate at higher temperature. An example of CVD is nickel carbonyl ($NiC0_4$) coating, as thick as 2.5 mm, on glass windows and containers to make them explosion or shatter resistant.
  2. Diamond CVD coating processing is used to increase the surface hardness of cutting tools. The process is performed at temperatures higher than 700°C which softens most tool steels. Thus, the application of diamond CVD is limited to materials which do not soften at this temperature, such as cemented carbides.
  3. Plasma-assisted CVD coating processing is performed at lower temperature than diamond CVD coating. Diamond coatings or silicon carbide barrier coatings are applied on plastic films and semiconductors, including sub-¼µm semiconductors.

- *Chemical solution deposition* uses a liquid precursor, usually a solution of organometallic powders dissolved in an organic solvent. This is a relatively inexpensive, simple thin film process that is able to produce stoichiometrically accurate crystalline phases.

- Chemical vapour deposition generally uses a gas-phase precursor, often a halide or hydride of the element to be deposited. In the case of metal-organic CVD, an organometallic gas is used.

  The two most important CVD technologies are *Low Pressure CVD* (LPCVD) and *Plasma Enhanced CVD* (PECVD). The key features are:

  1. The LPCVD process produces layers with uniformity of thickness and material characteristics. The main processing problems are the high deposition temperature, greater than 600°C, and the relatively slow deposition rate. The PECVD process can operate at lower temperatures, down to 300° C, due to the extra energy supplied to the gas molecules by the ionised vapour precursor, or plasma in the reactor. However, the quality of the films tend to be inferior to processes running at higher temperatures. PECVD relies on electromagnetic means (electric current, microwave excitation), rather than a chemical reaction, to produce a plasma, as shown in figure 1.10.
  2. Most PECVD deposition systems can only deposit the material on one side of the wafers, on 1 to 4 wafers at a time. LPCVD systems deposit films on both sides of at least 25 wafers, simultaneously. A schematic diagram of a typical LPCVD reactor is shown in figure 1.11. PECVD films are conformal and deposited at lower temperatures than for LPCVD, although the film is not stoichiometric, prone to cracking and peeling, with by-products formed.

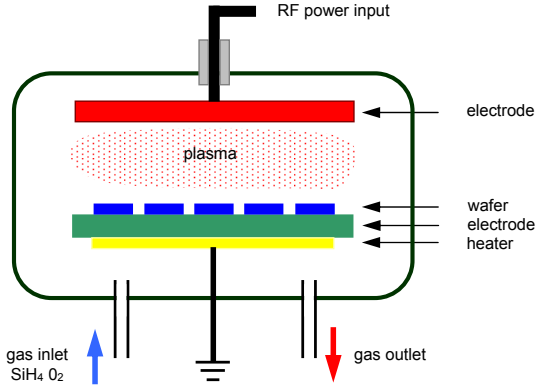


Figure 1.10. *Typical PECVD reactor.*

CVD is accomplished by placing the substrate wafers in a reactor chamber and heating them to a specific temperature. Controlled amounts of silicon or nitride source gases, usually carried by either nitrogen and/or hydrogen, are added to the reactor. Dopant gases may also be added if desired. A reaction between the source gases and the wafer occurs, thereby depositing the desired layer. Reaction temperatures between 500 to 1100ºC and pressures ranging from atmospheric to low pressure are used, depending on the specific deposition performed. Heating is usually accomplished with radio frequency, infrared or thermal resistance heating. Common source gases include silane SiH_4, silicon tetrachloride $SiC\ell_4$, ammonia NH_3, and nitrous oxide N_2O. Some dopant gases that are used include arsine AsH_3, phosphine PH_3, and diborane B_2H_6. The major categories of silicon CVD are shown by the following equations.

LPCVD Atmospheric or low pressure

Medium temperature (600°C to 1100°C)

Silicon Nitride, Si_3N_4: H_2 carrier gas (900°C to 1100°C)

$$3\ SiH_4 + 4\ NH_3 \rightarrow Si_3N_4 + 12\ H_2$$

Poly Silicon, Poly-Si, H_2 carrier gas (850 to 1000°C), N_2 carrier gas (600°C to 700°C)

$$SiH_4 + Heat \rightarrow Si + 2\ H_2$$

Silicon Dioxide, $Si0_2$:

$$SiH_4 + 4\ C0_2 \rightarrow Si0_2 + 4\ C0 + 2\ H_20 \quad N_2 \text{ carrier gas (500°C to 900°C)}$$

$$2\ H_2 + SiC\ell_4 + C0_2 \rightarrow Si0_2 + 4\ HC\ell \quad H_2 \text{ carrier gas (800°C to 1000°C)}$$

$$SiH_4 + C0 \rightarrow Si0_2 + 2\ H_2 \quad H_2 \text{ carrier gas (600°C to 900°C)}$$

Low Temperature (< 600°C)

Silicon Dioxide, $Si0_2$ or p-doped $Si0_2$, N_2 carrier gas (< 600°C)

$$SiH_4 + 2\ 0_2 + Dopant \rightarrow Si0_2 + 2\ H_20$$

Silicon Nitride, Si_3N_4, N_2 carrier gas (600°C to 700°C)

$$3\ SiH_4 + 4\ NH_3 \rightarrow Si_3N_4 + 12\ H_2$$

$$3\ SiH_4 + 2\ N_2O \rightarrow Si_3N_4 + 4\ H_2 + 2\ H_2O$$

PECVD Low Temperature Plasma Enhance (passivation) (< 600°C), RF or reactive sputtering

Silicon Dioxide, $Si0_2$:

$$SiH_4 + 2\ 0_2 \rightarrow Si0_2 + 2\ H_20 \qquad SiH_4 + 0_2 \rightarrow Si0_2 + 2\ H_2$$

Silicon Nitride, Si_3N_4:

$$3\ SiH_4 + 4\ NH_3 \rightarrow Si_3N_4 + 12\ H_2$$

$$3\ SiH_4 + 2\ N_20 \rightarrow Si_3N_4 + 4\ H_2 + 2\ H_20$$

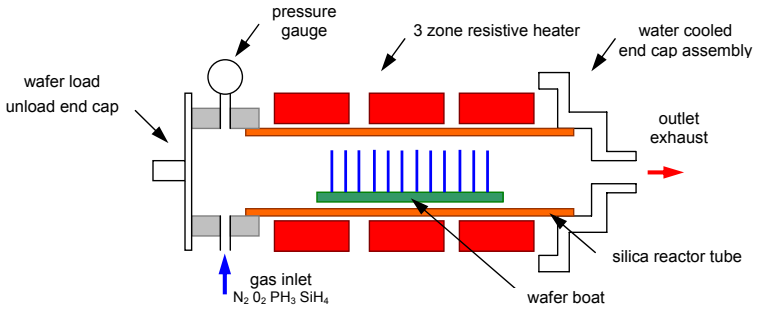


Figure 1.11. *Typical horizontal hot-wall LPCVD reactor.*

1.2.2 *Physical Vapour Deposition (PVD)*

Physical deposition uses mechanical or thermodynamic means to produce a thin film of solid. Physical deposition covers a number of deposition technologies in which material is released from a source and transferred to the substrate. Physical deposition coatings involve atom-by-atom, molecule-by-molecule or ion deposition of various materials on solid substrates in vacuum systems. Since most engineering materials bonded together by relatively high energies, and chemical reactions are not used to store these energies, physical deposition systems tend to require a low-pressure vapour environment to function properly.

The material to be deposited is placed in an energetic, entropic environment, so that particles of material escape its surface. Adjacent to this source is a cooler facing surface which draws energy from these particles as they arrive, allowing them to form a solid layer. The system process is in a vacuum deposition chamber, to allow the particles to travel unhindered. Since particles tend to follow a straight trajectory, films deposited by physical means are commonly directional, rather than conformal.

PVD comprises the standard technologies for deposition of metals. It is more common than CVD for metals since it can be performed with lower process risk and cheaper materials costs. The film quality is inferior to CVD, which for metals means higher resistivity and for insulators more defects and traps. The step coverage is also not as good as with CVD.

The choice of deposition method (specifically evaporation versus sputtering) may be arbitrary, and may depend more on what technology is available for the specific material.

Physical deposition includes:

- A *thermal evaporator* uses an electric resistance heater to melt the material and raise its vapour pressure to a useful range, where it starts to boil and evaporate. An atomic cloud is formed by the evaporation of the coating metal in a vacuum environment to coat all the surfaces in the line of sight between the substrate and the facing target (source). The vacuum allows the vapour to

reach the substrate without reacting with or scattering against other gas-phase atoms in the chamber, and reduces the absorption of impurities from the residual gas in the vacuum chamber. Only materials with a higher vapour pressure than the heating element can be deposited without contamination of the film. The method is often used in producing thin, ½ µm, decorative shiny coatings on plastic parts. The thin coating, however, is fragile and wears poorly. The thermal evaporation process can also coat a thick, 1 mm, layer of heat-resistant materials, such as MCrAℓY - a metal, chromium, aluminium, and yttrium alloy, on jet engine parts. Molecular beam epitaxy is a particular sophisticated form of thermal evaporation. A schematic diagram of a typical system for e-beam evaporation is shown in figure 1.12.

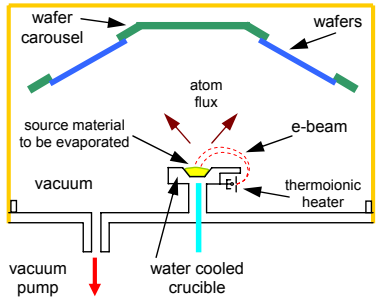


Figure 1.12. *Typical system for e-beam evaporation of materials.*

The principle is the same for all evaporation technologies, only the method used to the heat (evaporate) the source material differs. There are two main evaporation technologies, namely e-beam evaporation and resistive evaporation, each referring to the heating method.

- An *electron beam evaporator* fires a high-energy beam from an electron gun to boil a small spot of material; since the heating is not uniform but local, lower vapour pressure materials can be deposited. The beam is usually bent through a 270° angle in order to ensure that the gun filament is not directly exposed to the evaporant flux. Typical deposition rates for electron beam evaporation range from 1 to 10 nm/s.
- In *resistive evaporation*, a tungsten boat, containing the source material, is heated electrically with a high current to make the material evaporate. Many materials are restrictive in terms of what evaporation method can be used (that is, aluminium is quite difficult to evaporate using resistive heating), which typically relates to the phase transition properties of that material.

- *Sputtering* relies on a plasma (usually a noble gas, such as Argon) to knock material from a *target* or *source*, a few atoms at a time, at a much lower temperature than with evaporation. The coatings, such as ceramics, metal alloys, organic and inorganic compounds, involve connecting the work-piece and the substance to a high-voltage dc power supply in an argon-vacuum system at 10^{-2} to 10^{-3} mmHg. The gas plasma is established between the substrate (work-piece) and the target (donor) and transposes the sputtered-off target ionised atoms to the surface of the substrate. Because the target is kept at a relatively low temperature, unlike evaporation, this is a flexible deposition technique. It is especially useful for compounds or mixtures, where different components would otherwise tend to evaporate at different rates. Sputtering's step coverage is virtually conformal, producing thin, less than 3 µm, hard thin-film coatings; for example, titanium nitride (TiN) which is harder than the hardest metal. Sputtering is applied on cutting tools, injection moulding tools, and common tools such as punches and dies, to increase wear resistance and service life. When the substrate is non-conductive, for example, a polymer, radio-frequency (RF) sputtering is used. A schematic diagram of a typical RF sputtering system is shown in figure 1.13a. As for evaporation, the same basic principle applies to all sputtering technologies. The differences typically relate to the manner in which the ion bombardment of the target is realized. Magnetron splutter disposition, figure 1.13b, is used to deposit Aℓ, titanium, and tungsten, although CVD is difficult for alloys, Aℓ-Cu-Si.
- *Pulsed laser deposition* systems work by an ablation process. Pulses of focused laser light vaporize the surface of the target material and convert it to plasma; this plasma usually reverts to a gas before it reaches the substrate.
- *Cathodic Arc Deposition* or *Arc-PVD* is a kind of ion beam deposition where an electrical arc is created that blasts ions from the cathode. The arc has a high power density resulting in a high

level of ionization (30 to 100%), multiply charged ions, neutral particles, clusters, and macro-particles (droplets). If a reactive gas is introduced during the evaporation process, dissociation, ionization, and excitation occur during interaction with the ion flux and a compound film is deposited.

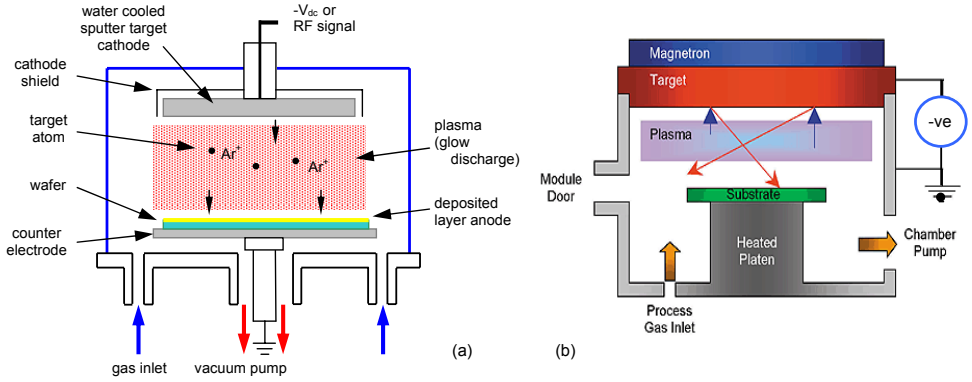


Figure 1.13. *Typical sputtering systems: (a) RF (ac) plasma and (b) dc plasma magnetron.*

1.3 Thermal oxidation and the masking process

The oxide of silicon, silica, or ***silicon dioxide*** (SiO_2), is an important planar processing ingredient and a extensively used dielectric in semiconductor manufacturing because it facilitates stable insulation and conformal passivation layers, with a high electric field breakdown strength of 10MV/cm, a resistivity of up to 10^{20} Ωcm, and a 9eV energy band gap. It is a useful and convenient deposition process, used many times on the silicon wafer surface during device fabrication. Besides the passivation and glass layer deposited over the surface of the die to protect it from mechanical damage and corrosion, dielectric layers are also used for isolating components or structures in the active circuit from each other, and as dielectric structures for MOS transistors and capacitors. Silicon dioxide is used as an insulating barrier between the gate metal and channel of insulated gate semiconductor power switching devices.

The formation of SiO_2 on a silicon surface is accomplished through a process called thermal oxidation, which is a technique that uses high temperatures, usually between 700°C to 1300°C, to promote the growth rate of oxide layers. The thermal oxidation of SiO_2 consists of exposing the silicon substrate to a rich oxidizing atmosphere of oxidisers, O_2 or H_2O, at elevated temperature, producing oxide films with thicknesses ranging from 6nm to 1µm. Oxidation of silicon is not difficult, since silicon naturally tends to form a stable oxide even at room temperature, provided an oxidizing environment is present. The elevated temperature used in thermal oxidation accelerates the oxidation process, resulting in thicker oxide layers per unit of time. This process affords thickness and property control of the SiO_2 layer.

The silicon wafers placed in a furnace containing oxygen gas for three to four hours at 1000°C to 1200°C, form a surface oxide layer of SiO_2, usually less than 1µm thick. Wet oxidation, with water added, is about 20 times faster (100nm to 120nm per hour) than dry oxidation (14nm to 25nm per hour) but the oxide quality is lower. The wafer is effectively encapsulated by silica glass, which prevents penetration by impurity atoms, except gallium atoms. Selective diffusions or implanting are made in the silicon by opening windows through the oxide by selective etching with hydrofluoric HF acid following a photo-resist lithography masking process - see section 1.5.

The oxidation furnace (or diffusion furnace, since oxidation is a diffusion process involving oxidant species), provides the heat needed to elevate the oxidizing temperature and the typical furnace consists of:

- a heating system;
- a temperature measurement and control system;
- fused quartz process tubes where the wafers undergo oxidation;
- a system for moving process gases into and out of the process tubes; and
- a loading station used for loading (or unloading) wafers into (or from) the process tubes.

The heating system consists of several heating coils that control the temperature around the furnace tubes. The wafers are placed in quartz glassware called boats, which are supported by fused silica paddles inside the process tube. A boat can contain many wafers, typically 50 or more. The oxidizing agent (oxygen or steam) then enters the process tube through its source end, subsequently diffusing to the wafers where the oxidation occurs. A schematic diagram of a typical wafer oxidation furnace is shown in figure 1.14.

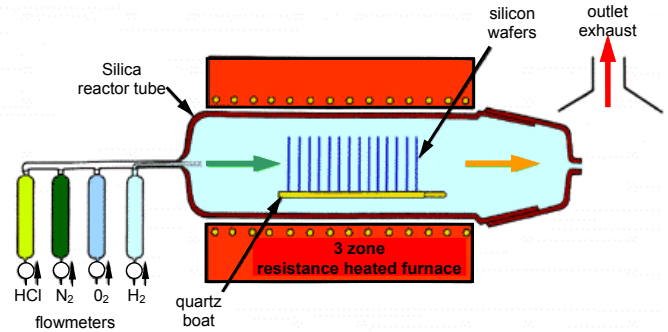


Figure 1.14. *Typical wafer oxidation furnace.*

Depending on the oxidant species used, namely 0_2 or H_20, the thermal oxidation of $Si0_2$ may either be in the form of dry oxidation, wherein the oxidant is 0_2 or wet oxidation, wherein the oxidant is H_20. The reactions for dry and wet oxidation are characterised by:

In *dry oxidation*, where the oxidising agent is oxygen, 0_2:

$$Si_{(s)} + O_{2(g)} \rightarrow SiO_{2(s)} \tag{1.27}$$

During dry oxidation, the silicon wafer reacts with the ambient oxygen (and hydrogen chloride at near atmospheric pressure), forming a layer of silicon dioxide on its surface, usually less than 100nm thick.

In *wet oxidation*, used for thick oxides, where the oxidising agent is water vapour, H_20:

$$Si_{(s)} + 2H_2O_{(g)} \rightarrow SiO_{2(s)} + 2H_{2(g)} \tag{1.28}$$

Hydrogen and oxygen gases are introduced into a torch chamber where they react to form water molecules, which are then made to enter the reactor under high pressure where they diffuse toward the wafers. The water molecules react with the silicon to produce the oxide and hydrogen gas byproduct.

The oxidation reactions occur at the Si-$Si0_2$ interface, that is, silicon at the interface is consumed as oxidation takes place. As the oxide grows, the Si-$Si0_2$ interface moves into the silicon substrate. Consequently, the Si-$Si0_2$ interface will always be below the original Si wafer surface. $Si0_2$ formation therefore proceeds in two directions relative to the original wafer surface. Oxidation is the only deposition technology which actually consumes some of the substrate as it proceeds. The amount of silicon consumed by silicon dioxide formation is predictable from the relative densities and molecular weights of Si and $Si0_2$. The thickness of silicon consumed is 44% of the final thickness of the oxide formed, thus an oxide that is 100nm thick will consume about 44nm of silicon from the substrate.

Oxidation processes that have short durations (and during the first 50nm of oxide growth), may be modelled by a Linear Growth Law equation: $x_o = C\times(t + \tau)$, where x_o is the growing oxide thickness, C is the linear rate constant, t is the oxidation time, and τ is the initial time displacement to account for the initial oxide layer in situ at the start of the oxidation process. As the process proceeds the oxide growth rate decreases.

For oxidation processes that have long durations (where the oxide thickness reaches 100nm), the rate of oxide formation may be modelled by a Parabolic Growth Law equation: $x_o^2 = B\times t$, where x_o is the growing oxide thickness, B is the parabolic rate constant, and t is the oxidation time. This equation shows that the oxide thickness grown is proportional to the square root of the oxidizing time, which confirms that oxide growth is hampered as the oxide thickness increases. This is because the oxidizing species diffusion rate decreases as it has to travel a greater distance through the oxide to the Si-$Si0_2$ interface as the oxide layer thickens.

Together the linear and parabolic growth equations are known as the Linear-Parabolic Model. This oxide growth model is accurate over a wide temperature range (700°C to 1,300°C), oxide thicknesses (20nm to 2µm), and oxidant partial pressures (0.2 to 2.5 atmospheres). An increase in pressure increases the

oxide growth rate, but importantly, allows the temperature to be decreased for a given growth rate. For every 10 atmospheres of pressure, the temperature can be reduced by 30°C.

Oxide growth is accelerated by an increase in oxidation time, oxidation temperature or oxidation pressure. Other factors that affect thermal oxidation growth rate for $Si0_2$ include:

- the crystallographic orientation of the wafer;
- the wafer's doping level;
- the presence of halogen impurities in the gas phase;
- the presence of plasma during growth; and
- the presence of a photon flux during growth.

Uses for dielectric layers include:

- masking for diffusion and ion implant processes;
- diffusion from doped oxides;
- overcoating of doped films to prevent dopant loss and migration;
- gettering of impurities (see section 1.12); and
- mechanical and chemical protection.

There are other commonly-used dielectric and isolation materials besides $Si0_2$.
Silicon dioxide doped with phosphorus (commonly referred to as P-glass, phospho-silicate glass or PSG) is used because it inhibits sodium impurity diffusion and exhibits a smooth topography. Adding boron to PSG results in boro-phospho-silicate glass. BPSG, flows at lower temperatures than PSG; 850°C to 950°C for BPSG as opposed to 950°C to1100°C for PSG.
Polysilicon $Si0_2$ with enough oxygen content is also semi-insulating and is used in circuit and surface junction passivation. Alternately, silicon nitride is an excellent moisture barrier while stoichiometric silicon nitride is used in oxidation masks and for MOS gate dielectric. These dielectric layers are usually deposited by sputtering or chemical vapour deposition (CVD). The layer material deposited depends on the processing reactants.

The oxidising process is restricted to materials that can be oxidized, and only films that are oxides or nitrides of that material are possible. Silicon nitride, like silicon dioxide, is an amorphous insulating material that is an excellent moisture and contamination barrier, highly resistant to diffusion, not prone to delamination or cracking, and forms a progressive conformal layer on silicon. The oxidant is pure ammonia gas NH_3 or an ammonia plasma. Although superior to silicon dioxide, it has a much higher dielectric constant 7.5 as opposed to 3.85 for silicon dioxide, so is not favoured for power device insulated gate oxide structures. The disadvantages of silicon nitride are thermal related, namely higher processing temperatures are needed (950 to 1200°C) and the thermal expansion of silicon nitride is twice that of silicon dioxide. The relative properties of silicon dioxide, $Si0_2$, and silicon nitride, Si_3N_4, at 300K, are shown in Table 1.3.

LPCVD $\quad 3SiCl_2H_{2(g)} + 4NH_{3(g)} \rightarrow Si_3N_{4(s)} + 6H_{2(g)}$

Table 1.3: Properties of silicon dioxide (SiO_2) and silicon nitride (Si_3N_4) at 300K

Properties		$Si0_2$	Si_3N_4
Structure		amorphous	amorphous
Melting Point	°C	≈ 1600	≈ 1900
Density	g/cm^3	2.2	3.1
Refractive Index		1.46	2.05
Dielectric Constant		3.9	7.5
Dielectric Strength	V/cm	10^7	10^7
Infrared Absorption Band	µm	9.3	11.5 - 12.0
Energy Gap at 300K	eV	9	≈ 5.0
Linear Coefficient of Thermal Expansion, $\Delta L/L/\Delta T$	1/K x 10^{-6}	0.5	3.3-
Thermal Conductivity at 300 K λ	W/cm-K	0.014	0.030
DC Resistivity at 25°C	Ohm-cm	10^{14} - 10^{16}	≈ 10^{14}
DC Resistivity at 500°C	Ohm-cm	-	2 x 10^{13}
Etch Rate in buffered HF	nm/min	100	½ - 1

1.4 Polysilicon deposition

Polysilicon Deposition is the process of depositing a thin-film layer of polycrystalline silicon on a semiconductor wafer, similar to epi-deposition but processed at much lower temperatures.

Polysilicon, poly-Si is compatible with high temperature processing and interfaces with thermal $Si0_2$. One of its primary uses is as gate electrode material for metal-oxide type devices because it is more reliable than Aℓ. Polysilicon gate electrical conductivity may be enhanced by depositing a metal (such as tungsten) or a metal silicide (such as tungsten silicide) over the gate. It can also be deposited conformally over steep topography. Heavily-doped poly thin-films are used in bipolar emitter structures. Lightly-doped poly-Si film may be employed as a resistor, a conductor or as an ohmic contact for shallow junctions, with the desired electrical conductivity attained by doping the polysilicon material.

Polysilicon deposition is achieved by thermal decomposition, called pyrolysis of silane, SiH_4, inside a low-pressure reactor at a temperature of 580°C to 650°C, with the deposition rate exponentially increasing with temperature. This pyrolysis process involves the basic reaction: $SiH_{4\,(g)} \rightarrow Si_{(s)} + 2H_{2\,(g)}$.

There are two common low-pressure processes for depositing polysilicon layers:

- using 100% silane at a pressure of 25 Pa to 130 Pa; and
- using 20% to 30% silane (diluted in nitrogen) at the same total pressure.

Both of these processes can deposit polysilicon on 10 to 200 wafers per run, at a rate of 10 to 20 nm/min and with thickness uniformities of better than ± 5%. The critical process variables for polysilicon deposition include temperature, pressure, silane concentration, and dopant concentration. Wafer spacing and load size have only minor effects on the deposition process.

The rate of polysilicon deposition R_d increases rapidly with temperature, since it follows the Arrhenius equation:

$$R_d = R_{do} e^{-qE_a/kT} \tag{1.29}$$

where R_d is the deposition rate,
E_a is the activation energy in electron volts, eV,
T is the absolute temperature in degrees Kelvin, K,
k is the Boltzmann constant, q is the electron charge, and
R_{do} is a constant.
The activation energy for polysilicon deposition is about 1.7eV.

Based on equation (1.29), the rate of polysilicon deposition increases as the deposition temperature increases. There will be a minimum temperature at which the rate of deposition becomes faster than the rate at which unreacted silane arrives at the surface. Beyond this temperature, the deposition rate can no longer increase with temperature, since it is now starved of silane from which the polysilicon is being generated. Such a reaction is *mass-transport-limited*. When a polysilicon deposition process becomes mass-transport-limited, the reaction rate is dependent on reactant concentration, reactor geometry, and gas flow.

When the rate at which polysilicon deposition occurs is slower than the rate at which unreacted silane arrives, the deposition is *surface-reaction-limited*. A deposition process is then primarily dependent on reactant concentration and reaction temperature. Surface-reaction-limited processes result in excellent thickness uniformity and step coverage. The logarithm of deposition rate plotted against the reciprocal of the absolute temperature in the surface-reaction-limited region is a straight line with slope $-qE_a/k$.

At reduced pressures, polysilicon deposition below 575°C is too slow to be practical. Above 650°C, poor deposition uniformity and excessive roughness are encountered due to unwanted gas-phase reactions and silane depletion. Pressure can be varied inside a low-pressure reactor either by changing the pumping speed or changing the inlet gas flow into the reactor. If the inlet gas is composed of both silane and nitrogen, the inlet gas flow, and hence the reactor pressure, may be varied either by changing the nitrogen flow with a constant silane flow, or changing both the nitrogen and silane flow to change the total gas flow while keeping the gas ratio constant, equation (1.29).

Polysilicon doping, if needed, is also performed during the deposition process. The electrical characteristics of a poly-Si thin film depends on its doping. As in single-crystal silicon, heavier doping results in lower resistivity. Poly-Si is more resistive than single-crystal silicon for any given level of doping mainly because the grain boundaries in poly-Si hamper carrier mobility. Common dopants for polysilicon include arsenic, phosphorus, and boron. For better doping confinement, polysilicon is usually deposited undoped, with the dopants introduced after deposition.

There are three ways to dope polysilicon, namely, diffusion, ion implantation, and in-situ doping. *Diffusion doping* consists of depositing a heavily-doped silicon glass over the undoped polysilicon. This glass served as the source of dopant for the poly-Si. Dopant diffusion takes place at a high temperature

of 900°C to 1000°C and attains the lowest resistivities. Resistance is reduced if silicides (WSi_2, $TaSi_2$, $CoSi_2$, etc.) are used and minimised (for minimum *R-C* delay) if a metal gate is used for both the gate and interconnections. *Ion implanting* is more precise in terms of dopant concentration control and involves directly bombarding the poly-Si layer with high-energy ions. *In-situ doping* consists of adding dopant gases, such as phosphine, arsine or diborane, to the CVD reactant gases during the epi polysilicon deposition process. Adding phosphine or arsine results in slower deposition, while adding diborane increases the deposition rate. The deposition thickness uniformity degrades when dopants are added during deposition.

1.5. Lithography – optical and electron

The fabrication of features on silicon wafers requires that several different layers, each with a different specific pattern, be deposited on the surface sequentially, and that doping of the active regions be done in controlled amounts over small regions of precise areas. The various patterns used in depositing layers and doping regions on the substrate are defined by a process called *lithography*. One important aspect of lithography is *photoresist* processing, which is the process of covering areas that either need to be subsequently removed or retained with a light sensitive film - the photoresist. The process of material removal following a photolithographic process is known as *etching*.

Photoresist layers have two basic functions:

- precise pattern formation and
- protection of the substrate from chemical attack during the etch process.

Performance metrics

- Resolution: minimum feature dimension that can be transferred with fidelity to a resist film.
- Registration: how accurately patterns on successive masks can be aligned (or overlaid) with respect to previously defined patterns.
- Throughput: number of wafers that can be exposed/unit time for a given mask level.

Photoresist materials consist of three components:

- a matrix material (also known as resin), which provides body and binder for the photoresist;
- the inhibitor (also referred to as sensitizer), which is the photoactive ingredient; and
- the solvent, which maintains the resist liquid until it is applied to the substrate.

The lithography process consists of the following steps:

- Dehydration and priming;
- A layer of photoresist material is first spin-coated on the surface of the wafer;
- Soft baking;
- The resist layer is then selectively exposed to radiation such as ultraviolet light, electrons or X-rays, with the exposed areas defined by the exposure tool, mask or computer data.
- The photoresist layer is subjected to photo-development which removes unwanted areas of the resist-layer, exposing the corresponding areas of the underlying layer. Depending on the resist type, the development stage may remove either the exposed or unexposed areas. The areas with no resist material left on them are then subjected to additive or subtractive processes, allowing the selective deposition or removal of material on the substrate. Hard bake.
- Post-development inspection.

During development, unwanted areas in the photoresist are dissolved by the developer. When the exposed areas become soluble in the developer, a positive image of the mask pattern is produced on the resist. Such a resist is therefore called a *positive photoresist*. *Negative photoresist* layers result in negative images of the mask pattern, where the exposed areas are not soluble in the developer. Wafer fabrication may employ both positive and negative photoresists, although positive resists offer higher resolution capabilities.

1 Dehydration and priming

Prior to the application of resist to a wafer, the wafer must be free of moisture and contaminants, both of which cause a multitude of resist processing problems. Dehydration baking is performed to eliminate any moisture adsorbed by substrate surfaces, since hydrated substrates result in adhesion failures. The bake is usually performed at between 400°C to 800°C. Convection ovens are used for baking up to 400°C, while furnace tubes are used for baking up to 800°C. After dehydration baking, the wafer is coated with a pre-resist priming layer of hexamethyldisilazane with glycol-ether solvent designed to enhance the adhesion properties of the wafer. Resist coating must follow the priming, within an hour.

2 Spin Casting

Resist coating, or the process of producing a uniform, adherent, and defect-free resist film of the correct

thickness over the wafer, is usually performed by *spin-coating*. In the spin casting process the material to be deposited is dissolved in liquid form in a solvent. The viscous material (a polymer like UV positive resist ortho-diazoketone) is applied to the substrate centre by spraying or spinning. Most spin-coating processes reach speeds of 2000 to 7000 rpm for a duration of 20 to 60 seconds. The thickness that can be cast on a substrate ranges from a single monolayer of molecules, by adhesion promotion, to tens of micrometres. The control on film thickness, typically 350nm to 2µm for UV exposed photoresists, can be sustained within ±10%. Thickness of the photoresist is given by

$$t = c\, S_{\%s} \left(\frac{\nu}{\omega^2 r^2} \right)^{1/3} \qquad (1.30)$$

where t = thickness
c = constant
$S_{\%s}$ = fraction of solids
ν = viscosity
ω = angular velocity
r = radius

The spin casting process is illustrated in figure 1.15. Other materials such as polyimide and spin-on glass can be applied by casting. Once the solvent is evaporated, a thin film of the material remains on the substrate.

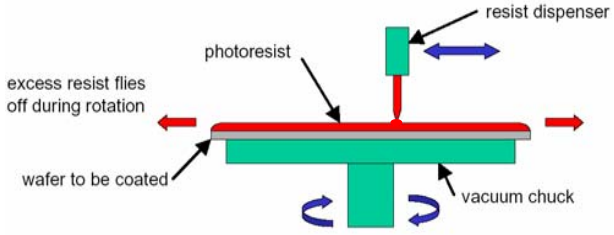


Figure 1.15. *The spin casting process as used for photoresist in photolithography.*

3 Soft baking

The soft bake causes the photoresist to cure and the remaining solvents to evaporate. Soft baking is performed using an oven (for example, convection, IR, hot plate), sometimes in a N_2 atmosphere. Soft-bake ovens provide well-controlled and uniformly distributed temperatures and a bake environment with a high degree of cleanliness. The temperature range for soft baking is between 80 to 100°C for 5 to 30 minutes, while the exposure time is established based on the heating method used and the resulting properties of the soft-baked resist.

4 Alignment and Exposure

After photoresist coating and soft baking, the wafer undergoes exposure to some form of radiation that produces a pattern image on the resist. A photomask is aligned and placed on the coated wafer with precision instruments. Two exposure forms are common:

- Optical lithography exposure which uses intense mercury arc lamp UV light and
- Electron beam lithography exposure which uses a focussed laser-type beam.

Each use specific resists, positive and negative resists, which react when exposed.
The pattern is formed on the wafer using a mask, which defines which areas of the resist surface will be exposed to radiation and those that will be covered. The chemical properties of the resist regions impinged by radiation change in a manner that depends on the type of resist used. Irradiated regions of positive photoresists become more soluble in the developer, so positive resists form a positive image of the mask on the wafer. Negative resists form a negative image of the mask on the wafer because the exposed regions become insoluble in the developer.

5 Development

Development, which is the process step that follows resist exposure, leaves behind the correct resist pattern on the wafer which will serve as the physical mask that covers wafer areas to be protected from chemical attack during subsequent etching, implantation, lift-off, etc. The wafers are developed with aqueous alkaline solutions of either sodium hydroxide or potassium hydroxide. Metal free developers such as tetramethylammonium hydroxide avoid sodium metal contamination. The developer is applied by either immersion, spraying, atomization or puddle developing, causing the un-polymerized areas of the photoresist to be dissolved and removed. A good development process has a duration of less than a minute, results in minimum pattern distortion or swelling, keeps the original film thickness of protected areas intact, and faithfully recreates the intended pattern.

Regardless of the developer method used, it is followed by thorough solvent rinsing with acetone and drying to ensure that the development action does not continue after the developer has been removed from the wafer surface.
After developing an additional baking process or *hard bake* is performed, at 120 to 180°C for 20 to 30 minutes, to harden the remaining photoresist to an enamel like finish. The photoresist is then ready to protect the underlying $Si0_2$ during etching, etc.

6 Inspection

Post-development inspection ensures that the resist processing steps have produced the desired results. This is performed using an optical microscope, although scanning electron microscope SEM and laser-based systems are also used. Process aspects that this inspection step check include:

- use of the correct mask;
- resist film quality;
- adequate image definition;
- dimensions of critical features;
- defects and their densities; and
- pattern registration.

1.5.1 Optical Lithography

Optical Lithography refers to a lithographic process that uses visible or ultraviolet (300 to 600nm) light to form patterns on the photoresist through printing. Printing is the process of projecting the image of the patterns onto the wafer surface using a light source and a photo mask. There are three types of printing:

- contact;
- proximity; and
- projection printing, each of which will now be described.

Printers or aligners are the equipment used for printing.

Patterned masks, usually composed of glass or chromium, are used during printing to cover areas of the photoresist layer that should not be exposed to light. Development of the photoresist in a developer solution after exposure to light produces a resist pattern on the wafer, which defines which wafer areas are exposed for material deposition or removal.

Both negative and positive photoresist are used.

Positive resists have two major components:

- a resin, ortho-diazoketone with xylene solvent and
- a photoactive compound dissolved in a solvent, sodium or potassium hydroxide.

The photoactive compound in its initial state is an inhibitor of dissolution. Once this photoactive dissolution inhibitor is reacted with light, the resin becomes soluble in the developer.

Negative photoresists also consist of two parts:

- a chemically inert, azide based, aliphatic polyisoprene rubber with xylene solvent and
- a photoactive agent, xylene.

When exposed to light, the photoactive agent reacts with the rubber, promoting cross-linking between the rubber molecules that make them less soluble in the developer. Such cross-linking is inhibited by oxygen, so this light exposure process is usually performed in a nitrogen atmosphere.

A disadvantage of negative resists is that exposed portions swell as the unexposed areas are dissolved by the developer. This swelling, which is volume increase due to the penetration of the developer solution into the resist material, results in distortion of the pattern features. This swelling phenomenon limits the resolution of negative resist processes. The unexposed regions of positive resists do not exhibit swelling and distortion, to the same extent as the exposed regions of negative resists. This allows positive resists to attain better image resolution.

i. Contact printing refers to the light exposure process wherein the photomask is pressed physically against the resist-covered wafer with pressure. This pressure is typically in the range of 0.05 to 0.3 atmospheres. Light with a wavelength of about 400nm is used in contact printing.
Good contact printing is capable of attaining resolutions of better than ¼µm. However, the contact between the mask and the resist diminishes the uniformity of the attainable resolution across the wafer.
To alleviate this problem, masks used in contact printing must be thin and flexible to allow better contact over the entire wafer.
Contact printing also results in defects in both the masks used and the wafers, necessitating regular mask replacement (whether thick or thin) after a certain amount of use. Mask defects include pinholes, scratches, intrusions, and star fractures. Despite these drawbacks, contact printing is widely used.

ii. *Proximity printing* is another optical lithography technique. It involves no contact between the mask and the wafer, hence masks used with this technique have longer useful lives than those used in contact printing. During proximity printing, the mask is usually only 10 to 50 μm from the wafer surface.
Minimum line width (or critical dimension, c_d):

$$c_d \cong \sqrt{\lambda g}$$

where λ = wavelength and g = gap

The resolution achieved by proximity printing is poorer than that of contact printing. This is due to the diffraction of light when passing through the slits that make up the pattern in the mask, which then traverses the gap between the mask and the wafer.
This type of diffraction is Fresnel diffraction, or near-field diffraction, since it results from the small gap between the mask and the wafer. Proximity printing resolution is improved by reducing the gap between the mask and the wafer and by using light of shorter wavelengths.

iii. *Projection printing* is the third technique used in optical lithography. It also involves no contact between the mask and the wafer. This technique employs a large gap between the mask and the wafer, such that Fresnel diffraction is no longer involved. Instead, far-field diffraction occurs, which is known as Fraunhofer diffraction.
Projection printing is the technique employed in most modern optical lithography equipment. Projection printers use a precision objective lens between the mask and the wafer, which collects diffracted light from the mask and projects it onto the wafer. The capability of a lens to collect and project diffracted light onto the wafer is measured by its numerical aperture, N_A. The N_A values of lenses used in projection printers typically range from 0.16 to 0.40.
The resolution achieved by projection printers depends on the wavelength and coherence of the incident light and the N_A of the lens. The resolution achievable by a lens is governed by Rayleigh's criterion, which defines the minimum distance between two images for them to be resolvable. Thus, for any given N_A, there exists a minimum resolvable dimension where this line resolution is given by:

$$\ell_m = c\frac{\lambda}{N_A}$$

where c is a process dependent factor, typically 0.6 to 0.8, and
N_A = numerical aperture, which is

$$N_A = \overline{n}\sin\theta$$

where $\overline{n}$ is the index of refraction.

Using a lens with a higher N_A results in better image resolution, but the penalty is that the depth of focus of a lens is inversely proportional to the square of the N_A, so improving the resolution by increasing the N_A reduces the depth of focus of the system. Poor depth of focus causes some points on the wafer to be out of focus, since no wafer surface is perfectly flat. Thus projection printing aligner design compromises between resolution and depth of focus.

1.5.2 Electron Lithography

Electron Beam Lithography (EBL) refers to a lithographic process that uses a focused beam of electrons to form the patterns needed for material deposition on (or removal from) the wafer, in contrast with optical lithography which uses light for the same purpose. Electron lithography offers higher patterning resolution than optical lithography because of the shorter wavelength (sub 100nm) of the 10 to 50 keV electrons employed.

Since a small-diameter focused beam of electrons can be scanned over a surface, an EBL system does not use masks (unlike optical lithography, which uses photomasks to project the patterns). An EBL system draws the pattern over the resist wafer using the electron beam as its drawing pen. Thus, EBL systems produce the resist pattern in a sequential manner, making it slow compared to optical systems.

A typical EBL system consists of the following parts:

- an electron gun or electron source that supplies the electrons;
- an electron column that shapes and focuses the electron beam;
- a mechanical stage that positions the wafer under the electron beam;
- a wafer handling system that automatically feeds wafers to the system and unloads them after processing; and
- a computer system that controls the equipment.

The resolution of optical lithography is limited by diffraction, which is not a problem for electron lithography. This is because of the short wavelengths of the electrons in the energy range used by EBL systems. However, the resolution of an electron lithography system is constrained by electron scattering in the resist and by various aberrations in its electron optics. Resolution is as low as 10 to 25nm.

Just like optical lithography, electron lithography also uses negative (copolymer-ethyl-acrylate) and positive (polymethylmethacrylate) 100nm thick resists, which in this case are referred to as electron beam resists (or e-beam resists). E-beam resists are e-beam-sensitive materials that are used to cover the wafer according to the defined pattern. The same polymer based photo resists are used for x-ray (½ to 5nm wavelengths) exposure.

Positive electron resists produce an image that is the same as the pattern drawn by the e-beam (positive image), while negative ones produce the inverse image of the pattern drawn (negative image). Positive resists undergo bond breaking when exposed to electron bombardment, while negative resists form bonds or cross-links between polymer chains under the same situation. As a result, areas of the positive resist that are exposed to electrons become more soluble in the developer solution, while the exposed areas of the negative resist become less soluble. The positive resists form positive images - because its electron-exposed areas result in exposed areas on the wafer after dissolving in the developer. In the case of negative resists, the electron-exposed areas become the unexposed areas on the wafer, forming a negative image.

The resolution achievable with any resist is limited by two major factors:

- the tendency of the resist to swell in the developer solution and
- electron scattering within the resist.

Resist swelling occurs as the developer (isopropyl alcohol) penetrates the resist material. The resulting volume increase can distort the pattern, such that close adjacent lines merge. Resist contraction after the resist has undergone swelling can also occur during rinsing. However, this contraction is often not enough to bring the resist back to its intended form, so the swelling distortion remains even after rinsing. Unfortunately, a swelling/contraction cycle weakens the adhesion of the smaller features of the resist to the substrate, which can create undulations in narrow lines. Reducing resist thickness decreases the resolution-limiting effects of swelling and contraction.

When electrons strike a material, they penetrate the material and lose energy from atomic collisions. These collisions can cause the striking electrons to scatter, termed *scattering*. The scattering of electrons may be backward (or back-scattering, wherein electrons bounce back), but scattering is often forward through small angles with respect to the original trajectory. During electron beam lithography, scattering occurs as the electron beam interacts with the resist and substrate atoms. This electron scattering has two major effects:

- it broadens the diameter of the incident electron beam as it penetrates the resist and substrate and
- it gives the resist an unintended extra doses of electron exposure as back-scattered electrons from the substrate reflect back into the resist.

Thus, scattering effects during e-beam lithography result in wider images than what can be ideally produced from the e-beam diameter, degrading the resolution of the EBL system. Closely spaced adjacent lines can add electron exposure to each other, a phenomenon known as the *proximity effect*.

The advantages of EBL are:

- Generation of submicron resist geometries;
- Highly automated and precisely controlled operation;
- Greater depth of focus than that available from optical lithography; and
- Direct patterning onto the wafer without using a mask.

EBL disadvantages are:

- Low throughput;
- Expensive resists; and
- Proximity effect: backscattering of electrons irradiates adjacent regions and limits minimum spacing between features.

1.6 Etching

In wafer fabrication, etching refers to a process by which material is removed from the wafer, that is, either from the silicon substrate itself or from any film or layer of material exposed on the wafer.

The rate at which the etching process occurs is known as the *etch rate*. The etching process is said to be *isotropic* if it proceeds in all wafer directions at the same rate. If it proceeds in only one direction, then it is completely *anisotropic*. Each case is illustrated in figure 1.16.
There are two types of etching processes:

- *Wet etching* where the material is dissolved when immersed in a chemical solution and
- *Dry etching* where the material is sputtered or dissolved using reactive ions or a vapour phase etchant.

Since etching processes generally neither completely isotropic or completely anisotropic, an etching process needs to be described in terms of its level of isotropy. Wet etching, or etching with the use of chemicals, is generally isotropic. On the other hand, dry etching processes that employ reactive plasmas are generally anisotropic.

1.6.1 Wet Chemical Etching

Wet etching is an etching process that utilizes liquid chemicals or etchants to remove materials from the wafer, usually in specific patterns defined by photoresist masks on the wafer. Materials not covered by these masks are etched away by the chemicals while those covered by the masks remain intact. These masks are deposited on the wafer in a previous wafer fabrication 'lithography' step, as in section 1.5.

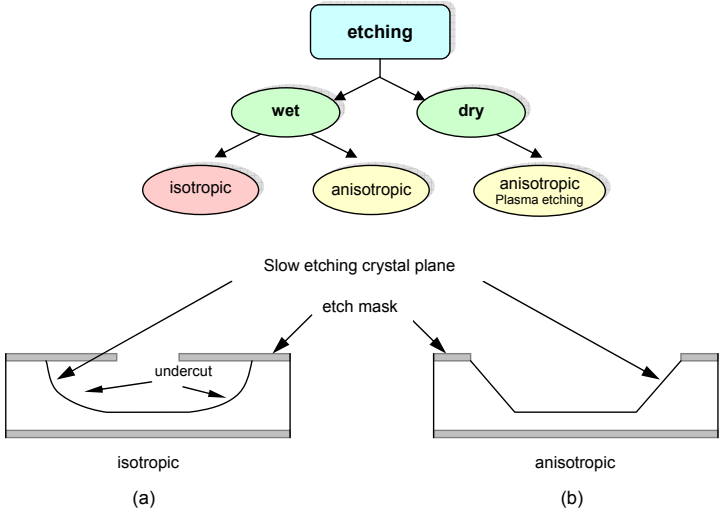


Figure 1.16. *Difference between: (a) isotropic and (b) anisotropic wet etching.*

Wet etching is the simplest etching technology. A simple wet etching process may consist of dissolution of the material to be removed in a liquid solvent, without changing the chemical nature of the dissolved material. A wet etching process usually involves one or more chemical reactions that consume the original reactants and produce new species. Typical $Si0_2$ etch rates are 30nm/min.

A basic wet etching process involves three steps:

- diffusion of the etchant to the surface for removal;
- reaction between the etchant and the material being removed; and
- diffusion of the reaction by-products from the reacted surface.

Reduction-oxidation (redox) reactions are commonly encountered in wafer fabrication wet etching processes. That is, an oxide of the material to be etched is first formed, which is then dissolved, leading to the formation of new oxide, which is again dissolved, and so on until the material is consumed.

Wet etching is generally isotropic, that is, it proceeds in all directions at the same rate. An etching process that is not isotropic is referred to as *anisotropic*. In semiconductor fabrication, a high degree of anisotropy is desired in etching because it results in a more faithful copy of the mask pattern, since only the material not directly under the mask is attacked by the etchant. Isotropic etchants, on the other hand, etch away a portion of material that is directly under the mask (usually in the shape of a quarter-circle), since its horizontal etching rate is the same as its vertical rate, as shown in figure 1.16a. When an isotropic etchant eats away a portion of the material under the mask, the etched film is said to have *undercut* the mask. The amount of undercutting is a measure of an etching parameter known as the *bias*. Bias is the difference between the lateral dimensions of the etched image and the masked image. Thus, the mask used in etching must compensate for whatever bias an etchant is known to produce, in order to create the desired feature on the wafer. Because of the isotropic nature of wet etching, it results in high bias values that are not practical for use in pattern images that have features measuring less than 3 microns. Thus, wafer feature patterns that are smaller than 3 microns are not wet-etched.

Another important consideration in any etching process is the *selectivity* of the etchant. An etchant not only attacks the material being removed, but the mask and the substrate (the surface under the material being etched) as well. The selectivity of an etchant refers to its ability to remove only the material intended for etching, while leaving the mask and substrate materials intact. Selectivity, S_f, is the ratio between the different etch rates of the etchant for different materials. A good etchant has a high selectivity value with respect to both the mask, S_{fm}, and the substrate, S_{fs}, that is, its etching rate for the film being etched must be much higher than its etching rates for both the mask and the substrate.

Despite the resolution limitations of wet etching, it has found widespread use because of its following advantages:

- low cost, the cost per wafer for dry etching is 1 to 2 orders of magnitude higher;
- high reliability;
- high throughput; and
- excellent selectivity in most cases with respect to both mask and substrate materials.

Automated wet etching systems add even more advantages:

- greater ease of use;
- higher reproducibility; and
- better efficiency in the use of etchants.

Wet etching has disadvantages including:

- limited resolution;
- higher safety risks due to the direct chemical exposure of personnel;
- high cost of etchants in some cases;
- problems related to the resist's loss of adhesion to the substrate;
- problems due to the possible formation of bubbles which inhibit the etching process; and
- problems related to incomplete or non-uniform etching.

Silicon (single-crystal or poly-crystalline), which is hydrophobic, may be isotropic wet-etched using a mixture of nitric acid ($HN0_3$) and hydrofluoric acid (HF) housed in polypropylene temperature controlled baths. The nitric acid consumes the silicon surface to form a layer of silicon dioxide, which in turn is dissolved away by the HF. The over-all reaction is:

$$Si + HNO_3 + 6\,HF \rightarrow H_2SiF_6 + HNO_2 + H_2 + H_2O$$

Silicon dioxide, which is hydroscopic, is isotropic wet-etched using a variety of HF based solutions or vapour. The over-all reaction is:

$$SiO_2 + 6\,HF \rightarrow H_2 + SiF_6 + 2\,H_2O$$

H_2SiF_6 is a gas soluble in water. Specifically, water-diluted HF with some buffering agents such as ammonium fluoride (NH_4F) is a commonly used $Si0_2$ etchant formulation. Etchants dissolve Si slower in the <111> plane because for silicon this plane has more bonds per unit area. HF is also used for silicon nitride Si_3N_4 (or H_3P0_4 at 180°C giving a 10nm/minute etch rate) and CVD oxide etching. Anisotropic etching can be achieved with K0H, with silicon and silicon dioxide.

Wet etching of aluminium and aluminium alloy layers use slightly heated (35 to 45°C) solutions of phosphoric acid, acetic acid, nitric acid, and water. The nitric acid consumes some of the aluminium material to form an aluminium oxide layer. This oxide layer is then dissolved by the phosphoric acid and water, as more $A\ell_20_3$ is formed simultaneously in maintaining the cycle.
Other materials on the wafer may be wet-etched by using the appropriate etching solutions. Generally nitric acid, including aqua regia (HCℓ plus $HN0_3$) is involved with metal (Aℓ, Cr, Ni, Au, Ag, etc.) etching.

Table 1.4: Comparison between dry and wet etching

Etching	Wet	Dry
Method	Chemical Solutions	Ion Bombardment or Chemical Reactive
Environment and Equipment	Atmosphere, Bath	Vacuum Chamber
Advantage	• Low cost, easy to implement • High etching rate, high throughput • Good selectivity for most materials • Highly selective	• Highly selective • Capable of defining small feature size (<100 nm)
Disadvantage	• Inadequate for defining feature size • Potential of chemical handling hazards • Wafer contamination issues	• High cost, hard to implement • Low throughput • Poor selectivity • Potential radiation damage
Directionality	Isotropic (except for etching Crystalline Materials)	Anisotropic (etching mainly normal to surface)

1.6.2 Dry Chemical Etching

Dry etching refers to the removal of material, typically a masked pattern of semiconductor material, by exposing the material to a bombardment of ions (usually a plasma of nitrogen, chlorine and boron trichloride) that dislodge portions of the material from the exposed surface. Unlike with many of the wet chemical etchants used in wet etching, the dry etching process typically etches directionally or anisotropically.
Dry etching does not utilize any liquid chemicals or etchants to remove materials from the wafer, generating only volatile by-products in the process. Dry etching may be realised by any of the following:

- through chemical reactions that consume the material, using chemically reactive gases or plasma;
- physical removal of the material, usually by momentum transfer or
- a combination of both physical removal and chemical reactions.

Drying etching offers better control the etching process and reduced contamination levels.

The dry etching technology can split in three separate classes:

- reactive ion etching (RIE);
- sputter etching; and
- vapour phase etching.

Plasma etching is an example of a purely chemical dry etching technique. While physical sputtering and ion beam milling are examples of purely physical dry etching techniques. Lastly, reactive ion etching is an example of dry etching that employs both physical and chemical processes.
Like wet etching, dry etching also follows the resist mask patterns on the wafer, that is, it only etches away materials that are not covered by mask material, and are therefore exposed to its etching species, while leaving areas covered by the masks almost intact. These masks were previously deposited on the wafer by a lithography fabrication step - see section 1.5.

i. Plasma etching

Plasma etching, which can etch virtually any substrate compatible material, is a purely chemical dry etching technique that consists of the following steps:

- generation of reactive species in a plasma;
- diffusion of these species to the surface of the material being etched;
- adsorption of these species on the surface;
- occurrence of chemical reactions between the species and the material being etched, forming volatile by-products;
- desorption of the by-products from the surface; and
- diffusion of the desorbed by-products into the bulk of the gas.

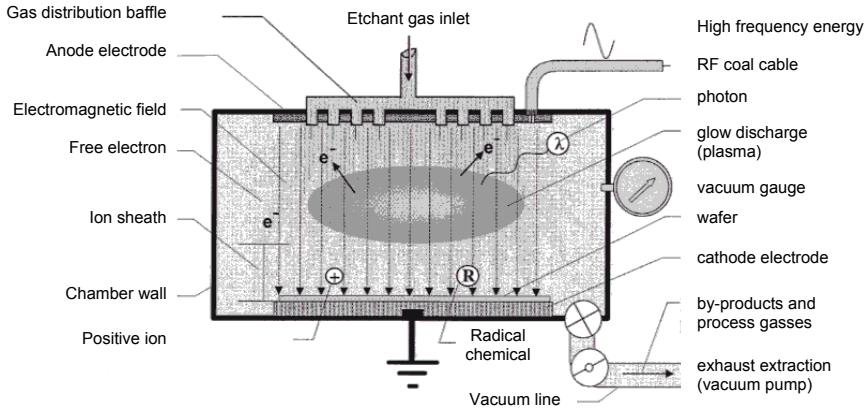


Figure 1.17. *Typical dry plasma etching system.*

Two kinds of plasma etching reactor systems are in use — the barrel (cylindrical), and the parallel plate (planar), as shown in figure 1.17. Both reactor types operate on the same principles and vary primarily in configuration only. The typical reactor consists of a vacuum reactor chamber made usually of aluminium, glass or quartz. A radio-frequency (RF) energy source is used to activate fluorine-based or chlorine-based gases which act as etchants. Wafers are loaded into the chamber, a pump evacuates the chamber (10^{-1} to 5 Torr), and the reagent gas is introduced. The RF energy ionizes the gas and forms the etching plasma, which reacts with the wafers to form volatile products which are pumped away.
The reactant gas is $CF + 0_2$ for silicon, C_2F_6 for $Si0_2$ and $CF_4 + Ar$ for Si_3N_4. Similar reactants are used to etch metals, for example $CC\ell_4$ is used to etch aluminium Aℓ and chromium Cr.
The desorption of the reaction by-products from the material surface being plasma etched is as important as the occurrence of the chemical reactions that consume the material. If such desorption does not occur, then etching cannot occur even if the chemical reactions have been completed.
The selectivity of the species used in dry etching that employs chemical reactions is important. Selectivity refers to the ability of the reactive species to etch away only the material intended for removal, while leaving all other materials intact. The species used must not attack the mask material over the material being etched as well as the material beneath it.

The reactive species used in dry chemical etching must be selected to meet the following criteria:

- high selectivity against etching the mask material over the layer being etched;
- high selectivity against etching the material under the layer being etched;
- high etch rate for the material being removed; and
- excellent etching uniformity.

They should also allow a safe, clean, and automation-ready etching process.

Another important consideration in any etching process is its anisotropy, or property of etching in one direction only. A completely anisotropic etching process that removes material in the vertical direction only is desirable, since it will follow the mask patterns on the wafer faithfully, leaving any material covered by mask material basically untouched. Most etching techniques employing purely chemical means to remove the material (whether through wet or dry etching) do not exhibit high anisotropy. This is because chemical reactions can and do occur in all directions. Thus, chemical reactions can attack in the horizontal direction and consume a portion of the material covered by the mask, termed *undercutting*, as show in figure 1.16a.
If maximum anisotropy is of utmost concern, then dry etching techniques that employ physical removal of material must be considered. One such technique is physical sputtering, which involves purely physical removal of material by bombarding it with highly energetic but chemically inert species or ions. These energetic ions collide with atoms of the material as they hit the material's surface, dislodging surface atoms in the process.

ii. Reactive ion etching

Reactive ion etching (RIE), which is sometimes referred to as reactive sputter etching, is a combination of chemical and physical etching, and involves bombarding the exposed surface material to be etched with highly energetic chemically reactive ion species. These species are usually oxidizing and reducing agents produced from process gases that have been ionized and fragmentized by a glow discharge. Such high-energy ion bombardment dislodges atoms from the material (just like purely physical sputtering), in effect achieving material removal by sputtering.

In addition to sputter-removal, the bombarding ions used in RIE are chosen to chemically react with and remove the exposed material being bombarded to produce volatile reaction by-products that can be pumped out of the system. This is the reason why RIE is widely used in wafer fabrication - it achieves the required anisotropy (by means of sputter-removal) and the required selectivity (through chemical reactions). Table 1.5 presents some examples of the process gases employed in the reactive ion etching of common wafer related materials.

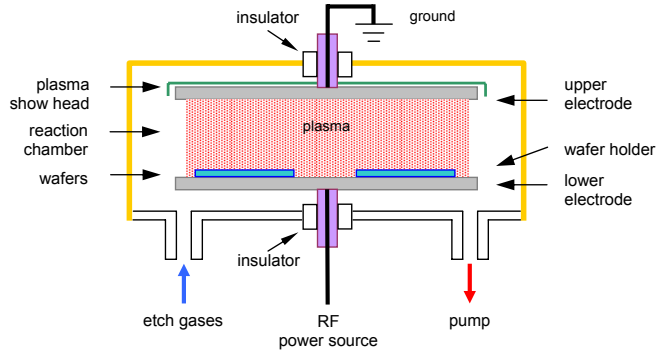


Figure 1.18. *Typical parallel-plate reactive ion etching system.*

Table 1.5: Examples of gases used in the RIE of common wafer materials

Material to be Etched	Examples of Gases used in the RIE	mask	selectivity	Etch rate A/min	Etch product
Polysilicon/silicon, Si	$CF+0_2$; CF_4; SF_6; $C\ell_2$; $CC\ell_3F$; $BC\ell_2+C\ell_2$	Resist (Cr, Ni, Aℓ)	20:1 40:1	500	$SiC\ell_2$, $SiC\ell_4$
Silicon dioxide, $Si0_2$	C_2F_6; CF_4; SF_6; HF, CF_4+H_2; CHF_3+0_2	Resist (Cr, Ni, Aℓ)	10:1 30:1	200	SiF_4
Silicon nitride, Si_3N_4	$CF_4+Ar/0_2/H_2$; CHF_3	Resist (Cr, Ni, Aℓ)	10:1 20:1	100	SiF_4
Aℓ; Aℓ doped with Si, Cu, Ti	$CC\ell_4$; $CC\ell_4+C\ell_2$; $BC\ell_3$; $BC\ell_3+C\ell_2$; $C\ell_2$	Resist SI_3N_4	10:1	300	$A\ell_2C\ell_6$, $A\ell C\ell_3$
Tungsten W, V, Ti, Ta, Mo	Fluorinated gases, CF_4				WF_6 ….
Refractory Silicides	Fluorinated plus chlorinated gases (with or without 0_2)				SiF_4
Resist / polymer	0_2	SI_3N_4 (Cr, Ni)	50:1	500	

In RIE, the substrate is placed inside a reactor in which several chemically reactive gases (CF_4 or $CC\ell_4$) are introduced at low pressure (10^{-4} to 10^{-3} Torr). A plasma is struck, by electrical discharge, in the gas mixture using an RF power source, stripping the gas molecules down to ions. The ions are accelerated towards, and react at, the surface of the material being etched, forming another gaseous material, which is removed by the low pressure in-line vacuum pressure. This is the chemical part of reactive ion etching. There is also a physical part which is similar in nature to the sputtering deposition process. If the ions have high enough energy, a few hundred eV, they can knock atoms out of the material to be etched without a chemical reaction. It is a complex task to develop dry etch processes that balance chemical and physical etching, since there are many parameters to adjust. By changing the balance it is possible to influence the anisotropy of the etching, since the chemical part is isotropic and the physical part is highly anisotropic, the combination can form sidewalls that have shapes from rounded to vertical. A schematic of a typical reactive ion etching system is shown in figure 1.18.

A subclass of RIE, increasing in use, is deep RIE. In this process, etch depths of hundreds of microns can be achieved with almost vertical sidewalls. Two different gas compositions are alternated in the reactor. The first gas composition creates a polymer on the substrate surface, and the second gas composition etches the substrate. The polymer is immediately sputtered away by the physical part of the etching, but only on the horizontal surfaces and not the sidewalls. Since the polymer only dissolves slowly in the chemical part of the etching, it builds up on the sidewalls and protects them from etching. As a result, etching aspect ratios of 50 to 1 can be achieved. The process can be used to etch completely through a silicon substrate, and etch rates are 3 to 4 times higher than wet etching.

iii. Sputter etching

Sputter etching (ion milling) is essentially RIE without reactive ions. The systems used are similar in principle to sputtering deposition systems. The difference is that the substrate is subjected to the ion bombardment instead of the material target used in sputter deposition, as shown in figure 1.19.
The wafer to be etched is attached to a negative electrode, or target, in a glow-discharge circuit. Positive argon ions bombard the wafer surface, resulting in the dislocation of the surface atoms. Power is provided by an RF energy source.

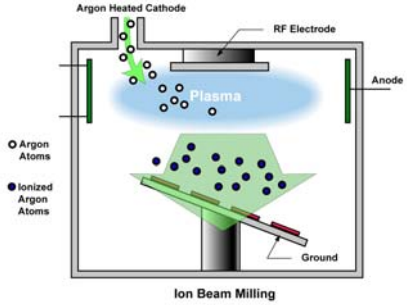


Figure 1.19. *Typical sputter etching system.*

Targeting the layer to be etched with incident ions that are perpendicular to its surface ensures that only the uncovered material is removed. Unfortunately, such a purely physical process is also non-selective, that is, it also attacks the mask layer covering the material being etched, since the mask is also directly hit by the bombarding species. For this reason, physical sputtering is not common as a dry etching technique for wafer fabrication.

A good balance between isotropy and selectivity may be achieved by employing both physical sputtering and chemical means in the same dry etching process, such as reactive ion etching.

iv. Vapour phase etching

Vapour phase etching is another dry etching method, which can be performed with simpler equipment than what RIE requires. In this process, the wafer to be etched is placed inside a chamber, into which one or more gases are introduced. The material to be etched is dissolved at the surface in a chemical reaction with the gas molecules. The two most common vapour phase etching technologies are silicon dioxide etching using hydrogen fluoride (HF) and silicon etching using xenon diflouride (XeF_2), both of which are isotropic in nature. The vapour phase process must not have bi-products form in the chemical reaction that condense on the surface and interfere with the etching process.

Dry etching technology is expensive to operate compared to wet etching. If feature resolution in thin film structures or vertical sidewalls for deep etchings in the substrate are required, consider dry etching.

1.7 Lift-off Processing

In semiconductor wafer fabrication, the term *lift-off* refers to the process of creating patterns on the wafer surface through an additive process, as opposed to the more familiar patterning techniques that involve subtractive processes, such as etching. Lift-off is used in patterning metal films for interconnections.
Lift-off consists of forming an inverse image of the pattern desired on the wafer using a stencil layer, which covers certain areas on the wafer and exposes the remainder. The layer to be patterned is then deposited over the stencilled wafer. In the exposed areas of the stencil, the layer material is deposited directly on the wafer substrate, while in the covered areas, the material is deposited on the stencil film.
After the layer material has been deposited, the wafer is immersed in a liquid that dissolves the stencil layer. Once the stencil is dissolved by the liquid, the layer material lifts off (hence the term *lift-off*), leaving behind the layer material that was deposited directly onto the wafer substrate itself, which forms the final pattern on the wafer.

The lift-off process as a patterning technique offers the following advantages:

- composite layers consisting of several different materials may be deposited one material at a time and then patterned with a single lift-off;
- residues that are difficult to remove are prevented in the absence of etching of the patterned layer; and
- sloped side walls become possible, resulting in good step coverage.

On the other hand, the main disadvantage of the lift-off process is the difficulty of creating the required stencil patterns for successful lift-off.

Materials that are used as stencil film for lift-off include:

- a single photoresist layer;
- two photoresist layers;
- a photoresist-aluminium-photoresist layer;
- polyimide/molybdenum layer;
- polyimide/polysulphone/$Si0_2$ layer; and
- inorganic dielectric-photoresist layer.

Any deposited film can be lifted-off, provided:

- During film deposition, the substrate does not reach temperatures that burn the photoresist.
- The film quality is not absolutely critical. Photoresist will outgas slightly in vacuum systems, which may adversely affect the deposited film quality.
- Adhesion of the deposited film on the substrate is good.
- The film can be wetted by the solvent.
- The film is thin enough and/or grainy enough to allow solvent to seep beneath.
- The film is not elastic and is thin and/or brittle enough to tear along adhesion lines.

The key to successful lift-off is to ensure the existence of a distinct break between the layer material deposited on top of the stencil and the layer material deposited on the wafer substrate. Such a separation allows the dissolving liquid to reach and attack the stencil layer. One technique to create such 'breaks' is cold evaporation over steep steps.

1.8 Resistor Fabrication

Circuits on semiconductor wafers may require resistive components, of which there are several types:

- diffused resistors;
- ion-implanted resistors;
- thin-film resistors; and
- polysilicon resistors.

Diffused resistors are fabricated through p-type diffusion into an n-type background, which is usually accomplished simultaneously with base diffusion. The sheet resistance (see equation (1.5)) of a base diffusion is usually 100 to 200 Ohms per square (see example 1.1). As such, the use of base diffusion results in good layout proportions for resistors ranging in value from 50 to 10k Ohms. A diffused resistor is isolated from its background by the contact potential of its corresponding p-n junction or by a high reverse voltage if the tub is biased properly. Such practice allows several diffused resistors to be built in a single tub, resulting in savings in chip area.

Ion-implanted resistors exhibit sheet resistances that are as high as 5 kilo-Ohms per square, allowing significant reductions in chip area requirements of high resistance-valued resistors. Ion-implanted resistors are fabricated by first forming two base diffusions and then ion-implanting the resistor between them. Contacts are then formed on the base diffusions. Ion-implanted resistors are suited for low-power digital and linear circuits because of their high sheet resistance.

Thin-film resistors offer high precision and stability. They are fabricated by vacuum evaporation or sputtering of thin films of resistive materials directly on the substrate oxide layer. Materials used for thin-film resistors include nichrome, sichrome, and a variety of refractory silicides. These materials exhibit good adhesion on the oxide as thin films, and are usually built with a film thickness of about 10nm to 100 nm. Thin film resistance can be adjusted (increased) precisely by laser trimming.

Polysilicon resistors are fabricated from undoped polysilicon films that are deposited onto the wafer. These are implanted with the correct type and amount of impurity (n or p type), and then annealed at about 600 to 1000°C. Polysilicon resistors are used when high resistance is needed but wide tolerances are acceptable.

1.9 Isolation Techniques

The individual components and regions that comprise the circuit on a monolithic die need to have electrical isolation from each other in order to function. The most common techniques used for achieving component isolation during wafer fabrication are:

- by employing reverse-biased p-n junctions;
- through mesa isolation;
- by wafer bonding to an insulating substrate;
- by oxide isolation;
- by trenching; and
- through a combination of any of these processes.

A reverse-biased p-n junction has an extremely low leakage current, so is use as an isolation technique during wafer fabrication. By doping two adjacent regions with opposite types of conductivity and providing them with adequate reverse biasing, they become effectively isolated from each other. In such a situation, the coupling between the regions is only capacitive, which is an issue at high frequencies.

Another technique for achieving component isolation is known as *mesa isolation*. This involves the building of the components on an active film that is grown on an insulating (or semi-insulating) film, and then etching moats around the components. This results in the components becoming individual 'islands', or *mesas*, hence the name 'mesa isolation'. Circuits fabricated on silicon on insulators, as well as those made on epitaxial GaAs over semi-insulating GaAs substrate, are examples of applications of mesa isolation.

Wafer bonding to an insulative substrate may be considered a variant of mesa isolation. This isolation technique takes advantage of the fact that any two flat, smooth, clean, and hydrophilic surfaces can be bonded at ambient temperature without the use of external forces. Wafer bonding can be applied to widely dissimilar materials. Once the moats are etched around the mesas, isolation is provided by the insulating substrate.

Oxide isolation techniques consist of a series of material deposition and removal steps that lead to the formation of active single-crystal tubs that are surrounded by an oxide layer. Such oxide layers, once formed, provide near-perfect isolation between the active tubs.

Trenching is a process wherein anisotropic wet etching or reactive ion etching is employed to plough a trench around the active region. The trench is then filled with isolating material. Planarization is performed after filling the trenches, see section 1.11.

1.10 Wafer Cleaning

There are a number of wafer cleaning techniques or steps employed to ensure that a semiconductor wafer is always free of contaminants and foreign materials as it undergoes the wafer fabrication process. Different contaminants have different properties, and therefore have different requirements for removal from the wafer.

Basic Wet Concepts of Cleaning

Remove organic contamination and particles by oxidation:

$NH_40H:H_20_2:H_20$ = 1:1:5 to 1:2:7 at 70 to 90°C.

Remove metal contamination by forming a soluble complex:

$HC\ell:H_20_2:H_20$ =1:1:6 to 1:2:8 at 70 to 90°C.

Removing Metal Contamination

Wet cleaning process is the most effective method for removing metallic contamination:

HF: 0.5%, H_20_2: 10% clean, due to high oxidizing mechanism of H_20_2.

Ca contamination causes a rough surface and a defect density in the oxide:

The threshold value for Ca contamination is 10^9 atoms/cm^2.

Removal of Organic Contamination

Photoresist removal by plasma and wet cleaning ($H_2S0_4:H_20_2$ = 3:1 to 4:1 at 120 to 130°C).

Depletion of H_20 due to high wafer temperature, 120°C, causes unstable process control. The alternative is to add ozone into water, which can be used as a strong oxidizing agent that decomposes organic impurities. The oxide thickness increases as the immersion time increases and with the concentration of ozone.

Photoresist Stripping

Resist stripping, is the removal of unwanted photoresist layers from the wafer. Its objective is to eliminate the photoresist material from the wafer as quickly as possible, without allowing any surface materials under the resist to be attacked by the chemicals used. Resist stripping is classified into:

- organic stripping;
- inorganic stripping; and
- dry stripping.

Organic stripping employs organic strippers, which are chemicals that break down the structure of the resist layer. Organic strippers are phenol-based, but their short pot-life and difficulties with phenol disposal make low-phenol or phenol-free organic strippers more common. Glycol ethers or Na0H and silicates are used for positive resists.

Wet inorganic strippers, also known as oxidizing-type strippers, are used for inorganic stripping, usually to remove photoresist from non-metallised wafers, as well as post-baked and other hard-to-remove resists. Inorganic strippers are solutions of sulphuric acid H_2S0_4 and an oxidant (such as ammonium persulphate $Na_2S_20_8$, hydrogen peroxide H_20_2 or chromic $Cr0_3$), heated to about 125°C.

Dry stripping or *plasma ashing*, pertains to the removal of photoresist by dry etching using plasma etching equipment. The wafers are placed into a chamber under vacuum, and oxygen is introduced and subjected to radio frequency (13.56MHz) power which creates oxygen radicals. The radicals react with the resist to oxidize it to water, carbon monoxide, and carbon dioxide. The ashing step removes the top layer or skin of the resist, then additional wet or dry etching processes strip away the remaining resist. Its advantages over wet etching with organic or inorganic strippers include better safety, absence of metal ion contamination, decreased pollution issues, and a lower tendency to attach to underlying substrate layers.

Chemical Removal of Film Contaminants

Chemically bonded films of contaminant material can be removed from a wafer surface by chemical cleaning. Chemical cleaning comes in various forms, depending on the nature of the film contaminants that need to be removed and from which type of surface. Chemical cleaning is performed with a series of acids (HF, H_2S0_4, H_20_2, $HC\ell$, and $HN0_3$) and rinse baths involving de-ionised water. As an example, removal of film contaminants from a wafer with nothing but thermally grown oxide may consist of the following steps:

- preliminary cleaning;
- removal of residual organic contaminants and some metals;
- stripping of the hydrous oxide film created by the previous step;
- desorption of atomic and ionic contaminants; and
- drying.

Storage of cleaned wafers should be avoided but, if necessary, must be done using closed glass containers inside a nitrogen dry box.

Sputter Etching of Native Oxide Films
A thin oxide layer grows over silicon or aluminium when exposed to air, forming SiO_2 and $A\ell_2O_3$, respectively. These native oxide layers need to be removed from specific areas, because they exhibit adverse effects such as higher contact resistance or hampered interfacial reactions. In-situ sputter or plasma etching are the techniques commonly utilized to remove such native oxides from the wafer. *In-situ* means performing the sputter or plasma etching in the same vacuum environment where the overlying layer will be deposited.

Elimination of Particulates
Wafer contamination with insoluble particulates is a common problem. There are two frequently utilized techniques for removing particulates from a wafer:

- ultrasonic scrubbing and
- a combination of mechanical scrubbing and high-pressure spraying.

Ultrasonic scrubbing consists of immersing the wafer in a liquid medium, which is supplied with ultrasonic energy. The sonic agitation causes microscopic bubbles to form and collapse, creating shock waves that loosen and displace particles. Ultrasonic scrubbing requires a filtration system that removes the particles from the bath as they are detached. One drawback of ultrasonic scrubbing is that it can cause mechanical damage to substrate layers.

Mechanical scrubbing employs a brush that rotates and hydroplanes over a solvent applied on the wafer surface. This means that the brush does not actually contact the wafer, but the solvent moved by the rotating brush dislodges particles from the wafer surface. Simultaneously, high-pressure spraying of the wafer surface with a de-ionised water jet helps in clearing the wafer surface of particulate contamination.

1.11 Planarization

Planarization is the process of improving the flatness or planarity of the surface of a semiconductor wafer through various methods known as planarization techniques.

The starting raw wafers for semiconductor device fabrication are ideally flat or planar. However, as the wafer goes through the various device fabrication steps, layers of different materials, shapes, and depths are deposited over the wafer surface through different growth and deposition techniques. Also, portions of materials already deposited on the wafer need to be removed at different processing stages. This series of material growth, deposition, and removal steps decreases the flatness or planarity of the wafer surface.

Device fabrication that involves multi-layers of metallization aggravates the problem of wafer non-planarity. Narrower metallisation has necessitated the need for thicker metallization in order to meet the current requirements of the device. Modern fabrication techniques that increase the number of metal layers on the wafer while decreasing the width of the metal interconnects increases the problem of wafer non-planarity.

A decrease in the flatness of the wafer's surface introduces at least two problems to device fabrication. First, ensuring ample step coverage of fine lines so that no breaks in the continuity of the lines arise becomes more difficult as the wafer becomes less flat. Second, progressive loss of planarity eventually makes the imaging of fine-featured patterns on the wafer problematic.

There are several planarization techniques used in wafer fabrication. There are two categories of planarization techniques, namely, local planarization and global planarization. Local planarization refers to smoothing techniques that increase planarity over small areas. Global planarization involves techniques that decrease long-range variations in wafer surface topology, especially those that occur over the entire image field of the stepper.

Planarization techniques include:

- oxidation;
- chemical etching;
- taper control by ion implant damage;
- deposition of films of low-melting point glass;
- resputtering of deposited films to smooth them out;
- use of polyimide films;
- use of resins and low-viscosity liquid epoxies;
- use of spin-on glass materials;
- sacrificial etch-back; and
- mechanical-chemical polishing of the wafer.

1.12 Gettering

Gettering is the process of removing device-degrading impurities from the active regions of the wafer. Gettering, which can be performed during crystal growth or in subsequent wafer fabrication steps, is an important factor for enhancing the performance and yield of semiconductor devices. The mechanism by which gettering removes impurities from device regions can be described by the following steps:

- the impurities to be gettered are released into solid solution from whatever precipitate they are in;
- they undergo diffusion through the silicon; then
- they are trapped by defects such as dislocations or precipitates in an area away from active regions.

There are two general classifications of gettering, namely, extrinsic and intrinsic.

Extrinsic gettering refers to gettering that employs external means to create damage or stress in the silicon lattice in such a way that extended defects needed for trapping impurities are formed. These chemically reactive trapping sites are usually located at the wafer backside, which are removed if the backside is subsequently lapped to produce a thinner substrate.
Several methods can be used to achieve external gettering. For instance, the introduction of mechanical damage by abrasion, grooving or sandblasting can produce stresses at the backside of a wafer, which when annealed create dislocations that tend to relieve these stresses. These locations serve as gettering sites. The main drawback of this method is its tendency to initiate and propagate wafer backside microcracks that may compromise wafer mechanical strength.
Diffusing phosphorus into the wafer backside is another technique used for external gettering. Phosphorus diffusion into silicon results in phosphorus vacancies or dislocations that serve as trapping sites for impurity atoms, such as gold. Another effect of P diffusion is the creation of Si-P precipitates, which are capable of removing Ni impurities through interactions between Si self-interstitials and Ni atoms, nucleating $NiSi_2$ particles in the process.
The introduction of damage by a laser is another external gettering method. Scanning a laser beam across the wafer surface induces damage that is similar to mechanical damage, but the laser damage is controlled and cleaner. The laser subjects the irradiated areas to thermal shock, forming dislocation nests that serve as gettering sites.

Ion-bombardment to produce wafer backside damage is another method for external gettering, using high-energy ions to induce the necessary stress within the lattices of the wafer backside. Deposition of a polysilicon layer on the wafer backside has also been used for external gettering. Polysilicon layers introduce grain boundaries and lattice disorder that can act as traps for mobile impurities.

Intrinsic gettering refers to gettering that involves impurity trapping sites created by precipitating supersaturated oxygen out of the silicon wafer. The precipitation of supersaturated oxygen creates clusters that continuously grow, progressively introducing stress into the wafer.
Eventually these stresses reach the point where they need to be relieved. Dislocation loops or stacking faults form to provide the necessary stress relief. These dislocations and faults subsequently serve as trapping sites for impurities.
A basic requirement of intrinsic gettering is starting wafers that have sufficient, but not excessive oxygen levels (15 to 20 ppma).

The advantages of intrinsic gettering over extrinsic gettering are:

- it does not require subjecting the wafer to any treatment except for heating;
- its volume of impurity sink is significantly larger than that of external gettering on the wafer backside;
- its gettering regions are much closer to the device active operating regions.

1.13 Lifetime control

Two basic processes have been developed to reduce the lifetime of carriers in power semiconductor devices.

- Thermal diffusion of gold or platinum or
- Bombardment of the silicon with high-energy particles such as electrons, protons and helium.

The diffusion of gold or platinum occurs more rapidly than the diffusion of group III and V dopants, hence the precious metal is diffused at 800°C to 900°C, just prior to metallization, which is performed at a lower temperature. The higher the precious metal diffusion temperature, the higher the solubility and the lower the carrier lifetime. Disadvantages of precious metal diffusion include:

- devices cannot be tested prior to or immediately after impurity diffusion and
- small temperature changes cause a wide variation in device characteristics.

In high-resistivity silicon used to fabricate high-voltage power devices, irradiation bombardment causes defects composed of complexes of vacancies with impurity atoms of oxygen and of two adjacent vacancy sites in the lattice. The advantages of the use of irradiation in order to reduce carrier lifetime in power semiconductor devices are:

- irradiation is performed at room temperature, after fabrication;
- irradiation can be accurately controlled hence a tighter distribution of electrical characteristics results;
- overdose annealing can be performed at only 400°C; and
- it is a clean, non-contaminating process.

Attention has been focussed on MeV proton irradiation, which has high costs and long processing scan times, but offers the most accurate and precise form of lifetime control. The consequences of lifetime control are an improvement in switching speed (dynamic characteristics) at the expense of increased leakage and on-state voltage (steady-state characteristics) and displacement damage.

1.14 Silicide formation

Metallization refers to the metal layers that electrically interconnect the various device structures fabricated on the silicon substrate. Interconnection paths that possess low resistivities and the ability to withstand subsequent high temperature processes are critical to semiconductor manufacturing. Thin-film aluminium is the most widely used material for metallization because of its low resistivity and its adhesion compatibility with $Si0_2$. The resistivity of Aℓ is low enough for IC interconnection purposes, but its low melting temperature of 660°C and the low Aℓ-Si (11.3%:88.7%) eutectic temperature of 577°C, restrict subsequent processes to operating temperatures of less than 500°C. Aluminium alloys (lightly doped Aℓ) such as Aℓ-Cu (or TiN barrier metal) are preferred to pure aluminium for metallization because these inhibit problems like electromigration and junction spiking – see figure 1.20.
Aℓ metal layers are usually deposited through Physical Vapour Deposition (PVD) by sputtering.

Upon exposure to oxygen, aluminium readily forms a native thin oxide on its surface, $A\ell_20_3$, even at ambient temperature. The presence of such an oxide layer can increase the contact resistance of the Aℓ layer. It can also inhibit the sputtering of an Aℓ target or etching of an Aℓ thin film, resulting in processing difficulties. Aℓ can readily suffer from corrosion in the presence of a corrosive contaminant and moisture. For instance, if phosphorus-doped silicon dioxide is deposited over Aℓ lines, phosphoric acid results if moisture ingresses through the glass. The acid corrodes the Aℓ connects.

For these reasons, instead of using Aℓ, low-resistivity interconnections are usually fabricated using materials known as *refractory metal silicides* (MSi_x), which can handle much higher processing temperatures than Aℓ.
The formation of refractory metal silicides (such as WSi_2, $TiSi_2$, $MoSi_2$, and $TaSi_2$) at the wafer surface can be accomplished in four ways:

i. by deposition of the pure metal onto a Si layer (which can be a single-crystal or poly-crystalline Si substrate);
ii. simultaneous evaporation of the silicon and the refractory metal from two sources (or co-evaporation);
iii. sputter-deposition of the silicide, either from a composite target or by co-sputtering; and
iv. chemical vapour deposition (CVD).

i. The *silicide formation* technique of directly depositing a refractory metal on a silicon surface to form the required silicide layer employs the process of direct metallurgical reaction. After depositing the metal on the silicon, the wafer is exposed to high temperatures that promote the chemical reactions between the metal and the silicon needed to form the silicide. In such a metallurgical reaction, metal-rich silicides generally form first, and continue to grow until all the metal is consumed. When the metal has been consumed, silicides of lower metal content start appearing, which can continue to grow by consuming the metal-rich silicides. To illustrate this with titanium Ti as the metal, TiSi is the first silicide to form on Si, typically appearing at a temperature above 500°C and peaking at 700°C. $TiSi_2$ only starts to appear at 600°C and peaks at 800°C. Beyond 800°C, TiSi would be fully converted into $TiSi_2$, at which point the system attains stability.
Silicide formation by direct metallurgical reaction consumes silicon from the substrate onto which the metal is placed. Thus, it is important that enough silicon is available when this technique is used to form silicide layers.

ii. *Co-evaporation*, another technique for silicide formation, consists of the simultaneous deposition of the metal and the silicon under high vacuum conditions. The metal and silicon are vaporized through one of several possible heating techniques: with an electron beam, by rf induction, with a laser or by resistive heating. However, e-beam heating is the preferred technique because refractory metals (Ti, Ta,

Mo, W) have high melting points (1670 to 2996°C) while silicon has a low vapour pressure. With the use of two e-beam guns whose power supplies are controlled, the proper metal-to-Si ratio can be achieved.

Aspects of an evaporation process that ensure the deposition of films with repeatable properties include:

- the evaporation base pressure (must be < 1 micro-torr);
- evaporation rates;
- purity of the elements; and
- residual gases present in the evaporation chamber.

iii. *Sputter deposition*, the third technique, is for silicide formation. Sputtering is a deposition process wherein atoms or molecules are ejected from a target material by high-energy particle bombardment so that the ejected particles can condense on a substrate as a thin film. As in co-evaporation, the correct sputtering rates of the metal and Si must be determined and applied to ensure proper deposition of the silicide film. The step coverage of co-sputtered films is superior to that of evaporated films.
Sputter-deposition of silicides has various forms. For instance, sputtering from two targets using multi-pass sputtering systems achieves the appropriate mixture of metal and Si in a layered structure. Sintering then completes the chemical reaction between the metal and Si to form the silicide. Sputtering from a composite target (MSi_x) can also be performed, allowing better compositional control, but vulnerability to contamination are associated with composite targets.

iv. *Chemical vapour deposition* CVD of silicides, the fourth technique for silicide formation, involves chemical reactions between vapours to form the silicide film, and offers advantages over the other techniques, specifically, better step coverage, higher purity of the deposited films, and higher wafer throughput. However, the availability of gas reactants whose chemical reactions will produce the desired silicide is necessary for CVD in silicide formation.
Silicides in Table 1.6 are highly conductive alloys of silicon and metals; contact materials in silicon device manufacturing; combine advantageous features of metal contacts (significantly lower resistivity than poly-Si) and poly-Si contacts (no electromigration), and have superior resistance to temperature than metals like Aℓ. NiSi is preferred because of its low sintering temperature and shallow penetration into the silicon, as indicated in Table 1.6.

Table 1.6: Low resistivity metal silicides

Metal silicide	composition	resistivity	sintered at	nm of Si per nm of metal
		μohm-cm	°C	
cobalt silicide	$CoSi_2$	16-20	900	3.65
titanium silicide	$TiSi_2$	13-16	900	2.25
tungsten silicide	WSi_2	60-80	1000	
tantalum silicide	$TaSi_2$	35-40	1000	
platinum silicide	PtSi	25-35	600-700	
nickel silicide	NiSi	14-20	400-700	1.85

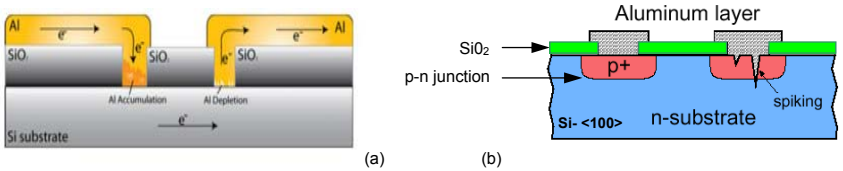
1.15 Ohmic contact

An ohmic contact is a resistive connection which is voltage independent (and technically has a Schottky barrier height of $\Phi_b \leq 0$). Aluminium is commonly evaporated in a vacuum at near room temperature, onto the wafer surface to form a metallised electrical contact. Deposit rates of ½ μm/minute are typical. Thermal annealing reduces contact resistivity.
If the silicon is n-type, a p-n Schottky junction is formed ($\Phi_b > 0$), which is undesirable as an ohmic contact. This junction forming aspect is discussed at the end of section 3.1.4. Ohmic metal contact to both p-type and n-type semiconductors with a large bandgap, like silicon carbide or gallium nitride, is technically more complicated – as introduced in section 1.20. IC metallisation processes cover the full range of different possibilities, including those applicable to power devices.

Metallization: Metal Deposition

After devices have been fabricated in the silicon substrate, a metallization process is the fabrication step in which proper interconnection of circuit elements and regions, is made.
Metal layers are deposited by a vacuum deposition technique on the wafer to form conductive pathways. The most common deposited metals include aluminium, nickel, chromium, gold, germanium, copper, silver, titanium, tungsten, platinum, and tantalum. Selected metal alloys may also be used.

Aluminium is the most common metal used to interconnect ICs, both to make ohmic contact to the devices and to connect these to the bonding pads on the chip's edge. Aluminium adheres well to both silicon and silicon dioxide, can be easily vacuum deposited - since it has a low boiling point, and has high conductivity. In addition to pure aluminium, alloys of aluminium are used to form IC interconnections for different performance-related reasons. For example, small amounts of copper are added to reduce the potential for electromigration effects (in which current applied to the device induces mass transport of the metal, accumulation and depletion, depending on current direction, as shown in figure 1.20a). Small amounts of silicon also are added to aluminium metallization to reduce the formation of metal 'spikes' that occur over contact holes. Copper has a higher conductivity (ρ_{Cu} = 1.67 μΩcm, $\rho_{Aℓ}$ = 1.65 μΩcm), is more malleable, and better electromigration resistance than Aℓ, but copper tends to oxidise, corrode, is not amenable to dry etching, and adheres poorly to, and contaminates, $Si0_2$. Anywhere aluminium comes in contact with silicon, some silicon is absorbed leaving voids. Random pits result which after annealing fill with aluminium forming spikes which penetrate through to the surface shorting shallow junctions, as shown in figure 1.20b.



(a) (b)

Figure 1.20. *Aluminium: (a) current direction dependant migration and (b) metallised junction spiking.*

Metallisation involves the following processing possibilities.

i. *Photolithographic* techniques similar to those used during device fabrication are used to deposit conductive patterns during metallization. The metal is deposited then covered with a patterned photoresist, and subsequently etched. Alternatively, the resist is applied first, followed by deposition of the metal. The wafer is then placed in a solvent that causes swelling of the resist. As the resist swells it lifts the overlaid metal away from the wafer surface.

ii. *Etching* deposited metals requires a more aggressive chemical attack than etching $Si0_2$ during device fabrication. Consequently, the resists need to be tougher. The process of *silyation* is frequently performed to harden the photoresist before it is subjected to etching. During silyation, silicon atoms are introduced into the surface of the organic resist. This process can be accomplished with either wet or dry procedures. Most commonly, a wet bath of either hexamethyldisilazane or silazone in xylene is used.

iii. Metal layers are *vacuum-deposited* onto wafers by one of the following methods:

a. Filament Evaporation;
b. Flash Evaporation;
c. Electron-beam Evaporation;
d. Sputtering or
e. Induction Evaporation.

a. Filament Evaporation

Filament evaporation, also called *resistive evaporation*, is the simplest method, is accomplished by gradually heating a filament of the metal to be evaporated. This metal may be in one of several different forms: pellets, wire, crystal, etc. Gold, platinum, and aluminium are metals typically used. The PMOS process uses aluminium wire.
The metal is placed in a basket. Electrodes are connected to either side of the basket or bell jar and a high current passed through it, causing the basket to heat because of thermal resistance. As the power, and therefore heat, is increased, the metallic filament partially melts and wets the filament and as the current through the filament is increased further, the metal eventually vaporizes. In this way, atoms of aluminium break free from the filament and condense on the cooler surface of the semiconductor wafers, forming the desired metal layer on the wafers. While filament evaporation is the simplest of all metallization approaches, problems of contamination during evaporation preclude its widespread use in IC fabrication.

The procedure can be summarised as follows:

- Metal sources – pellets, are placed on filaments (tungsten, molybdenum, quartz, graphite, etc.);
- Metals are heated via a resistive filament under vacuum conditions to their melting point;
- Metal pellets give off a vapour, the atoms of which are kinetic energy dependant on temperature;
- Metal atoms travel in a straight line from the source to a sample;

- Deposition rates of order of 1 nm/s are standard; and
- Contamination may result from the filament being at least the same temperature as the source.

b. Flash Evaporation

Flash or partial evaporation uses the principle of thermal-resistance heating to evaporate metals.
The evaporation process requires high temperature, and low pressure; and can be separated into three steps.

- The solid aluminium metal is changed into a gaseous vapour.
- The gaseous aluminium is transported to the substrate.
- The gaseous aluminium is condensed onto the substrate.

The sources are usually either powder or thin wires. In the latter case, wire is continuously fed from a spool until it contacts the heated ceramic bar. Upon contact, the metal evaporates and is subsequently deposited on the substrate.

The evaporation process does not produce a uniform layer of aluminium across the substrate. The deposition rate decreases radially away from the centre of the substrate.
Like filament evaporation, flash evaporation offers radiation-free coatings. This technique does offer some benefits beyond filament evaporation: contamination-free coatings, speed or good throughput of wafers, and the ability to coat materials or layers that are composite in nature.

c. Electron-beam Evaporation

Electron-beam evaporation, frequently called e-beam, functions by focusing an intense beam of electrons into a crucible, or pocket, in the evaporator that contains the aluminium to be deposited. As the beam is directed into the source area, the aluminium is heated to its melting point, and eventually, evaporation temperature. The benefits of this technique are speed and low contamination, since only the electron beam touches the aluminium source material. The process can be summarised as follows.

- Thermal emission of electrons from a filament source, usually tungsten, is used to heat samples to high temperatures.
- Electron beams are used when the required temperatures are too high for thermal evaporation.
- A magnetic field and rastering are used to steer the beam 270° into the metal source. This is done to allow shielding of the tungsten filament and to prevent contamination.
- Electrons striking metals can produce X-rays which sometimes causes damage to material layers on a wafer. An annealing stage restores any damaged areas.

Because the contact layers for power devices is relatively thick, typical 10μm, electron-beam evaporation is typically used because of the high deposition rate. To prevent the silicon from dissolving in the metal layer, 1 to 2% silicon is added, while 5 to 15% copper decreases electromigration. The technique is also applicable to Ni, Au, and Cu layers. Heating to 600 to 700°C forms silicides at the interface. A final Ni chemical disposition layer allows solder connection.

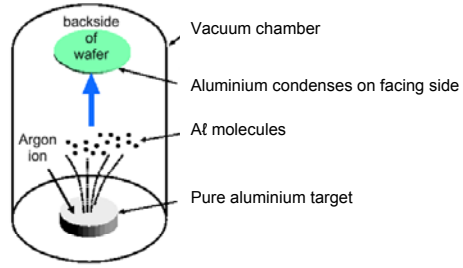


Figure 1.21. *Typical sputtering chamber.*

d. Sputtering

Aluminium sputtering is used commonly in metallization processes because the adhesion of the deposited metals is excellent. RF sputtering is done by ionizing inert gas particles, argon, in an electric field - producing a gas plasma, in a low-pressure or partial-vacuum atmosphere. Then the ions are directed toward the source or target - comprised of the metal to be subsequently deposited, where the energy of these gas particles physically dislodges, or sputters off, atoms of the source material. The dislodged atoms are then deposited in a thin film on the silicon substrate facing the target, figure 1.21.

The procedure can be summarised as follows.

- A parallel plate system generates high-energy ions that accelerate to bombard the source material – the material to be deposited.
- If the ion energy is high enough (typically 4 times the bond energy of the source) atoms will be knocked lose, ejected - sputtered. Typical bond energies are 5eV.
- the sputtered atoms reach the substrate and the gas providing ions must be inert, that is, must not chemically react with the sample substrate.
- the sputtered atoms condense and form a thin film over the substrate. Low pressures are incompatible with sputtering, thus the sample must be located close to the target source.
- Insulating materials must use an RF energy source.

Sputtering is a versatile process since almost any material can be deposited by this technique, using both direct current and radio-frequency voltages. Sputtering works is ideal for materials with extremely high melting points such as carbon, silicon, and alloys.

e. Induction Evaporation

Induction evaporation uses radio-frequency radiation to evaporate the metal in a crucible. The metal is then deposited as with the other methods.

iv. Metal *reactive ion etching* (RIE) is commonly used to etch metal layers. This process uses a combination of physical sputtering and chemically reactive species for etching at low pressures. RIE uses ion bombardment to achieve directional etching and also a chemically reactive gas (carbon tetrafluoride, carbon tetrachloride, boron trichloride, and others) to maintain good etched layer selectivity. A wafer is placed into a chamber and given a negative electrical charge. The chamber is heated and taken to a low pressure, and then filled with a positively charged plasma of the reactive gas. The opposing electrical charges cause the rapidly moving plasma molecules to align themselves and strike the wafer surface vertically, thereby reacting with and volatizing the exposed metal layer. After etching, the remaining photoresist is stripped in a similar manner as that considered in section 1.10.

Table 1.7: Summary of Pros and Cons of metal layer deposition and evaporation methods

Method	Advantages	Disadvantages
E-Beam Evaporation	• high temperature materials • good for lift-off • highest purity	• some CMOS processes sensitive to radiation • alloys difficult • poor step coverage
Filament Evaporation	• simple to implement • good for lift-off	• limited source material (no high temperatures) • alloys difficult • poor step coverage
Sputter Deposition	• better step coverage of alloys • high temperature materials • less radiation damage	• possible grainy films • porous films • plasma damage and contamination
Flash Evaporation	• contamination free • composite layers	• shadowing or step coverage

v. Alloying and Annealing. After the metallised interconnections have been deposited and etched, a final step of alloying and annealing may be performed. To perform alloying, the metallised substrate is placed in a low-temperature diffusion furnace. Usually aluminium is placed in the furnace to form a low-resistance contact between the aluminium metal and silicon substrate. Finally, either during the alloy step or following it, the wafers are exposed to a gas mixture containing hydrogen in a diffusion furnace at 400 to 500°C. This annealing step is designed to optimize and stabilize the characteristics of the device by combining hydrogen with uncommitted atoms at or near the silicon-silicon dioxide interface.

1.16 Glassivation

Glassivation is the deposition of the final passivating layer on top of the die to protect it from mechanical damage and corrosion. This final layer is often composed of an amorphous insulating material, or glass.
Silicon nitride (Si_3N_4) is a common glassivating material because of its suitability for this purpose, as highlight by the features given in Table 1.3. Highly resistant to diffusion, it is almost impenetrable to moisture and ionic contaminants such as sodium, Na. It can also be deposited with a low residual compressive stress, making it less prone to delamination or cracking. Its interfacing with the underlying metal layers is conformal. Finally, it can be prepared with a low pinhole density. Since silicon nitride has a high dielectric constant, it is not popular as an interlayer dielectric for the simple reason that it results in a high inter-metal capacitance.

Silicon nitride can be deposited using plasma-enhanced chemical vapour deposition (PECVD) or low-pressure chemical vapour deposition (LPCVD), using silane and ammonia gases NH_3, although the former is the technique of choice for glassivation purposes because it allows a lower processing temperature.
Sometimes a layer known as 'p-glass' is deposited with a layer of $Si0_2$ doped with phosphorous. Phosphine gas is used as a source of phosphorous for this type of deposition.

1.17 Back side metallisation and die separation

A final processing step called back-lapping is sometimes performed. The backside of the wafer may be lapped or ground down using a wet abrasive solution under pressure. Backside metallization with a metal such as gold may be deposited on the back of the wafer with sputtering. This makes attachment of the separated die to the package easier in the final assembly.
After sorting and testing, the individual dies are physically separated. Diamond scribing, laser scribing, and diamond wheel sawing are used for die separation. Diamond scribing involves scoring a line across the wafer surface with a diamond tip. The wafer is then bent along the line, causing a fracture and separation. Laser scribing is similar except that a laser is used to score the fracture line. Diamond sawing involves wet-cutting the substrates with a high-speed circular diamond saw. Sawing may be used to either partially cut and scribe the surface, or can be used to completely cut through the wafer.

1.18 Wire bonding

Wire bonding is the process of electrically connecting the silicon to the package electrical pins or legs. Power devices use high purity aluminium rather than gold, Au, although copper is being increasingly used. Wire bond reliability depends mainly on the proper choice of the bonding wire.
The first choice consideration involves the type of package. Gold wire cannot be used in hermetic packages because it will not be able to withstand the high temperature of hermetic sealing. Aluminium wire is the standard choice for hermetic assembly. For plastic packages, however, gold wire is the more logical choice because it is faster, easier to use, and therefore more cost-effective.
The diameter of the wire is the next important consideration. Thinner wires will be required by circuits with smaller bond pad openings, while circuits that draw large currents or require thermo-mechanical robustness require thicker wires.
Another consideration when choosing a bonding wire is its tensile strength. The wire is subjected to tensile stresses throughout its lifetime, e.g., during bonding, during encapsulation, and during usage. The higher the tensile strength the better.
The elongation property of the wire is also an important consideration in the wire selection process. Wires with higher elongation are more difficult to control during loop formation at wire bonding. Thus, it is better to choose a wire that does not elongate much during the bonding process.
The last major consideration is the length of the heat-affected zone of the wire. When the end of the wire is melted by flame-off to form the free air ball prior to bonding, the high temperature enlarges the grain structures of the zone closest to the ball. These larger grain structures are more vulnerable to shearing stresses that cut across the wires. Wires with a longer heat-affected zone cannot be used in low-loop wire bonding because the heat-affected zone may be subjected to the shearing stresses of loop formation. Normally the wire manufacturer will indicate whether the wire is for low-loop or for high-loop applications.

Table 1.8. Properties of various Wire Types

(172.41 / resistivity = % IACS International Annealed Copper Standard)

Property		Cu	Au	Al	Ag
Electrical Conductivity	% IACS	103.1	73.4	64.5	108.4
Thermal Conductivity	W/m K	398.0	317.9	243.0	428.0
Thermal Expansion Coefficient	mm/m K	16.5	14.2	23.6	19.0
Tensile Elastic Modulus	GPa	115	78	62	71

Copper wire is becoming one of the preferred materials for wire bonding. Copper wire of smaller diameter can achieve the same performance as larger diameter gold wire. Copper wire is also more economical than gold wire.

With proper set-up, copper wire can be successfully wedge-bonded and can be used as an alternative to aluminium wire, especially in applications where higher current-carrying capacity is needed or complex geometry problems are encountered.

Copper wire is harder than gold and aluminium, so it has a greater tendency to contribute to die damage if the bonding parameters are not maintained under tight control. It is also inherent for copper to oxidize, which if left unchecked can lead to storage and shelf life issues.

Wire bonding methods
In the case of ICs, after separation into individual dies, the functional devices are attached with an epoxy material to a lead frame assembly. Once attached to the lead frame, electrical connections must be provided between the die and assembly leads. This is accomplished by attaching aluminium or gold leads via thermal compression or ultrasonic welding.

The three wire bonding methodologies, as shown in figure 1.22, are:
a Wire bonding – which is applicable to power devices
b Flip-chip bonding
c Tape-automated bonding

a. Wire bonding
As summarised in Table 1.9, Au or Aℓ wires are wire bonded between pads and the substrate using:
- Ultrasonic (100 to 500 microns diameter wire for power devices),
- Thermo-compression bonding or
- Thermo-sonic.

Ultrasonic bonding
Due to problems with thermo-compression bonding:
- Oxidation of Aℓ makes it difficult to form a good ball.
- In the case of ICs, epoxies cannot withstand high temperatures.

Ultrasonic is a lower temperature bonding alternative, which relies on pressure and rapid mechanical vibration to form bonds. The approach is:
i. The wire fed from a spool through a hole in the bonding tool
ii. Wire is lowered into position as 20-60 kHz ultrasonic vibration causes the metal to deform and flow.
iii. Tool is raised after the bond is formed.
iv. Clamp pulls and breaks wire.

Thermo-compression bonding
- Fine wire (15 to 75 μm diameter) fed from a spool through a heated capillary.
- H_2 torch or electric spark melts the wire end, forming a ball.
- Ball is positioned over the chip bonding pad, capillary is lowered, and ball deforms into a 'nail head'.
- Capillary raised and wire fed from spool and positioned over substrate; bond to package is a wedge produced by deforming the wire with the edge of the capillary.
- Capillary is raised and wire is broken near the edge of the bond.

Thermo-sonic bonding
- Combination of thermo-compression and ultrasonic
- Temperature maintained at approximately 150°C
- Ultrasonic vibration and pressure used to cause metal to flow to form weld
- Capable of producing 5 to10 bonds/s

Table 1.9: Three wire bonding processes

Wire bonding	Pressure	Temperature	Ultrasonic energy	Wire
Ultrasonic	Low	25°C	Yes	Au, Aℓ
Thermo-compression	High	300-500°C	No	Au,
Thermo-sonic	Low	100-150°C	Yes	Au

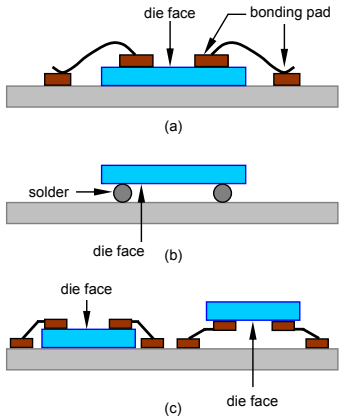


Figure 1.22. *Bonding methodologies: (a) wire; (b) basic flip-chip; and (c) tape-automated.*

b. Flip-chip bonding
Direct interconnection where the IC die is mounted upside-down onto a module or PCB, as shown in figure 1.22b. Connections are made via solder bumps located over the surface of die. The I/O density is limited only by the minimum distance between adjacent bond pads. A number of different flip-chip bonding techniques are shown in figure 1.23.
The bonding process, shown in figure 1.23c for stud bump bonding, is:
- Die are placed face down on the module substrate so that I/O pads on the chip are aligned with those on the substrate
- Solder reflow process is used to simultaneously form all the required connections,
- Drawback: bump fabrication process itself is fairly complex and capital intensive.
- Solderless flip-chip technology is another alternative; involves stencil printing of organic polymer onto a die.

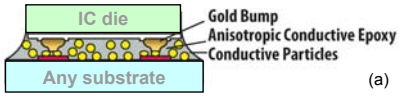
c. Tape-automated bonding
Tape-automated bonding is an interconnect technology between the substrate and the die, using a prefabricated carrier with copper leads adapted to the die pads instead of single wires.
This prefabricated carrier, usually a tape consists of a perforated polyamide film, like a camera film, and of the same dimensions, which has a transport perforation and stamped openings for the die and the connection leads. Copper foil is glued on this film, where the copper is structured by photolithography, forming a flexible circuit.
The advantage of this process is the creation of freestanding fingers, as seen in figure 1.22c, in the tape openings, which are then soldered or welded to bumps previously created on die pads (inner lead bonding). The mounted die can be burned-in, tested, and afterwards punched out from the tape. No mechanical protection is needed, the bumps sealing hermetically to the die, and the leads have a mechanical strength of about ten times the strength of a bonded wire.

Advantage:
- All bonds formed simultaneously, improving throughput.

Disadvantages:
- Requires multilayer solder bumps with complex metallurgy.
- A particular tape can only be used for a chip and package that matches its interconnect pattern.

Anisotropic Adhesive Attachment
(Z-axis conductive epoxy)

- Ideal for PCB and flex circuits
- High I/O
- Tight pitch
- Cost-effective flip chip solution
- Utilises off-the-shelf wire bondable ICs

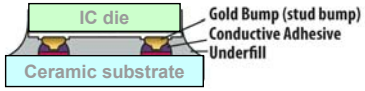


(a)

Stud Bump Bonding

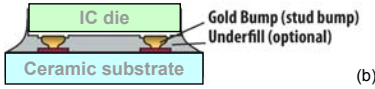
- Ideal for high I/O flip chip to ceramic substrate
- Mid-process replacement of faulty chips
- Underfill required
- Proven technology with reliability data
- Utilises off-the-shelf wire bondable ICs

(c)



Thermal-Sonic Bonding (gold to gold interconnect)

- Ideal for high frequency applications and MEMs to ceramic substrates
- I/O limited to ≈32 or less
- Underfill option
- Low temperature process
- Lead free

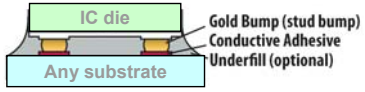


(b)

Solder Mounting

- Standard flip chip technology
- Solder bump devices
- Underfill option
- Z-axis control for ultimate strength
- High volume cost-effective solution

(d)



Process for Stud Bump Bonding (e)

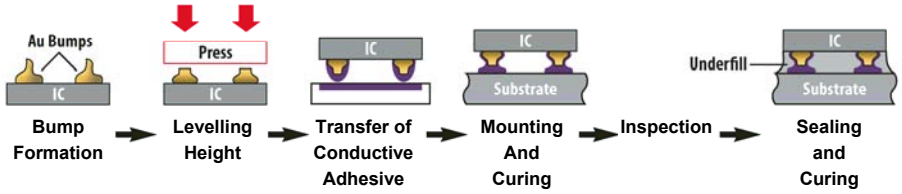


Figure 1.23. *Flip-chip bonding methodologies:*
(a) anisotropic adhesive attachment (Z-axis conductive epoxy);
(b) thermal-sonic bonding (gold-to-gold interconnect); (c) stud bump bonding;
(d) solder mounting; and (e) full process for stud bump bonding.

1.19 Types of wafer silicon

Silicon is the most common material used for semiconductors. After oxygen, silicon is the second-most abundant element in the earth's crust. It is not poisonous, and it is environment friendly, its waste does not represent any problems. However, to be useful as a semiconductor material, silicon must be refined to a chemical purity of better than 99.9999%. Pure silicon is not a natural state but is refined from various silicon dioxides $Si0_2$ such as quartzite gravel, which is the purest silica, or crushed quartz, silicates.

1.19.1 Purifying silicon

Silicon dioxide of either quartzite gravel or crushed quartz is placed in an electric arc furnace. A carbon arc is then applied to release the oxygen, at temperatures over 1800°C, leaving the products carbon dioxide and molten silicon. The reduction to silicon is via the reaction:

$$SiO_2 + SiC \rightarrow Si + SiO + CO$$
$$SiO_2 + 2C \rightarrow Si + 2CO \quad (1.31)$$

This process yields a metallurgical-grade, medium-grey metallic looking, 99% pure silicon.
Silicon with a one percent impurity is not useful in the semiconductor industry. Impurities of the order 10^{14} make major changes in the electrical behaviour of silicon. Since there are about $5x10^{22}$ atoms/cm^3 in a silicon crystal, a purity of better than 1 part in 10^8 or 99.999999% pure material is needed. Next the silicon is crushed and reacted with HCℓ gas, in the presence of copper-containing catalyst, to make trichlorosilane, a high vapour pressure liquid that boils at 31.8°C as in:

$$Si + 3HCl_{(gas)} \xrightarrow{Cu} SiHCl_3 + H_2 \quad (1.32)$$

Many of the impurities in the silicon (aluminium, iron, phosphorus, chromium, manganese, titanium, vanadium, and carbon) also react with the HCℓ, forming various chlorides, which are highly reactive. Each of these chlorides have different boiling points, so by fractional distillation it is possible to separate out the $SiHC\ell_3$ from most of the impurities. The pure trichlorosilane is then reacted with hydrogen gas at an elevated temperature of about 1100°C to form pure electronic grade silicon (1 pp10^{-9} of impurities).

$$SiHCl_3 + H_2 \rightarrow 2Si + 3HCl \quad (1.33)$$

Although the resultant silicon is relatively pure, it is in a polycrystalline form that is not suitable for semiconductor device manufacture. This so-called electronic grade polysilicon, EGS, requires further processing. The presented example is one way of producing pure silicon. There are other production procedures with different chemical reactions used, yet the end-product is the same - pure silicon.
The next step in the process is to grow single crystal silicon from the polycrystalline silicon via one of three methods, as considered in section 1.19.3.

- the Czochralski process;
- the float-zone process or
- the ribbon silicon process.

In *single-crystal silicon*, the molecular structure, which is the arrangement of atoms in the material, is uniform and defect free, because the entire structure is grown from the same crystal. This uniformity is ideal for transferring electrons efficiently through the material. To make an effective semiconductor, however, silicon is doped with other elements to make it n-type or p-type.
Multi-crystalline silicon, in contrast, consists of several smaller crystals or grains, which introduce boundaries. These boundaries impede the flow of electrons and encourage them to recombine with holes, thereby reducing the efficiency of the silicon. However, multi-crystalline silicon is much less expensive to produce than single-crystalline silicon.

1.19.2 Crystallinity

The crystallinity of a material indicates how perfectly ordered the atoms are in the crystal structure. Silicon, as well as other semiconductor materials, can come in various crystalline forms:

- single-crystalline,
- multi-crystalline,
- polycrystalline or
- amorphous.

Table 1.10: Types of crystalline silicon and formation processing methods

Type of Silicon	abbreviation	Crystal Size Range	Deposition Method
single-crystal silicon	sc-Si	> 10cm	Czochralski, float-zone
Multicrystalline silicon	mc-Si	1mm-10cm	Cast, sheet, ribbon
polycrystalline silicon	pc-Si	< 1mm-1mm	Chemical-vapour deposition
microcrystalline silicon	μc-Si	< 1mm	Plasma deposition

In a single-crystal material, the atoms making up the structure of the crystal are repeated in a regular, orderly manner from layer to layer. In contrast, a multi-crystalline material is composed of numerous smaller crystals, with the orderly arrangement disrupted when moving from one crystal to another. Multi and poly crystalline silicon are particularly important in areas like photovoltaic cells, light-emitting, and laser diodes, when cost is an important factor. One classification scheme for silicon uses approximate crystal size and methods typically used to grow or deposit such material, as shown in Table 1.10.

1.19.3 Single crystal silicon

Several different processes can be used to grow an ingot or boule of single or mono-crystal silicon. The most established and dependable processes are the *Czochralski method* and the *float-zone technique*. The *ribbon-growth* technique is used for lower cost and quality silicon crystal growth.

1.19.3i *Czochralski process*

The most commonly used process for creating the boule is called the *Czochralski* method, as illustrated in figure 1.24a. Electronic grade polysilicon silicon is heated in a quartz crucible to 1400°C in an argon atmosphere, using RF or resistance heating. A starter seed of single-crystal silicon on a puller contacts the top surface of molten polycrystalline silicon at 1415°C to 1420°C. As the seed crystal is slowly withdrawn - pulled and rotated, if the temperature gradient of the melt is adjusted so that the melting/freezing temperature is just at the seed-melt interface, atoms of the molten silicon solidify in the pattern of the seed and extend its single-crystal structure, forming a cylindrical boule of near perfect, pure silicon.

- The ingot pull is unusually pure, because impurities either burn or tend to be drawn into the liquid silicon. An argon atmosphere precludes any oxygen impurity. The rod and crucible are rotated in opposite directions to minimise the effects of convection in the melt. The pull-rate (1µm to 1mm/s), the rotation-rate (10 to 40 turns per minute), and the temperature gradient are carefully optimised for a particular wafer diameter (up to 30mm) and lattice structure orientation growth direction (direction <111>, along the diagonal of the sides of the cube crystal structure, for bipolar devices). Lengths of boule of several metres are attainable.
- A small amount of boron (or phosphorous) is usually added during the Czochralski process to pre-dope the substrate silicon.

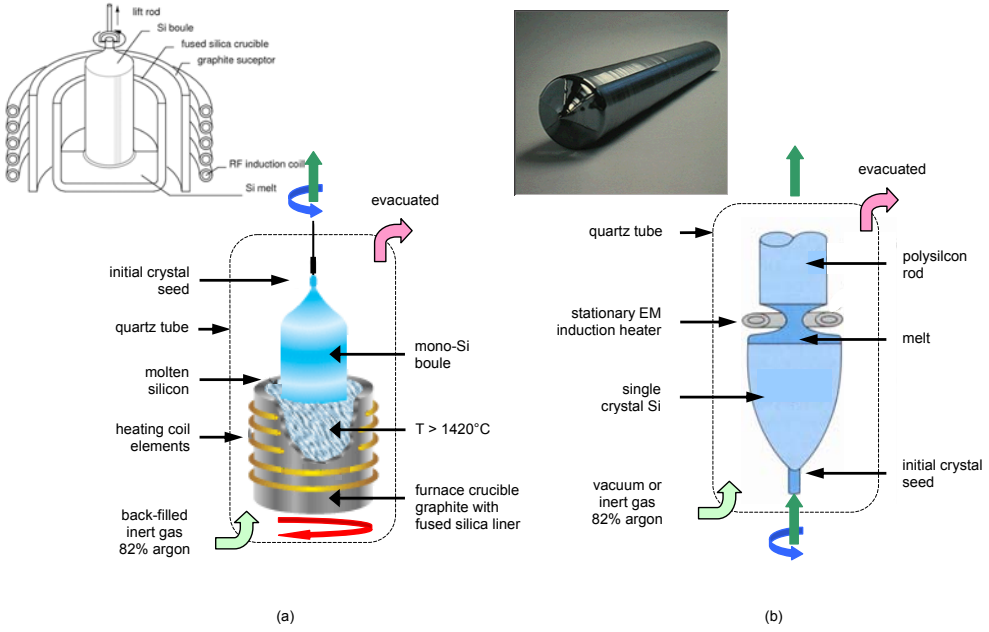


Figure 1.24. *Techniques for producing a single crystal silicon boule: (a) Czochralski method and (b) the float-zone method.*

1.19.3ii *Float-zone process*

The *float-zone process* produces purer crystals than the Czochralski method, because the pull is not contaminated by the crucible used in growing Czochralski crystals. In the float-zone process, as illustrated in figure 1.24b, a polycrystalline silicon rod is set atop a seed mono-crystal and then effectively lowered through an electromagnetic induction coil. The coil's magnetic field induces an electric field in the rod, heating and melting the interface between the rod and the seed. Single-crystal silicon forms at the interface, growing upward as the rod is slowly raised. In this floating zone technique, the molten silicon is unsupported, maintaining itself through surface tension.
Rods of mono-silicon measure up to 10cm in diameter and 1m in length.

Wafer preparation

The next step is the same for both single-crystal formation methods. The boule ends are cropped using a water-lubricated, single-blade diamond saw. The ingot is then ground to a uniform diameter in a lathe, and each end is bevelled with a sand belt to reduce the possibility of shattering the ingot. X-ray diffraction can then be used to determine the crystal structure orientation, which is marked by grinding the length of the cylindrical side of the boule.
The cylindrical single-crystal ingot is sawed, using a multi-blade, inner-diameter saw in conjunction with a wet lubricant, into thin wafers for further processing. The sawing wastes 20% to 50% of the silicon as sawdust, known as *kerf*.
The sliced wafers are mechanically lapped under pressure using a counter-rotating machine to achieve flatness and parallelism on both wafer sides. Most lapping operations use slurries of either aluminium oxide or silicon carbide. The edges of the individual wafers are also rounded by wet automatic grinders.
After lapping, the wafers are etched with a solution containing nitric, acetic, and hydrofluoric acids (HF, CH_3C00H, and $HN0_3$). This etching process removes external surface damage and reduces the thickness of the wafer.
Next, the wafers are polished using an aqueous mixture of colloidal silica and sodium hydroxide. The wafers are mounted onto a metal carrier plate that is attached by vacuum to the polishing machine. The chemical polishing process usually involves two or three grinding and polishing steps with progressively finer slurry, which decreases wafer thickness and results in a mirror-like lustre finish. Sometimes carrier pads must be stripped from the metal carrier plates. The pads are usually stripped with solvents such as methylene chloride, methyl ethyl ketone or a glycol ether mixture.
Finally, the wafers are cleaned to remove any particles or residue remaining on the exterior surface of the polished wafer. Various cleaning steps and solutions containing ammonia, hydrogen peroxide, hydrofluoric acid, hydrochloric acid (NH_3, H_2O_2, HF, and HCℓ), and deionised water may be used.
The finished wafers are inspected and packaged for shipping, since most semiconductor manufacturers purchase wafers from specialist wafer producers.

1.19.3iii *Ribbon silicon*

Although single-crystal silicon technology is well developed, the Czochralski and float-zone processes are complex and expensive, as are the ingot-casting processes discussed under multi-crystalline silicon. Another crystal-producing process is ribbon silicon growth, where the single crystals cost less than from other processes, because they form the silicon directly into thin, usable wafers of single-crystal silicon. By forming thin crystalline sheets directly, sawing and slicing steps of cylindrical boules are avoided.
One ribbon growth technique, termed edge-defined film-fed growth, starts with two crystal seeds that grow and capture a sheet of material between them as they are pulled from a source of molten silicon. A frame entrains a thin sheet of material when drawn from a melt. This technique does not waste much material, but the quality of the material is not as high as Czochralski process and float zone produced silicon. The resultant silicon quality is inferior for large-area, high-voltage, power semiconductor switching devices.

1.19.4 Multi-crystalline Silicon

Multi-crystalline (or poly-crystalline) silicon describes when the active portion of the silicon comprises several relatively large crystals, called grains, up to a square centimetre or so in area.
Having several large crystals in a cell introduces a problem. Charge carriers can move around relatively freely within one crystal, but at the interface between two crystals, called the grain boundary, the atomic order is disrupted. Free electrons and holes are much more likely to recombine at grain boundaries than within a single crystal.

There are several ways to minimize the problems caused by grain boundaries:

- adjusting growth conditions through treatments such as annealing (heating followed by a slow cooling rate stage) the semiconductor material so that grains are columnar and as large as possible. The impurities are also better distributed;
- designing cells so that the charge carriers are generated within or close to the built-in electric field; and
- filling broken bonds at grain edges with elements such as hydrogen or oxygen, which is called passivating the grain boundaries.

Multi-crystalline silicon based devices are generally less efficient than those made of single-crystal silicon, but they can be less expensive to produce. Multi-crystalline silicon is produced in a variety of ways.

- The most common commercial methods involve a casting process in which molten silicon is directly cast into a mould and allowed to slowly solidify into an ingot. The starting material can be a refined lower-grade silicon, rather that the higher-grade semiconductor grade required for single-crystal material. The mould is usually square, producing an ingot that can be cut and sliced into square cells, minimising wasted silicon.
- The procedure of extracting pure multi or poly-crystalline silicon from tri-chlorine-silane can be (among others) performed in special furnaces. Furnaces are heated by electric current, which flows through (in most cases) silicon electrodes. The 2m long electrodes measure 8mm in diameter. The current flowing through electrodes can reach up to 6000A. The furnace walls are additionally cooled preventing the formation of any unwanted reactions due to gas side products. The procedure results in pure poly-crystalline silicon used as a raw material for solar cell production. Poly-crystalline silicon can be extracted from silicon by heating it up to 1500°C and then cooling it down to 1412°C, which is just above solidification of the material. The cooling is accompanied by the origination of an ingot of fibrous-structured poly-crystalline silicon of dimensions 40x40x30 cm. The structure of poly-crystalline silicon in regions of the material is uniform, yet it is not conformal to the structure in other parts.

1.19.5 Amorphous Silicon

Amorphous silicon is produced in high frequency furnaces in a partial vacuum atmosphere. In the presence of a high frequency electrical field, gases like silane, B_2H_6 or PH_3 are blown through the furnaces, supplying the silicon deposit with boron and phosphorus.
Amorphous solids, like common glass, are materials whose atoms are not arranged in any particular order. They do not form crystalline structures, and they contain large numbers of structural and bonding defects. Economic advantages are that it can be produced at lower temperatures and can be deposited on low-cost substrates such as plastic, glass, and metal. These characteristics make amorphous silicon the leading thin-film material.
Since amorphous silicon does not have the structural uniformity of single or multi crystalline silicon, small structural deviations in the material result in defects such as *dangling bonds*, where atoms lack a neighbour to which they can bond. These defects provide sites for electrons to recombine with holes, rather than contributing to the electrical circuit. Ordinarily, this kind of material would be unacceptable for electronic devices, because defects limit the flow of current. However, amorphous silicon can be deposited so that it contains a small amount of hydrogen, 5% to 10%, in a process called hydrogenation. The result is that the hydrogen atoms combine chemically with many of the dangling bonds, as shown in figure 1.25, essentially neutralising or removing them and permitting electrons to move through the material.

Staebler-Wronski Effect
Instability currently retards amorphous silicon exploitation in some semiconductor applications. In the case of photo-voltaic cells, the amorphous cells experience an electrical output decreases over a period of time when first exposed to sunlight. The electrical output stabilizes with a net output loss of 20%. The reason is related to the amorphous hydrogenated nature of the material, including tiny microvoids or atomic-level gaps in the amorphous silicon structure several angstroms in diameter (1 angstrom =10^{-10} m). Other causes include oxygen or carbon impurities that are in the cells and ordinary stresses in the system that break silicon-silicon bonds in the region of the imperfections.

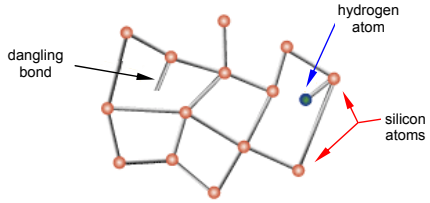


Figure 1.25. *Amorphous silicon showing the dangling bonds and hydrogen sites.*

Devices suffering from light induced degradation can recover their effectiveness if they are annealed at 150°C for a few minutes. Annealing is also effective at the normal operating temperatures of silicon, about 50° to 80°C. This is called self-annealing.

1.20 Silicon Carbide

Wide bandgap semiconductors (GaN (III-V), SiC, diamond, etc.) have better high voltage and temperature characteristics than silicon devices. However, because silicon carbide, SiC, sublimes at high temperature, ≈1800ºC, processing is more difficult than for silicon (which melts at a lower temperature of 1415°C). The similar chemistry properties of silicon and silicon carbide (both in group IV) means that many of the existing processes for silicon can be applied to silicon carbide, but with some refinement and higher processing temperatures. The exception is thermal diffusion which is not effective if a good SiC surface morphology is to be retained.

The SiC crystal boules are grown by seeded sublimation using the physical vapour transport (PVT) method. Alternatively, chemical vapour deposition (CVD) can be used, where SiH_4, C_3H_8, and H_2 are typically injected into the chamber. This process is mainly used for producing SiC epitaxial growth. A hot walled CVD reactor can deposit 100µm at a rate of 1 to 5 µm/hour at 1200ºC to 1500ºC. Crystal defects (micropipes, stack faults, etc.) occur at a rate of less than 1 per cm^2. Proprietary defect healing technology can significantly decrease the defect rate. The main single crystal polytypes for power switching device fabrication are 4H-SiC and 6H-SiC (this lattice structure terminology is based on the Ramsdell notation).

Nitrogen for n-type and aluminium or boron for p-type can be used in epitaxial growth and ion implantation. Substrates usually have an n or p epitaxial drift layer. Typical n-type epitaxy (50µm) can be thicker than a p-type layer (10µm), and the n-type epitaxy has a thin 1µm n-type buffer or fieldstop.

Ion implantation is shallow, typically less than 1µm, and requires high temperature and 30 to 300keV. Subsequent annealing in argon is at 1650ºC in the presence of a silicon over pressure. The lower the temperature, the longer the annealing time. Contact metallization can use sintered nickel on highly n-doped SiC, which is annealed at 1150ºC for a few minutes (with deuterium) (then overlaid with thick gold). A nickel, aluminium, and titanium Schottky metal sintered combination is suitable for p^+ region metallization, which is overlaid with gold. Aluminium metallisation is used at the Schottky diode anode.

SiO_2 is an electrical-insulator that can be grown on both Si and SiC. Oxide growth for SiC is slower than that on silicon and involves nitridation of nitric oxide, N_20, at 1300ºC. Because of the physical and chemical stability of silicon carbide, acid wet etching is ineffective and dry reactive ion etching tends to be used for etching processes.

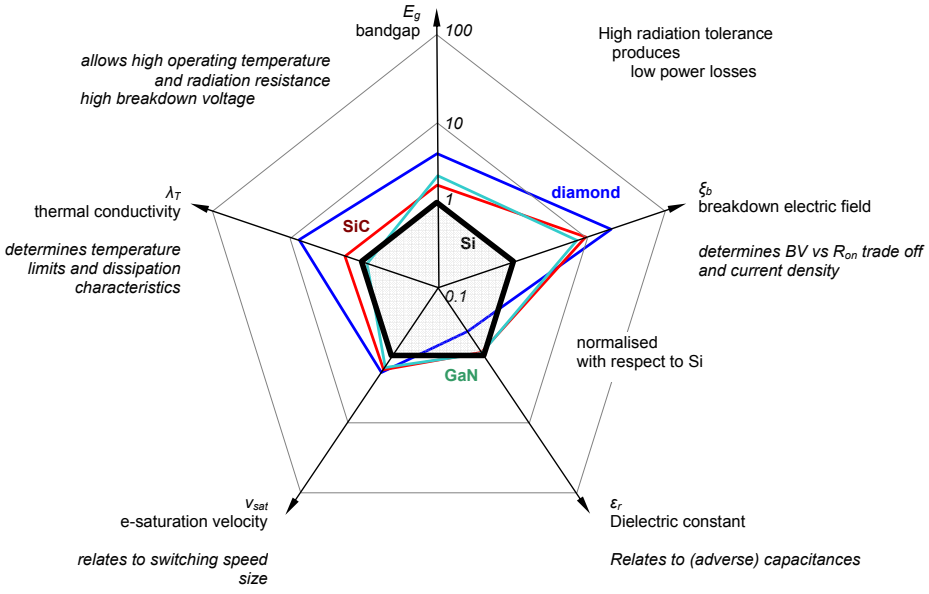


Figure 1.26. *Key electrical and thermal normalised characteristics of group IV monocrystalline silicon, diamond, gallium nitride, and 3 polytypes of silicon carbide, at room temperature.*

1.21 Si and SiC physical and electrical properties compared

The processing of silicon is a mature, cost efficient technology, with 300mm wafers and submicron resolution established within the microelectronics industry. So-called wide bandgap semiconductors like silicon carbide (processed on 100mm wafers) offer promising high voltage and temperature power switching device possibilities as material quality and process yields improve.
The non-repetitive peak current through SiC bipolar junctions is limited due to the significant positive temperature coefficient of the forward voltage drop. Therefore device size has to be chosen large enough to avoid over-current destruction. Avalanche capability is not better than Si. SiC bipolar current flow can lead to defect growth and finally to the destruction of the device, as experienced by pin diodes.
Figure 1.26 shows and allows normalised, with respect to silicon, comparison of the key physical and electrical properties of the main semiconductor materials applicable to power switching device fabrication.

material	Bandgap Energy	Dielectric Constant	Electron/ Hole Mobility	Breakdown Electric Field	Saturated Electron Drift Velocity	Thermal Conductivity	Figure of Merit *w.r.t* Si	Coefficient of Linear Thermal Expansion
	E_g	ε_r	μ_n / μ_p	ξ_b	v_{sat}	λ_T	*FoM*	α
	eV	pu	cm^2/Vs	MV/cm	10^7 cm/s	W/mK	$\lambda_T \times (\xi_b \times v_{sat})^2$	x 10^{-6} K^{-1}
Si	1.12	11.9	1400 / 450	0.30	1.0	150	1	2.6
GaAs	1.43	13.1	8500/400	0.46	1.0	46		6.86
GaN	3.45	9.0	1000 / 350	2.0	2.5	110	407	5.6
3C-SiC	2.36	9.72	800 / 320	1.3	2.7	360	2381	2.8
4H-SiC	3.26	10.1	900 / 120	2.2	2.0	370	3241	5.2
6H-SiC	3.03	9.66	400 / 90	2.5	2.0	490	1307	5.2
diamond	5.45	5.5	2200 / 1800	10.0	2.7	2200	54000	0.8

The higher

- the energy bandgap, E_g, the higher the possible operating temperature before intrinsic conduction mechanisms produce adverse effects;
- the avalanche breakdown electric field, ξ_b, the higher the possible rated voltage;
- the thermal conductivity, λ, the more readily heat dissipated can be removed; and
- the saturation electron drift velocity, v_{sat}, and the electron mobility, μ_n, the faster possible switching speeds.

Although the attributes of wide bandgap materials are evident, processing is more difficult than with Si and some of the parameters vary significantly with a wide operating (and processing) temperature range. SiC performance characteristic figures are slightly better than those for GaN, except, importantly, GaN has better carrier mobility. GaN growth is complicated by the fact that nitrogen tends to revert to the gaseous state, and therefore only thin layers are usually grown on sapphire or SiC substrates. Lattice-substrate boundary misfit occurs because of the significant difference in molecule sizes and packing. This limitation is more accentuated with GaN on silicon. There is a 17% misfit in molecule package and a 56% mismatch in thermal expansion ($\alpha_{GaN} = 5.59\times10^{-6}$ and $\alpha_{Si} = 3.59\times10^{-7}$ @ 300K). To prevent cracking during processing cooling, an intermediate transition layer like AℓN, is introduced. GaN does not have a native oxide. A low thermal conductivity does not bode well from high power, high temperature power devices. Wide band gap based, low-voltage (<100V) GaN, lateral RF transistors are viable. The GaN direct band structure and short carrier lifetimes tend to imply limitation to majority carrier devices, that is, high voltage bipolar devices are problematic. But boundary imperfections may prove problematic with high-voltage, power devices where the principle current flow is usually vertically through the structure, hence through the imperfect mechanical and electrical lattice boundary. In the case of GaN, ohmic contacts involve annealing Ti/Aℓ/Ni/Au, while Schottky contacts can be Ni/Au. SiN_x passivation is deposited by plasma enhanced CVD.

Some of the physical and electrical parameters and their values in figure 1.26 will be explained and used in subsequent chapters. Other useful substrate data is given in Table 1.11.

Table 1.11: Other useful substrate material data

parameter			Si	SiC	GaN	Diamond
maximum operating temperature	T_{max}	°C	300	1240		1100
melting temperature	T_{melt}	°C	1415	sublime >1800	2500	phase change
density	ρ	g / cm^3	2.33	3.17-3.21	6.15	3.52
electrical resistivity	ρ_e	Ω m	10^{-3}	10^6	10^{-3}	$>10^{11}$

Reading list

Streetman, B. G. and Banerjie, S. K., *Solid State Electronic Devices*, Prentice-Hall International, 6th Edition, 2005.

Van Zeghbroeck, B., *Principles of Semiconductor Devices*, http://ece-www.colorado.edu/~bart/book/

Zetterling, C. M., *Process technology for Silicon Carbide devices*, IEE, 2002.

http://www.siliconfareast.com/

http://www.semiconductorglossary.com/

Chemical Reactions associated with Wafer Fabrication

Fab Area	Chemical Reaction	Reaction Equation(s)	Comments
Epitaxy	Hydrogen reaction of $SiCl_4$ to deposit a silicon epitaxial layer	$SiC\ell_4 + 2\ H_2 \rightarrow Si + 4\ HC\ell$	reversible process Si may also be etched using HCℓ
Epitaxy	Silane (SiH_4) reaction to deposit a silicon epitaxial layer	$SiH_4 \rightarrow Si + 2H_2$	can be done at a relatively lower temperature
SiO_2 Thermal Oxidation	Silicon dioxide (SiO_2) deposition through dry thermal oxidation	$Si + O_2 \rightarrow SiO_2$	deposition temperature usually between 700-1300°C
SiO_2 Thermal Oxidation	Silicon dioxide (SiO_2) deposition through wet thermal oxidation	$Si + H_2O \rightarrow SiO_2 + 2H_2$	deposition temperature usually between 700-1300°C
SiO_2 CVD	SiO_2 deposition through CVD reaction between silane (SiH_4) and 0_2	$SiH_4 + O_2 \rightarrow SiO_2 + 2H_2$	low-temperature deposition process
SiO_2 CVD	SiO_2 deposition through PECVD reaction between silane (SiH_4) and N_20	$SiH_4 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2H_2$	deposition temperature usually between 200-400°C
SiO_2 CVD	SiO_2 deposition through LPCVD reaction between dichlorosilane ($SiH_2C\ell_2$) and N_20	$SiH_2C\ell_2 + 2\ N_2O \rightarrow SiO_2 + 2N_2 + 2HC\ell$	high deposition temperature of almost 900°C
Si_3N_4 CVD	Silicon nitride (Si_3N_4) deposition through LPCVD reaction between dichlorosilane ($SiCl_2H_2$) and ammonia (NH_3)	$3\ SiC\ell_2H_2 + 4\ NH_3 \rightarrow Si_3N_4 + 6H_2 + 6\ HC\ell$	deposition temperature usually between 700-800°C
Si_3N_4 CVD	Si_3N_4 deposition through PECVD reaction between silane (SiH_4) and NH_3	$SiH_4 + NH_3 \rightarrow Si_xN_yH_z + H_2$	deposition temperature usually between 200-350°C
Poly-Si CVD	Polysilicon deposition through LPCVD reactions of SiH_4	SiH_4 + surface site → SiH_4 (adsorbed); SiH_4 (adsorbed) → $SiH_2 + H_2$; $SiH_2 \rightarrow Si + H_2$	deposition temperature usually between 580-650°C; over-all reaction: $SiH_4 \rightarrow Si + 2H_2$
W CVD	Tungsten (W) deposition through CVD reaction between WF_6 and Si	$WF_6 + 3\ Si \rightarrow 2W + 3\ SiF_4$	deposition done between 200-400°C
W CVD	Tungsten (W) deposition through CVD reaction between WF_6 and H	$WF_6 + 3\ H_2 \rightarrow W + 6HF$	deposition done between 250-500°C
WSi_2 CVD	Tungsten silicide (WSi_2) deposition through CVD reaction between WF_6 and SiH_4	$WF_6 + 2\ SiH_4 + \rightarrow WSi_2 + 6HF + H_2$	deposition at an elevated temperature
$TiSi_2$ CVD	Titanium silicide ($TiSi_2$) deposition through CVD reaction between $TiCl_4$ and SiH_4	$TiC\ell_4 + 2\ SiH_4 + \rightarrow TiSi_2 + 4HC\ell + 2H_2$	deposition at an elevated temperature
Si Wet Etching	Removal of Si through wet etching with $HN0_3$ and HF	$Si + HNO_3 + 6\ HF \rightarrow H_2SiF_6 + HNO_2 + H_2 + H_2O$	acetic acid is preferred as a buffering agent
SiO_2 Wet Etching	Removal of $Si0_2$ through wet etching with HF	$SiO_2 + 6\ HF \rightarrow SiF_6 + H_2 + 2H_2O$	usually at room temperature to prevent HF attack of Si