Chapter 3

UEEA2223/UEEG4223
Integrated Circuit Design

The Manufacturing Process
Chapter 3 The Manufacturing Process

3.0 Introduction

3.1 Crystal Growth

3.2 Integrated Circuit Fabrication
  3.3.1 Wafer Cleaning
  3.3.2 Oxidation Process
  3.3.3 Lithography
  3.3.4 Etching
  3.3.5 Polycrystalline Process
  3.3.6 Diffusion
  3.3.7 Ion Implantation
  3.3.8 Metallization
  3.3.9 Sintering/Annealing

3.3 Fabrication of Bipolar Junction Transistor

3.4 Fabrication of CMOS Transistor

3.5 Packing Integrated Circuit
  3.5.1 Packaging Materials
  3.5.2 Interconnect Level
    3.5.3.1 Interconnect Level 1 - Die-to-Package-Substrate
    3.5.3.2 Interconnect Level 2 - Package Substrate to Board
    3.5.3.3 Multi-Chip Modules - Die to Board
  3.5.3 Thermal Consideration in Packaging

3.6 Pictorial View of Wafer Preparation and Wafer Fabrication

Exercises

Bibliography
Figure 3.1: Basic tetrahedral structure........................................................................50
Figure 3.2: Diamond structure..................................................................................50
Figure 3.3: Wafer flat for identifying the orientation and dopant type.......................51
Figure 3.4: A Czochralski silicon rod puller...............................................................51
Figure 3.5: Equilibrium segregation coefficients for dopant in silicon.........................54
Figure 3.6: Doping distribution near the solid-molten interface....................................55
Figure 3.7: The number of particle and diameter of particle for various classes of cleanliness..................................................................................................................57
Figure 3.8: Cleaning step and composition of RCA solution...........................................58
Figure 3.9: Cross section depicting the process sequence for obtaining semi-recessed and fully recessed LOCOS..........................................................................................60
Figure 3.10: Process step showing shallow trench isolation structure...........................60
Figure 3.11: Process step of deep filled trench with LOCOS...........................................61
Figure 3.12: Chemical mechanical polishing technique................................................62
Figure 3.13: The fabrication process steps of junction isolation bipolar junction transistor66
Figure 3.14: Step by step fabrication of n, p, n bipolar transistor..................................67
Figure 3.15: CMOS process steps...............................................................................68
Figure 3.16: n-well implant.........................................................................................69
Figure 3.17: n-well........................................................................................................69
Figure 3.18: Active area definition...............................................................................69
Figure 3.19: Substrate etch for FOX region....................................................................70
Figure 3.20: Gate oxide growth...................................................................................70
Figure 3.21: Polysilicon gate formation.........................................................................70
Figure 3.22: LDD implant of p-MOS transistor.............................................................71
Figure 3.23: LDD implant of n-MOS transistor.............................................................71
Figure 3.24: Oxide deposition......................................................................................72
Figure 3.25: Etchback and spacer formation...............................................................72
Figure 3.26: Deep p+ implant......................................................................................72
Figure 3.27: Deep n+ implant......................................................................................73
Figure 3.28: Oxide cuts and plugs.................................................................................73
Figure 3.29: Contact plug fill and metallization 1.........................................................74
Figure 3.30: Metal 1 oxidating.....................................................................................74
Figure 3.31: Oxide and planarization............................................................................75
Figure 3.32: High level metal layers and vias...............................................................75
Figure 3.33: Top metal and passivation layer...............................................................76
Figure 3.34: Rent’s constant for varies class of chip and system Figure........................77
Figure 3.35: Interconnect hierarchy in traditional integrated circuit packaging.............79
Figure 3.36: Wiring bonding connecting pad and lead..................................................79
Figure 3.37: Typical conductance and inductance of package type and wire...............80
Figure 3.38: Automated tap bonding (a) polymer with imprinted wire pattern and (b) die attaché using solder bump...............................................................80
Figure 3.39: Flip-chip bonding...................................................................................81
Figure 3.40: Printed circuit board mounting approach. (a) through-hole mounting and (b) surface mounting..................................................................................................................81
Figure 3.41: Commonly use package (1) leadless carrier, (2) DIP, (3) PGA, (4) small outline IC, (5) quad flatpack, and (6) PLCC..........................................................82
Figure 3.42: Parameters of various chip carriers...........................................................83
Figure 3.43: Ball grid array packaging; (a) cross-section, (b) photo of PGA bottom........83
Figure 3.44: An Avionics processor module. Courtesy of Rockwell International ........... 84
Figure 3.45: Silicon ingot ................................................................................................. 87
Figure 3.46: Cleaning room attire .................................................................................. 87
Figure 3.47: Cleaning station ........................................................................................... 88
Figure 3.48: Epitaxial furnace ......................................................................................... 88
Figure 3.49: Wafer mask and 10X reticle ......................................................................... 89
Figure 3.50: Lithography – pattern generator ................................................................. 89
Figure 3.51: Lithography – step and repeat machine ...................................................... 90
Figure 3.52: Diffusion furnace ......................................................................................... 90
Figure 3.53: Ion implanter .............................................................................................. 91
Figure 3.54: Etch station ................................................................................................. 91
Figure 3.55: Wafer with fabricated integrated circuit ..................................................... 92
Figure 3.56: Chemical-mechanical polishing machine .................................................. 92
Chapter 3

The Manufacturing Process

3.0 Introduction

Jack Kilby was the first person developed miniaturized transistor circuit in 1958. It was then followed by Robert Noyce and Gordon Moore, who built first planar miniaturized transistor in 1960. Thereafter with the aid of computer and advancement in photolithography, integrated circuit IC was fast developed into ultra large scale integration ULSI where millions of transistors can be built in a silicon chip of 3.0cm x 3.0cm area.

Integrated circuit is mainly built on silicon due to cheap, developed processes, and abandon on the earth crust. Beside these reasons, the by-product produced by the process is not toxic and can be disposed easily. There are many other available materials that can be used to build integrated circuit such GaAs, SiC, GaN, SiGe and etc. Unless, it is mentioned, this lecture shall be presented using silicon as the building element for the integrated circuit.

An integrated circuit IC consists of a single crystalline chip or it may be called monolithic, typically 300µm to 600µm thick and covering a surface area of 3.0cm x 3.0cm containing both active and passive elements. The process used to fabrication an IC covers wafer preparation, photolithography, epitaxial growth, impurity diffusion, ion implantation, oxidation, etching, metallization deposition, and annealing.

Today there are many fabrication technologies developed. These technologies are used to fabricate silicon based integrated circuit. Among the common types are the bipolar junction isolation JI, dielectric isolation DI, self align junction isolation SAJI, oxide isolation OI, vertical dual MOS, SiGe, complimentary metal oxide semiconductor CMOS, BiCMOS, and etc. We will briefly discuss some of them and some component fabrications.

Semiconductor such germanium, silicon, and gallium arsenide have a lattice structure, which consists of two interlocking face-centered cubic lattices, and have total of eight atoms per unit cell. This structure can be broken down into primitive cell of tetrahedral shape shown in Fig. 3.1 and diamond structure
3 The Manufacturing Process

shown in Fig. 3.3. Diamond structure that has different atom type such as gallium arsenide is called *zinc blende*.

![Figure 3.1: Basic tetrahedral structure](image1)

![Figure 3.2: Diamond structure](image2)

The most common two types of silicon crystal orientations used to fabricate the integrated circuit are (111) and (100) types. The selection of the type depends on the type circuit to be built. (111) orientation is widely used for fabricating bipolar device, whilst (100) type is mainly used to fabricate MOSFET. (100)-type has the surface state charge 30-40% less than (111) type. Thus, it is ideal to use for MOSFET fabrication in which the device is sensitive to surface trapping and thereby affecting the mobility of the device. Bipolar device is a current driven device. Thus, surface state charge would not have significant impact on the speed of the device. Moreover, choosing the right orientation would assist
3 The Manufacturing Process

scribing in dicing of the silicon wafer. Figure 3.3 shows the way to identify orientation and dopant type.

![Diagram of wafer flats](image)

Figure 3.3: Wafer flat for identifying the orientation and dopant type

3.1 Crystal Growth

Before the fabrication of the integrated circuit, the preparation of silicon or gallium arsenide is required. The preparation of wafer involves several process steps. They are distillation and reduction/synthesis, crystal growth, grind/saw/polish, and electrical and mechanical characterizations. We shall not discuss the process of making gallium arsenide GaAs crystal. We shall concentrate on the process of making silicon crystal.

The starting materials are silicon dioxide for silicon and gallium and arsenic for gallium arsenide wafer. They are chemically processed to form high-purity crystal polycrystalline semiconductor for which single crystal is formed. The single crystal ingot is shaped to define diameter and is sawed into wafer. The wafer is then etched and polished to provide smooth, specular surface where device is fabricated.
3 The Manufacturing Process

Pure form of sand SiO$_2$ called *quartzile* is placed in a furnace with various forms of carbon like coke, coal, and wood chip. Although there are numbers of reaction take place, the overall reaction follows equation (3.1).

$$\text{SiC} + \text{SiO}_2 \rightarrow \text{Si} + \text{SiO}↑ + \text{CO}↑$$  \hspace{1cm} (3.1)

This process produces metallurgical grade silicon with purity of about 98.0%. The next step, the silicon is pulverized and treated with hydrochloric acid gas HCl at 300$^\circ$C to form trichlorosilane SiHCl$_3$. The chemical reaction follows equation (3.2).

$$\text{Si} + 3\text{HCl} \rightarrow \text{SiHCl}_3 + \text{H}_2↑$$  \hspace{1cm} (3.2)

Trichlorosilane, a liquid at room temperature, with boiling point of 32$^\circ$C is fractional distilled to remove unwanted impurities. The purified SiHCl$_3$ is then used in a hydrogen reduction reaction to prepare the electronic grade silicon EGS.

$$\text{SiHCl}_3 + \text{H}_2 \rightarrow \text{Si} + \text{HCl}↑$$  \hspace{1cm} (3.3)

The reaction takes place in a reactor containing resistance heated silicon rod, which serves as the nucleation point for deposition of EGS in polycrystalline form of high purity is the raw material used to prepare device quality single crystal. Pure EGS has impurity concentration generally has impurity concentration in part per billion range.

The polycrystalline silicon is melt in the argon Ar atmosphere in quartz crucible. Right type and amount of dopant is then added. With right temperature control and the aid of “seed” silicon rod of right diameter is formed by rotation and pulling in Czochralski puller as shown in Fig. 3.4. Figure 3.4(a) shows the photograph of a modern computer-controlled Czochralski crystal puller. Figure 3.4(b) is the schematic drawing showing the components of the puller.

Once thermal equilibrium is established, the temperature at the vicinity of the seed is reduced the molten silicon begins to freeze out onto the seed crystal. Subsequently, seed is slowly rotated and withdrawn at the rate of a few millimeter per minute to form a cylindrically shaped single crystal of silicon, which is known as *ingot*. Typically, 4 to 6 inch diameter and 1 to 2 meter in length type of ingot can be formed. In today’s process, ingot of diameter as large as 12 inch is commonly produced to save cost and improve productivity. However, for large ingot as large as 12 inch in diameter, an external magnetic
field is applied around the crucible and it is used to control the concentration of defects, impurities, and oxygen.

In the crystal growth process, boron and phosphorous are the most common dopants for making silicon $p$- and $n$-type semiconductor materials respectively. As a crystal is pulled from the molten silicon, the doping concentration incorporated into crystal is usually different from the doping concentration of the molten silicon at the interface. The ratio of these two concentrations is defined as the *equilibrium segregation coefficient* $k_0$, which defined as

$$k_0 \equiv \frac{C_S}{C_I}$$

where $C_S$ and $C_I$ are respectively the equilibrium concentration of the dopant in the solid and liquid near interface. Figure 3.5 shows the equilibrium segregation coefficient for common dopants used for silicon. The value below one means
3 The Manufacturing Process

that during the growth the dopants are rejected into the molten silicon. As the result, the dopant concentration of molten silicon becomes higher.

<table>
<thead>
<tr>
<th>Dopant</th>
<th>$k_0$</th>
<th>Type</th>
<th>Dopant</th>
<th>$k_0$</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>$8.0 \times 10^{-1}$</td>
<td>$p$</td>
<td>As</td>
<td>$3.0 \times 10^{-1}$</td>
<td>$n$</td>
</tr>
<tr>
<td>Al</td>
<td>$3.0 \times 10^{-3}$</td>
<td>$p$</td>
<td>Sb</td>
<td>$3.3 \times 10^{-2}$</td>
<td>$n$</td>
</tr>
<tr>
<td>Ga</td>
<td>$8.0 \times 10^{-3}$</td>
<td>$p$</td>
<td>Te</td>
<td>$3.0 \times 10^{-4}$</td>
<td>$n$ Deep-lying impurity level</td>
</tr>
<tr>
<td>In</td>
<td>$4.0 \times 10^{-4}$</td>
<td>$p$</td>
<td>Li</td>
<td>$1.0 \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>1.25</td>
<td>$n$</td>
<td>Cu</td>
<td>$4.0 \times 10^{-1}$</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>$7.0 \times 10^{-2}$</td>
<td>$n$</td>
<td>Au</td>
<td>$5.0 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.35</td>
<td>$n$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.5:** Equilibrium segregation coefficients for dopant in silicon

Consider a crystal being growth from the initial molten silicon of weight $M_o$ with an initial doping concentration $C_o$ (the weight of dopant per 1g of molten silicon) in the molten silicon. At a given time, a crystal of weight $M$ has been grown, the amount of the dopant remaining in the molten silicon by weight is $S$. For an incremental amount of the crystal with weight $dM$, the corresponding reduction of the dopant ($-dS$) from the molten is $C_S dM$, where $C_S$ is the doping concentration in the crystal by weight.

$$-dS = C_S dM$$  \hspace{1cm} (3.5)

The remaining weight of the molten silicon is $(M_o - M)$ and the doping concentration in liquid by weight $C_l$ is given by

$$C_l = \frac{S}{M_o - M}$$ \hspace{1cm} (3.6)

Substituting equation (3.5) and (3.6) into equation (3.4), it yields equation (3.7).

$$\frac{dS}{S} = -k_0 \left( \frac{dM}{M_o - M} \right)$$ \hspace{1cm} (3.7)
Given that the initial weight of the dopant $C_0M_o$, integration equation (3.7) yields equation (3.8).

$$\int_{S}^{M} \frac{dS}{S} = -k_o \int_{0}^{M} \left( \frac{dM}{M_o - M} \right)$$  \hspace{1cm} (3.8)

Solving equation (3.8) and combining equation (3.6) and equation (3.4), it yields equation (3.9).

$$C_S = k_o C_o \left( 1 - \frac{M}{M_o} \right)^{k_o-1}$$  \hspace{1cm} (3.9)

During the growth of silicon ingot, dopant is constantly being rejected into the molten silicon. If the rejection rate is higher than the rate at which the dopant can be transported away by diffusion or stirring, then a concentration gradient will develop at the interface as shown in Fig. 3.6. The segregation coefficient is $k_0 = C_S/C_I(0)$. We can define an effective segregation coefficient $k_e$, which is the ratio of $C_S$ and the impurity concentration far away from the interface.

$$k_e \equiv \frac{C_S}{C_I}$$  \hspace{1cm} (3.10)

![Figure 3.6: Doping distribution near the solid-molten interface](image)

Consider a small virtual stagnant molten layer of width $\delta$ in which the only flow that required to replace the crystal being withdrawn from the molten. Outside
3 The Manufacturing Process

the stagnant layer the concentration remains constant at $C_I$. In the layer, the concentration can be described by steady state continuation equation.

$$D \frac{d^2 C}{dx^2} + v \frac{dC}{dx} = 0$$

(3.11)

where D is the diffusion coefficient of the molten silicon and v is the velocity of the crystal growth. The solution of equation is $C = A_1 e^{-v x / D} + A_2$ with the constant to be determined by two boundary conditions. The first is at $x = 0$, $C = C_I(0)$ and second is determined by conservation of total number of dopant i.e. the sum of dopant flux at interface is zero. This condition yields equation.

$$D \left( \frac{dC}{dx} \right)_{x=0} + [C_I(0) + C_S] = 0$$

(3.12)

Substituting the conditions and $C = C_I$ at $x = \delta$, the solution for the concentration $C$ is

$$e^{-\delta v / D} = \frac{C_I - C_S}{C_I(0) - C_S}$$

(3.13)

The effective segregation coefficient $k_e$ is

$$k_e = \frac{C_S}{C_I} = \frac{k_0}{k_0 + (1 - k_0)e^{-\delta v / D}}$$

(3.14)

3.2 Integrated Circuit Fabrication

The fabrication of integrated circuit IC both MOS and bipolar devices, involves a numbers of repeated major process steps. The processes can be broadly classified into wafer cleaning process, oxidation process, lithography (imaging, resist-bleaching and resist development), etching, diffusion/ion implantation, chemical vapor deposition CVD, polycrystalline process or epitaxy, physical vapor deposition/metal deposition or evaporation/sputtering, and sintering/annealing.

Ultra-clean conditions must be maintained during fabrication process especially applicable to lithography process. Any dust particle on silicon wafer can cause defect in the final resist coating. Thus, a clean room facility is necessary for fabrication. Clean room is rated as Class according to the
maximum number of particles of 0.5µm diameter per cubic foot of air. Class 100 to Class 1 clean room is now being used for VLSI/ULSI fabrication process. Class 100 shall mean the air filtration system is able to filter all particles of diameter size greater than 0.5µm and filtered air has less than 100 dust particles of diameter size less than 0.5µm per cubic foot of air. Likewise Class 1,000 shall mean the air filtration system is able to filter all particles of diameter size greater than 10.0µm and filtered air has less than 10 dust particles of diameter size less than 10.0µm per cubic foot of air. Figure 3.7 shows the number of particle versus the diameter of particle expectation for different class of cleanliness for fabrication of the VLSI integrated circuit.

![Figure 3.7: The number of particle and diameter of particle for various classes of cleanliness](image)

### 3.3.1 Wafer Cleaning

By nature, there is a layer of native oxide grown on any wafer due to presence of oxygen in atmosphere and also due to presence of contaminants such as wax, resin, greasy film, sodium chloride, copper, and etc. Hydrofluoric acid is used to remove oxide that formed on surface of silicon wafer. Ammonium hydroxide, sulfuric acid, and hydrogen peroxide are typically used to remove organic contaminants, whilst hydrogen peroxide and hydrochloric acid are used to remove metal contaminants. De-ionized DI water is then used as solvent for cleaning or rinsing. The wafer is finally dried in nitrogen environment to prevent oxidation and contamination. The right proportional mixing of the above mentioned solvents are termed as RCA solution that was developed in 1965. The solutions are divided into solution clean 1 and solution clean 2.
3 The Manufacturing Process

Figure 3.8 shows the eight cleaning steps for cleaning the wafer to remove inorganic, organic, and native oxide contaminants before actual fabrication process steps begin. The figure also shows the composition of the various solutions and the temperature requirements during cleaning process.

<table>
<thead>
<tr>
<th>Step</th>
<th>Solution</th>
<th>Temperature</th>
<th>Type of Contaminant to be removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H$_2$SO$_4$ + H$_2$O$_2$ (4:1)</td>
<td>120$^\circ$C</td>
<td>Organic particle</td>
</tr>
<tr>
<td>2</td>
<td>DI water</td>
<td>25$^\circ$C</td>
<td>Rinse</td>
</tr>
<tr>
<td>3</td>
<td>NH$_3$OH + H$_2$O$_2$ + H$_2$O (1:1:5)</td>
<td>80$^\circ$C – 90$^\circ$C</td>
<td>Organic particle</td>
</tr>
<tr>
<td>4</td>
<td>DI water</td>
<td>25$^\circ$C</td>
<td>Rinse</td>
</tr>
<tr>
<td>5</td>
<td>HCl + H$_2$O$_2$ + H$_2$O (1:1:6)</td>
<td>80$^\circ$C – 90$^\circ$C</td>
<td>Inorganic ion</td>
</tr>
<tr>
<td>6</td>
<td>DI water</td>
<td>25$^\circ$C</td>
<td>Rinse</td>
</tr>
<tr>
<td>7</td>
<td>HF + H$_2$O (1:50)</td>
<td>25$^\circ$C</td>
<td>Native oxide</td>
</tr>
<tr>
<td>8</td>
<td>DI water</td>
<td>25$^\circ$C</td>
<td>Rinse</td>
</tr>
</tbody>
</table>

Figure 3.8: Cleaning step and composition of RCA solution

De-ionized water or DI water is highly purified and filtered water that all traces of ionic, particles, and bacterial contamination have been removed. The theoretical resistivity of pure water at 25$^\circ$C is 18.3Mohm-cm. A basic DI water system can achieve resistivity of 18.0Mohm-cm with fewer than 1.2 colonies of bacteria per milliliter and no particle size larger than 0.25µm.

3.3.2 Oxidation Process

Silicon dioxide SiO$_2$ is a very good dielectric material. Its dielectric constant is 3.9. Silicon dioxide is used mainly for masking where dopant cannot be diffused, passivation, and insulation.

Oxidation is a process of growing a thin layer of amorphous silicon dioxide. The process is also termed as chemical vapor deposition CVD. There are several methods to grow oxide named as the dry and wet methods. Thermal oxide is grown using oxygen and silicon yields dry oxide.

\[
\text{Si} + \text{O}_2 \longrightarrow \text{SiO}_2 \quad \text{(3.15)}
\]

or wet oxide is grown using steam and silicon.

\[
\text{Si} + 2\text{H}_2\text{O} \longrightarrow \text{SiO}_2 + 2\text{H}_2 \quad \text{(3.16)}
\]
Anodic oxide is formed in gaseous or liquid medium by electric field induced transportation of mobile ion. This method is also termed as low pressure chemical vapor deposition LPCVD.

Other methods are: low temperature (400°C to 500°C) chemical vapor deposition using silane SiH₄ and oxygen O₂ and high temperature (1,000°C) deposition using tetrachlorosilane SiCl₄ with CO₂, O₂, and H₂O.

\[ \text{SiH}_4 + \text{O}_2 \rightarrow \text{SiO}_2 + 2\text{H}_2 \]  \hspace{1cm} (3.17)

At this point, it is worth to mention that there is another passivation film, which is silicon nitride Si₃N₄ film and it is also used to passivate semiconductor device because it acts as a barrier to the diffusion of metal ions, particularly sodium ions. Reactive sputtering of ammonia or nitrogen is the most common way to deposit this type of film.

\[ 3\text{SiH}_4 + 4\text{NH}_3 \rightarrow 900°C \rightarrow \text{Si}_3\text{N}_4 + 12\text{H}_2 \]  \hspace{1cm} (3.18)

or

\[ 3\text{SiCl}_4 + 4\text{NH}_3 \rightarrow 550°C-1200°C \rightarrow \text{Si}_3\text{N}_4 + 12\text{HCl} \]  \hspace{1cm} (3.19)

The oxidation process described above generally forms an oxide film that covers the entire surface of the wafer. The ability to selective oxidize a particular area on surface of silicon is important in high-density bipolar and MOS processes. It is because selective oxidation has to improve the device pack density and more planar final structures. To achieve this, it utilizes the technique called localized oxidation of silicon LOCOS process.

There are two types of LOCOS processes, which would result semi-recessed oxide and fully recessed structure. Oxide grown on silicon wafer without pre-etch process that is not protected by silicon nitride Si₃N₄ is called semi-recessed structure. Full recessed oxide is formed by etching the silicon prior to oxidation. This process can yield a very planar surface after silicon nitride removal. The cross section depicting the process sequence of LOCOS for semi-recessed and fully recessed structures are shown in Fig. 3.9.

The analysis of oxidation process shows that the grown oxide thickness \( t_{\text{ox}} \) can be approximated by the quadratic equation

\[ t_{\text{ox}}(t) = \frac{A}{2} \left[ \sqrt{1 + \frac{4Bt}{A^2}} - 1 \right], \]

where A and B are coefficient depending on temperature, crystal orientation and gas
mixture. The initial phase of oxide growth is linear which is approximately equal to $B \frac{t}{A}$. As time goes on, the process is slower that follows equation $\sqrt{Bt}$.

One silicon atom is used to form one molecule of silicon dioxide $\text{SiO}_2$. Based on the density of silicon $N_{\text{Si}}$ and silicon dioxide $N_{\text{ox}}$, which are $5.0 \times 10^{22} \text{ cm}^{-3}$ and $3.3 \times 10^{22} \text{ cm}^{-3}$ respectively, the recession is 46% meaning every unit thickness of oxide formed required 0.46 unit thickness of silicon.

(a) Semi-recessed structure LOCOS  (b) Fully recessed structure LOCOS

**Figure 3.9:** Cross section depicting the process sequence for obtaining semi-recessed and fully recessed LOCOS

Trench isolation either shallow or deep refilled types are used in advanced MOS and bipolar processes. Shallow trench isolation STI process step is shown in Fig. 3.10.

(a) Stack and trench etch  (b) Pad oxide undercut  (c) Liner oxidation

(d) CVD oxide gap filled  (e) CMP and HF dip  (f) $\text{H}_2\text{PO}_4$ nitride strip

**Figure 3.10:** Process step showing shallow trench isolation structure
Deep trench refilled with polysilicon is used to form trench capacitor used in DRAM memory design. It is also used as isolation in SiGe heterojunction HBT technology. The process step of deep trench refilled with polysilicon with combination of LOCOS field oxidation is shown in Fig. 3.11.

The deep trench is formed from reactive ion etching process. This process can create deep trench of high aspect ratio. The surface of trench is passivated with a layer of thermally grown oxide. The rest of process steps should be self explanatory.

Chemical Mechanical polishing CMP process is introduced in early 1990’s and it is now used to achieve high planar topology for deep submicron process. It is used in both bipolar and MOS processes. The conceptual process technique is shown in Fig. 3.13. The wafer is mounted on a rotating pattern. Liquid slurry is continuous dispensed on the surface of polishing pad. A combination of vertical force between wafer and the abrasive pad as well as the chemical action of slurry is used to polish the surface to highly planar state. Polishing process stops when the nitride layer is fully exposed.
3 The Manufacturing Process

Figure 3.12: Chemical mechanical polishing technique

3.3.3 Lithography

It is a process encompasses all the steps involved in transferring a pattern from a mask to the surface of silicon wafer. This is also a process to create precise dimensioned open area on silicon dioxide/silicon nitride or metallic surface. Opening on silicon dioxide surface will allow diffusion or doping of other impurities. There are two methods to do this. They are optical lithography and electron beam or ion beam lithography.

The process steps of photolithography are clean wafer, deposit barrier layer of SiO₂, Si₃N₄, metal, coating with photoresist, soft bake, align mask, expose pattern, develop photoresist, hard bake, etch windows in barrier layer, and removal of photoresist. In brief, photolithography has three steps i.e. imaging, resist bleaching, and resist-development.

Before imaging is made, the wafer is coated with a layer of photoresist, an organic compound, which is either light sensitive or non-light sensitive. A mask, which is made of chrome coated glass or quartz containing the pattern of the IC design, is used either to block out the required part or non-required part on the wafer. Example of the photoresist is polyisoprene derivative. There are two types of photoresist. The positive type is light sensitive to ultra violet UV or other light source, acid-resist organic polymer, initially insoluble to developing solution and after exposure becomes soluble to developing solution. The negative photoresist is initially soluble to developing solution, after exposing to UV light becomes insoluble. In the modern lithography, deep UV light of wavelength either 193nm or 248nm, electron beam, ion beam, X-ray are used as the light source.
Selective exposing the photoresist needs a chrome-coated glass or quartz as mask used to cover the area so that it is opaque. The photoresist is developed in flowing aqueous alkaline or trichloroethylene and then bake in the oven.

### 3.3.4 Etching

After thin film is deposited on the wafer surface, it is selectively removed by etching to leave the desired pattern of the film on the wafer surface. Several techniques are used to transfer the resist pattern to the silicon wafer. Wet etch and dry etch are the common methods. With decrease feature size, dry etching is preferred than wet etching in which wet etching has over-etch problem.

**Wet Etch.** Silicon dioxide is removed by hydrofluoric acid HF or hydrochloric acid HCL, whilst aluminum interconnects can be removed by phosphoric acid.

\[
\text{SiO}_2 + 6\text{HF} \rightarrow \text{H}_2\text{SiF}_6 + 2\text{H}_2\text{O} \quad (3.20)
\]

The photoresist is removed with sulfuric acid H\textsubscript{2}SO\textsubscript{4}.

**Dry Etch.** Dry etch is based on reactive ion-etch process and plasma-etch. This process is widely used in VLSI fabrication.

### 3.3.5 Polycrystalline Process

Polycrystalline is also known as *epitaxial growth* is characterized by low resistivity of doped silicon. It is a process of growing another layer of doped silicon crystal on the existing silicon crystal. Several process steps are involved for making polycrystalline.

The chemical reaction of hydrogen with silicon tetrachloride SiCl\textsubscript{4} is a method, which has chemical equation (3.21).

\[
\text{SiCl}_4 + 2\text{H}_2 \xrightarrow{1200^\circ\text{C}} \text{Si} + 4\text{HCL} \quad (3.21)
\]

Deposition of silicon can be also achieved using LPCVD process using thermal decomposition of silane SiH\textsubscript{4} in hydrogen environment.

\[
\text{SiH}_4 \xrightarrow{\text{H}_2 \text{ atmosphere} \ 600^\circ\text{C}} \text{Si} + 2\text{H}_2\uparrow \quad (3.22)
\]
3 The Manufacturing Process

Liquid-phase epitaxy LPE and molecular beam epitaxy MBE are being widely used in compound semiconductor technology especially the silicon germanium technology.

3.3.6 Diffusion

Diffusion of impurities is typically done by placing semiconductor wafers in a carefully controlled, high temperature quartz-tube furnace and passing a gas mixture that contain the desired dopant through it. Thus, its purpose is to introduce dopant into silicon crystal. Mixture of oxygen and dopants such as diborane and phosphine are introduced in the furnace with the exposed wafer surface at temperature ranges between 800°C and 1,200°C for silicon and 600°C and 1,000°C for gallium arsenide. The number of dopant atoms that diffuse into the semiconductor is related to the partial pressure of the dopant impurity in the gas mixture.

Dopant can be introduced by solid source (e.g. BN for boron, As₂O₃ for arsenic, and P₂O₅ for phosphorus), gases source (e.g. B₂H₆, AsH₃, and PH₃), and liquid source (e.g. BBr₃, AsCl₃, and POCl₃). The chemical reaction for phosphorous diffusion is as follow.

\[
4\text{POCl}_3 + 3\text{O}_2 \rightarrow 2\text{P}_2\text{O}_5 + 6\text{Cl}_2 \uparrow \quad (3.23)
\]

The P₂O₅ forms a glass-on-silicon wafer and then reduced to phosphorous by silicon following the equation.

\[
2\text{P}_2\text{O}_5 + 5\text{Si} \rightarrow 4\text{P} + 5\text{SiO}_2 \uparrow \quad (3.24)
\]

The phosphorous is released and diffuses into silicon and chlorine Cl₂ gas is vented.

For diffusion in gallium arsenide, the high vapor pressure of arsenic requires special method to prevent loss of arsenic by decomposition or evaporation. These methods include diffusion in sealed ampules with over pressure of arsenic and diffusion in an open-tube furnace with doped oxide capping layer such as silicon nitride. Most of the studies on p-type diffusion have been confined to the use of zinc in the forms of Zn-Ga-As alloys and ZnAs₂ for the sealed-ampule approach or ZnO-SiO₂ for the open-tube approach. The n-type dopants in gallium arsenide include selenium and tellurium.
To complete the process, 'drive in' or re-distribution of dopant is done in nitrogen or wet oxygen where silicon dioxide SiO$_2$ is grown at the same time.

### 3.3.7 Ion Implantation

Like diffusion process, it is a process where dopant is introduced using ion species such as BF$_3$ and PF$_5$. The ion species are ionized and accelerated to mass separation magnet where they are targeted at the wafer surface. The implantation energies are between 1.0keV and 1.0MeV. It would result an ion distribution with average depths ranging from 10nm to 10µm. Ion doses vary from $10^{12}$ ions/cm$^2$ for threshold voltage adjustment to $10^{18}$ ions/cm$^2$ for formation of buried insulating layer. The main advantage over the diffusion is low temperature, more precise control and reproducibility of impurity doping, and shallow implant. However, owing to high-energy bombardment, Rapid thermal annealing RTA at 400$^\circ$C to 500$^\circ$C is required to allow the implanted atom to stay at the right substitutional site and at the same time to repair crystal damage.

### 3.3.8 Metallization

The most common methods of physical vapor deposition PVD of metals are evaporation, e-beam evaporation, plasma spray deposition, and sputtering. Aluminum film can be deposited by PVD or CVD. Since aluminum and its alloys have low resistivity, these metals satisfy the low resistivity from 3.7µΩ-cm for Al to 3.5µΩ-cm for its alloys requirements. Aluminum adheres well to silicon dioxide. However, the use of Al for shallow integrated circuits with shallow junctions often create problem such as spiking and electro-migration.

Aluminum deposition can be done on high temperature reactive sputtering using electrical discharge at 1,000 to 3,000V. This is a process of chemical vapor deposition. It can also be done in vacuum vapor deposition by heating tungsten filament or electron beam evaporation in vacuum. Other methods used are filament evaporation, electron-beam evaporation, flash evaporation etc.

Chemical vapor deposition forms thin film on the surface of a substrate by thermal decomposition or reaction of gaseous compounds. Polysilicon, silicon dioxide, and silicon nitride are commonly deposited using CVD method. In addition, refractory metals such as tungsten can be deposited using CVD.
3 The Manufacturing Process

3.3.9 Sintering/Annealing

Sintering is done in nitrogen environment to establish intimate contact between silicon and aluminum. Rapid thermal annealing RTA is a process using to repair damage due to implantation and move the implanted ion to right substitutional site. The process is also used to drive-in the impurity.

3.3 Fabrication of Bipolar Junction Transistor

Traditionally bipolar junction transistor is mostly fabricated using junction isolation process. However, nowadays oxide isolation, dielectric isolation, self aligned double polysilicon process etc are commonly used to fabricate advanced bipolar junction transistor. The flowchart of the process steps involving seven mask sets are shown in Fig. 3.13.

![Flowchart of the fabrication process steps of junction isolation bipolar junction transistor](image)

**Figure 3.13:** The fabrication process steps of junction isolation bipolar junction transistor

An pictorial illustration of the processes for fabricating two npn bipolar junction transistor are shown in Fig. 3.14. The transistors are built with a p-type substrate.
3 The Manufacturing Process

The $p^+$ is used as the isolation, which is called *junction isolation* and the $p$ substrate is tied to the most negative voltage of the circuit. It is usually tied to $V_{EE}$ voltage, which is ground most of the time. This creates the reverse biasing of the $p^+$ and $n$ epitaxial layer junction that forms a natural isolation.

![Figure 3.14](image)

Figure 3.14: Step by step fabrication of npn bipolar transistor

Figure 3.14(a) shows the wafer type for the fabrication of bipolar junction transistor is a $p$-doped (111) orientation type. Figure 3.14(b) illustrates the devices after finishing fabricating the $n^+$ deep diffused barrier layer. Figure
3 The Manufacturing Process

3.14(c) illustrates the formation of polycrystalline and silicon dioxide step. Figure 3.14(d) illustrates the formation of $p^+$ junction isolation. Figure 3.14(e) shows the top view of devices shown Fig. 3.14(d) after the removal of silicon dioxide. Figure 3.14(f) shows the formation $p$-base material for the transistors. Figure 3.14(g) shows the top view of devices shown in Fig. 3.14(f). Figure 3.14(h) shows the formation of $n^+$ emitter and $n^+$ contact diffusion for collector. Figure 3.14(i) shows the side view of the bipolar junction transistor with the aluminum interconnect put in-placed. From Fig. 3.14(i), the two npn bipolar junction transistors are connected as a current mirror circuit. Figure 3.14(j) shows the top view of devices shown in Fig. 3.14(i).

3.4 Fabrication of CMOS Transistor

There are many process steps for fabrication of CMOS transistors. Listed here is a 41-step SAJI processes which are shown in Fig. 3.15. In this section, the process steps mentioned in the figure are not fully covered sequentially.

<table>
<thead>
<tr>
<th></th>
<th>1 Initial cleaning</th>
<th>2 Pad oxidation</th>
<th>3 $Si_3N_4$ deposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$p$-well making</td>
<td>5 $Si_3N_4$ etching</td>
<td>6 $p$-well ion implant</td>
</tr>
<tr>
<td>7</td>
<td>Drive-in &amp; oxidation</td>
<td>8 $Si_3N_4$ removal</td>
<td>9 $n$-well implant</td>
</tr>
<tr>
<td>10</td>
<td>Well annealing</td>
<td>11 $SiO_2$ removal</td>
<td>12 Pad oxidation</td>
</tr>
<tr>
<td>13</td>
<td>Nitride deposition</td>
<td>14 LOCOS making</td>
<td>15 Nitride etch</td>
</tr>
<tr>
<td>16</td>
<td>Field oxidation</td>
<td>17 Nitride removal</td>
<td>18 Oxide removal</td>
</tr>
<tr>
<td>19</td>
<td>Gate oxidation</td>
<td>20 Poly gate deposition</td>
<td>21 Gate masking</td>
</tr>
<tr>
<td>22</td>
<td>Dry etch of poly gate</td>
<td>23 Oxidation</td>
<td>24 Source &amp; drain of $n$-MOS</td>
</tr>
<tr>
<td>25</td>
<td>Source &amp; drain of $p$-MOS</td>
<td>26 Undoped oxide deposition</td>
<td>27 Source/drain drive-in</td>
</tr>
<tr>
<td>28</td>
<td>BPSG glass deposition</td>
<td>29 Contact holes of metal 1</td>
<td>30 Metal 1 deposition sputtering</td>
</tr>
<tr>
<td>31</td>
<td>Metal 1 patterning</td>
<td>32 Deposition of $SiO_2$</td>
<td>33 $SiO_2$ coating</td>
</tr>
<tr>
<td>34</td>
<td>SOG curing</td>
<td>35 Dielectric layer deposition</td>
<td>36 Contact holes of metal 2</td>
</tr>
<tr>
<td>37</td>
<td>Metal 2</td>
<td>38 $SiO_2$ deposition</td>
<td>39 Deposition of $Si_3N_4$</td>
</tr>
<tr>
<td>40</td>
<td>Bonding pad</td>
<td>41 Annealing</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.15:** CMOS process steps

The fabrication of $n$-MOS and $p$-MOS transistors in a $p$-type wafer that normal has about 300 to 600µm thickness, is shown in this section. $N$-region must be introduced to accommodate the process of $p$-MOS and it is called an $n$-well process. The first process of the fabrication involves growth a thin layer thermal oxide. It is then followed creating an $n$-well region using mask #1 by etching
away the oxide. The similar mask is used to dope the \( n \)-well by the process of ion implantation. After ion implantation, the wafer is annealed to repair crystal damage as the result of ion implantation process. Finally the oxide is stripped. Figure 3.16 and 3.17 illustrate the process of \( n \)-well formation.

![Figure 3.16: \( n \)-well implant](image1)

Active area such as gate, drain and source and field oxide FOX formations are the next process step utilizing mask #2. As it is illustrated in Fig. 3.18, thermal oxide is growth first and growing silicon nitride \( \text{Si}_3\text{N}_4 \) is the next step. The mask #2 used to define the field oxide FOX region as the isolation is shown in Fig. 3.19. At this point of process, ion implantation is also used to adjust the threshold voltage of the \( n \)-MOS and \( p \)-MOS transistors.

![Figure 3.17: \( n \)-well](image2)

![Figure 3.18: Active area definition](image3)
Polysilicon gate is a next step utilizing mask #3. As it is illustrated in Fig. 3.20, the thin layer of thermal oxide is grown before the polysilicon layer. The mask #3 is then used to mask out unwanted polysilicon to create the gate of the \( n \)-MOS and \( p \)-MOS transistors as illustrated in Fig. 3.21.

Mask #4 and #5 are alternatively used to create lightly doped drain LDD aimed to reduce short channel effect due to hot electrons and hot holes phenomena. Figure 3.22 illustrates the lightly doped drain LDD process for \( p \)-MOS transistor, whilst Fig. 3.23 illustrates the lightly doped drain LDD process for \( n \)-MOS transistor. Note that the photoresist will block implant ion but it can penetrate the oxide layer.
The next process is spacer formation, which is simply growing oxide to cover the polysilicon gate. The processes are illustrated in Fig. 3.24 and 3.25 respectively. This process step is preparing for creating a self-aligned heavily doped drain and source for the \( n \)-MOS and \( p \)-MOS transistors.
Upon finishing the creation of spacer, the next step is deep $p^+$ and $n^+$ implant which are illustrated in Fig. 3.26 and Fig. 3.27 respectively. It is done by removing the old photoresist and lithography process and etching and boron implant for $p$-MOS transistor and arsenic implant for $n$-MOS transistor.
The photoresist is used to block the unwanted ion implant to the silicon surface. It is used alternatively to block boron during the implant to penetrate to source and drain of the $n$-MOS transistor. During the arsenic implant, it is used to block arsenic from penetrating to source and drain of $p$-MOS transistor.

Mask #6 is used to provide metal contact as illustrated in Fig. 3.28 and Fig. 3.29 respectively. The process is done by growing silicon diode and lithography to etch the oxide that will provide contact with drain and source of the transistors. The cut in oxide is then filled with metal to get the via. The via is used to connect the drain and source of the MOS transistors to metal 1 level just like the pillars of the bridge or road connected to the foundation and the surface of the road.
The rest of processes are mainly the metal process, which are illustrated in Fig. 3.30 through Fig. 3.33. Owing to the complexity of the circuit, the number of metal level can be varied.

The final layer is the passivation layer, whereby the surface top of the integrated circuit is covered with a layer of glass with selective opening for providing bond pad for wire connection to the device’s package. The illustration of the passivation layer with access cut is shown in Fig. 3.33.
Figure 3.31: Oxide and planarization

Figure 3.32: High level metal layers and vias
3 The Manufacturing Process

Figure 3.33: Top metal and passivation layer

3.5 Packing Integrated Circuit

Integrated circuit package plays a fundamental role in the operation and performance of a component. Besides providing a means of bringing the signal and supply wires in and out of the silicon die, it also removes the heat generated by the circuit and provides mechanical support. Its also protects the die integrated circuit against environmental conditions such as humidity and heat. Furthermore the package has a major impact on the performance and power dissipation of the integrated circuit like the microprocessor and signal processor. This influence is getting more pronounced as technology scaling down progressed due to reduction of internal signal delays and on-chip capacitance. Up to 50% of the delay of a high-performance computer is due to packaging delay caused by inductive and capacitive parasitic from packaging material. The increasing complexity of circuit integrated into a single die also translates into a need for ever more input-output pins because the number of connections is roughly proportional to the complexity of the circuitry on the chip. This relationship was first observed by E. Rent of IBM, who translated it into an empirical formula called Rent’s rule. This formula relates the number of input/output pins $P$ to the complexity of the circuit as measured by the number of gates.

$$P = kG^\beta$$

(3.25)
where $k$ is the average number of I/Os per gate, $G$ is the number of gates, and $\beta$ is the Rent exponent. $\beta$ varies between 0.1 and 0.7. The value is strongly dependent on the application area, architecture, and organization of the circuit, as shown in Fig. 3.34.

It is clearly shown that microprocessors display a very different input/output behavior as compared to memory devices. The observed rate of pin-count increase for integrated circuits varies between 8% to 11% per year and it has been projected that packages with more than 2,000 pins will be required by the year 2010. For all these reasons, traditional dual-in-line, through-hole mounted packages have been replaced by other approaches such as surface-mount, ball grid array, and multichip module techniques. It is useful for the circuit designer to be aware of the available options, and their advantages and disadvantages.

<table>
<thead>
<tr>
<th>Chip/System</th>
<th>$\beta$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static memory</td>
<td>0.12</td>
<td>6.00</td>
</tr>
<tr>
<td>Microprocessor</td>
<td>0.45</td>
<td>0.82</td>
</tr>
<tr>
<td>Gate array</td>
<td>0.50</td>
<td>1.90</td>
</tr>
<tr>
<td>High speed computer - chip</td>
<td>0.63</td>
<td>1.40</td>
</tr>
<tr>
<td>High speed computer - circuit</td>
<td>0.25</td>
<td>83.0</td>
</tr>
</tbody>
</table>

**Figure 3.34:** Rent’s constant for varies class of chip and system

Owing to its multi-functionality, a good package must comply with a large variety of requirements namely the electrical, mechanical, thermal, and cost requirements.

**Electrical requirements:** The pins should exhibit low capacitance - both inter-wire and to the substrate, resistance, and inductance. Large characteristic impedance should be tuned to optimize transmission line behavior and observe that intrinsic integrated circuit impedances are high.

**Mechanical and thermal properties:** The heat removal rate should be as high as possible. Mechanical reliability requires a good matching between the thermal properties of the die-integrated circuit and the chip carrier. Long-term reliability requires a strong connection from the die to the package as well as from the package to the printed circuit board.

**Cost:** Cost is always one of the most important properties. Ceramic packages have a superior performance over plastic packages but they are also
3 The Manufacturing Process

substantially more expensive. Increasing the heat removal capacity of a package also tends to raise the package cost. The least expensive plastic packaging can dissipate up to 1.0W of heat. More expensive but still considered cheap plastic packages can dissipate up to 3.0W. Higher heat dissipation requires more expensive ceramic package. Chips dissipating over 50.0W require special heat sink attachment. Extreme techniques such as fans and blowers, liquid cooling hardware, or heat pipes are needed for higher dissipation levels.

Packing density is a major factor in reducing the cost of printed circuit board. The increasing pin count either requires an increase in the package size or a reduction in the pitch between the pins. Both have a profound effect on the packaging economics.

Packages can be classified in many different ways. It is by the material used, the number of interconnection levels, and the method used to remove heat.

3.5.1 Packaging Materials

The most common materials used for the package body are ceramic and polymer (plastics). The later have the advantage of being substantially cheaper but suffer from inferior thermal properties. For instance, the ceramic alumina Al$_2$O$_3$ conducts heat better than SiO$_2$ and the polyimide plastic by factors of 30 and 100 respectively. Furthermore, its thermal expansion coefficient is substantially closer to the typical interconnect metals.

The disadvantage of alumina and other ceramics is their high dielectric constant, which results in large interconnect capacitances.

3.5.2 Interconnect Level

The traditional packaging approach uses a two-level interconnection strategy. The die is first attached to an individual chip carrier or substrate. The package body contains an internal cavity where the chip is mounted. These cavities provide ample room for many connections between chip and leads. The leads compose the second interconnect level and connect the chip to the global interconnect medium, which is normally a printed circuit board. Complex systems contain even more interconnect levels, since boards are connected together using backplanes or ribbon cables. The first two layers of the interconnect hierarchy are illustrated in the drawing of Fig. 3.35.
The interconnect techniques used at levels one and two of the interconnect hierarchy are shown here.

3.5.3.1 Interconnect Level 1 - Die-to-Package-Substrate

Traditionally wire bonding was the technique of choice to provide an electrical connection between die and package. In this approach, the backside of the die is attached to the substrate using glue with a good thermal conductance. Next, the chip pads are individually connected to the lead frame with aluminum or gold wire. An example of wire bonding is shown in Fig. 3.36. Although the wire-bonding process is automated, it has some major disadvantages.

Wire must be attached serially one after the other. The lead time is longer with increasing pin counts. Larger pin counts make it substantially more challenging to find bonding patterns that avoid shorts between the wires. Bonding wire has inferior electrical properties such as a high individual inductance (5nH or more) and mutual inductance with neighboring signals. The inductance of a bonding wire is typically about 1.0nH/mm, while the inductance per package pin ranges.
3 The Manufacturing Process

between 7.0 and 40.0 nH per pin depending on the type of package as well as the positioning of the pin on the package boundary. Typical values of the parasitic inductances and capacitances for a number of commonly used packages are summarized in Fig. 3.37.

<table>
<thead>
<tr>
<th>Package Type</th>
<th>Capacitance pF</th>
<th>Inductance nH</th>
</tr>
</thead>
<tbody>
<tr>
<td>68 pin plastic DIP</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>68 pin ceramic DIP</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>256 pin grid array</td>
<td>1-5</td>
<td>2-15</td>
</tr>
<tr>
<td>Wire bond</td>
<td>0.5-1</td>
<td>1-2</td>
</tr>
<tr>
<td>Solder bump</td>
<td>0.1-0.5</td>
<td>0.01-0.1</td>
</tr>
</tbody>
</table>

**Figure 3.37:** Typical conductance and inductance of package type and wire

The exact value of the parasitic component is hard to predict because of the manufacturing approach and irregular outlay. New attachment techniques are being explored as a result of these deficiencies. In one approach is called *Tape Automated Bonding* (TAB). The die is attached to a metal lead frame that is printed on a polymer film typically polyimide shown in Fig. 3.38(a). The connection between chip pad and polymer film wire is by solder bump as shown in Fig. 3.38(b). The tape can then be connected to the package body using a number of techniques. One approach is using pressure connection.

![Image](image.png)

**Figure 3.38:** Automated tap bonding (a) polymer with imprinted wire pattern and (b) die attaché using solder bump

The advantage of the TAB process is that it is highly automated. The sprockets in the film are used for automatic transport. All connections are made simultaneously. The printed approach helps to reduce the wiring pitch, which results in higher lead counts. Elimination of the long bonding wires improves the electrical performance.
Another approach is to flip the die upside-down and attach it directly to the substrate using solder bumps. This technique, called flip-chip mounting, has the advantage of a superior electrical performance as shown in Fig. 3.39. Instead of making all the I/O connections on the die boundary, pads can be placed at any position on the chip. This can help address the power and clock distribution problems, since the interconnect materials on the substrate are typically copper Cu or gold Au of a better quality than the aluminum Al on the chip.

![Flip-chip bonding](image)

**Figure 3.39:** Flip-chip bonding

### 3.5.3.2 Interconnect Level 2 - Package Substrate to Board

When connecting the package to the printed circuit board, through-hole mounting has been the packaging style of choice. A printed circuit board is manufactured by stacking layers of copper and insulating epoxy glass. In the through-hole mounting approach, holes are drilled through the board and plated with copper. The package pins are inserted and electrical connection is made with solder as shown in Fig. 3.40(a). The traditional package in this class was the dual-in-line package or DIP. The packaging density of the DIP degrades rapidly when the number of pins exceeds 64. This problem can be alleviated by using the pin-grid-array PGA package that has leads on the entire bottom surface instead of only on the periphery. PGA package can extend to large pin counts over 400 pins.

![Printed circuit board mounting approach](image)

**Figure 3.40:** Printed circuit board mounting approach. (a) through-hole mounting and (b) surface mounting

The through-hole mounting approach offers a mechanically reliable and sturdy connection. However, this setback is the packaging density. For mechanical and sturdy reasons, a minimum pitch of 3.54mm between the through-holes is required. Even with this pitch, PGAs with large numbers of pins would also
substantially weaken the board. In addition, through-holes limit the board packing density by blocking lines that might otherwise have been routed below them, which results in longer interconnections. PGAs with large pin counts hence require extra routing layers to connect to the huge number of pins. Although the parasitic capacitance and inductance of the PGA are slightly lower than that of the DIP, their values are still substantial. These shortcomings of the through-hole mounting can be solved by using the surface-mount technique. A chip is attached to the surface of the board with a solder connection without requiring any through-holes as shown in Fig. 3.41(b). Consequently, surface-mount increases packing density due to through-holes elimination that provides more wiring space and mechanically strengthen the PCB; the lead pitch is reduced; and chips can be mounted on both sides of the board.

The negative effects of the surface-mount are connection makes the component board connection weaker. It is also cumbersome to mount a component on a board, requires expensive mounting equipment, difficult for board repair, and finally testing of board is more complex, because the package pins are no longer accessible at the backside of the board. Signal probing becomes hard or even impossible.

A variety of surface-mount packages are currently in use with different pitch and pin-count parameters. Four of these packages are shown in Fig. 3.41: the small-outline package SOIC with gull wings, the plastic leaded package PLCC with J-shaped leads, the leadless chip carrier LCC, and quad flatpack QFP.

Figure 3.41: Commonly use package (1) leadless carrier, (2) DIP, (3) PGA, (4) small outline IC, (5) quad flatpack, and (6) PLCC
An overview of the most important parameters for a number of packages is shown in Fig. 3.42.

<table>
<thead>
<tr>
<th>Package Type</th>
<th>Typical Lead Spacing</th>
<th>Maximum Lead Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual in Line DIP</td>
<td>3.54mm</td>
<td>64</td>
</tr>
<tr>
<td>Pin Grid Array PGA</td>
<td>3.54mm</td>
<td>&gt;300</td>
</tr>
<tr>
<td>Small Outline IC SOIC</td>
<td>1.27mm</td>
<td>28</td>
</tr>
<tr>
<td>Plastic Leadless Chip</td>
<td>1.27mm</td>
<td>124</td>
</tr>
<tr>
<td>Carrier PLCC</td>
<td>0.75mm</td>
<td>124</td>
</tr>
</tbody>
</table>

**Figure 3.42:** Parameters of various chip carriers

Even surface-mount packaging is unable to satisfy the quest for evermore higher pin counts. This is worsened by the demand for power connections: today’s high performance chips, operating at low supply voltages, require as many power and ground pins as signal I/O. When more than 300 I/O connections are needed, solder balls replace pins as the preferred interconnect medium between package and board. An example of such a packaging approach, called ceramic *ball grid array* BGA is shown in Fig. 3.43. Solder bumps are used to connect both the die to the package substrate and the package to the board. The area array interconnect of the BGA provides constant input/output density regardless of the number of total package I/O pins. A minimum pitch between solder balls of as low as 0.8mm can be obtained, and packages with multiple 1,000’s of I/O signals are feasible.

![Ball grid array packaging](image)

**Figure 3.43:** Ball grid array packaging; (a) cross-section, (b) photo of BGA bottom

### 3.5.3.3 Multi-Chip Modules - Die to Board

The deep hierarchy of interconnect levels in the package is becoming unacceptable in today’s complex circuit designs with higher levels of
3 The Manufacturing Process

integration, large number of signals, and performance requirements. There is a need to reduce the number of levels.

At the meantime, attention is focused on the elimination of the first level in the packaging hierarchy, which is eliminating die to package level by mounting the die directly on the wiring backplanes board or substrate. It offers a substantial benefit when performance or density is a major concern. This packaging approach is called multichip module technique MCM. As the consequently, there is a substantial increase in packing density as well as improved performance. A number of the previously mentioned die-mounting techniques can be adapted to mount dice directly on the substrate. This includes wire bonding, Tape Automated Bonding TAB, and flip-chip, although the later two are preferable. The substrate itself can be varying over a wide range of materials, depending upon the required mechanical, electrical, thermal, and economical requirements. Materials of choice are epoxy substrates similar to printed circuit boards, metal, ceramics, and silicon. Silicon has the advantage of presenting a perfect match in mechanical and thermal properties with respect to the die material.

The main advantages of the MCM approach are the increased packaging density and performance. An example of an MCM module implemented using a silicon substrate; commonly dubbed silicon-on-silicon is shown in Fig. 3.44. The module that implements an avionics processor module and is fabricated by Rockwell International, contains 53 ICs and 40 discrete devices on a 3.2” × 3.2” substrate with aluminum polyimide interconnect.

Figure 3.44: An Avionics processor module. Courtesy of Rockwell International
The interconnect wires of the module are only an order of magnitude wider than what is typical for on-chip wires because similar patterning approaches are used. The module itself has 180 I/O pins. Performance is improved by the elimination of the chip-carrier layer with variety of parasitic components, and through a reduction of the global wiring lengths on the die, a result of the increased packaging density. For instance, a solder bump has an assorted capacitance and inductance of only 0.1pF and 0.01nH respectively. The MCM technology can also reduce power consumption significantly, since large output drivers and associated dissipation become redundant due to the reduced load capacitance of the output pads.

While MCM technology offers some clear benefits, its main disadvantage is economic. This technology requires some advanced manufacturing steps that make the process expensive. The approach is only justifiable when either dense housing or extreme performance is essential. In the near future, this argument might become obsolete as MCM approaches proliferate.

3.5.3 Thermal Consideration in Packaging

As the power consumption of integrated circuits rises, it becomes increasingly important to efficiently remove the heat generated by the chips. A large number of failure mechanisms in ICs are accentuated by increased temperatures. Examples are leakage in reverse biased diodes, electro-migration, and hot-electron trapping. To prevent failure, the temperature of the die must be kept within certain ranges. The temperature range for commercial graded devices during operation is 0° to 70°C. Military parts are more demanding and require a temperature range varying from -55° to 125°C.

The cooling effectiveness of a package depends upon the thermal conduction of the package material, which consists of the package substrate and body, the package composition, and the effectiveness of the heat transfer between package and cooling medium.

Standard packaging approaches use still or circulating air as the cooling medium. The transfer efficiency can be improved by adding finned metal heat sinks to the package. More expensive packaging approaches, such as those used in mainframes or supercomputers, force air, liquids, or inert gases through tiny ducts in the package to achieve even greater cooling efficiencies. As an example, a 40-pin DIP has a thermal resistance of 38°C/W and 25°C/W for natural and forced convection air. This means that a DIP can dissipate 2 watts (3 watts) of power with natural (forced) air convection, and still keep the
3 The Manufacturing Process

temperature difference between the die and the environment below 75°C. For comparison, the thermal resistance of a ceramic PGA ranges from 15° to 30°C/W. Since packaging approaches with decreased thermal resistance are prohibitively expensive, keeping the power dissipation of an integrated circuit within bounds is an economic necessity. The increasing integration levels and circuit performance make this task nontrivial. An interesting relationship shown in equation (3.26) has been derived by Nagata. It provides a bound on the integration complexity and performance as a function of the thermal parameters.

\[
\frac{N_G}{t_p} \leq \frac{\Delta T}{\theta E}
\]  

(3.26)

where \(N_G\) is the number of gates on the chip, \(t_p\) the propagation delay, \(\Delta T\) the maximum temperature difference between chip and environment, \(\theta\) the thermal resistance between them, and \(E\) the switching energy of each gate.

Fortunately, not all gates are operating simultaneously in real systems. The maximum number of gates can be substantially larger, based on the activity coefficient in the circuit. For instance, it was experimentally derived that the ratio between the average switching period and the propagation delay ranges from 20 to 200 in mini- and large-scale computers. Nevertheless, equation (3.26) demonstrates that heat dissipation and thermal concerns present an important limitation on circuit integration. Design approaches for low power that reduce either \(E\) or the activity coefficient are rapidly gaining importance.

3.6 Pictorial View of Wafer Preparation and Wafer Fabrication

Pictorial view of some process steps for wafer preparation and wafer fabrication are shown in figures below. The pictures shown here are examples.
Figure 3.45: Silicon ingot

Figure 3.46: Cleaning room attire
3 The Manufacturing Process

**Figure 3.47:** Cleaning station

**Figure 3.48:** Epitaxial furnace
Figure 3.49: Wafer mask and 10X reticle

Figure 3.50: Lithography – pattern generator
Figure 3.51: Lithography – step and repeat machine

Figure 3.52: Diffusion furnace
Figure 3.53: Ion implanter

Figure 3.54: Etch station
3 The Manufacturing Process

**Figure 3.55:** Wafer with fabricated integrated circuit

**Figure 3.56:** Chemical-mechanical polishing machine
Exercises

2.1. Tetrahedral is the primitive cell of silicon crystal structure. Its lattice constant is 5.43\,\text{Å}. Calculate the distance between two nearest silicon atoms.

2.2. Tetrahedral is the primitive cell of silicon crystal structure. Its lattice constant is 5.43\,\text{Å}. Calculate the adjacent angle between two covalent bonds of the silicon atoms.

2.3. Calculate the surface density of (111) and (110) planes of the silicon crystal.

2.4. Al\textsubscript{x}Ga\textsubscript{1-x}As has zinc blende crystal structure and its density is \((5.36 - 1.6x)\) \text{gcm}\textsuperscript{-3}. Given that atomic weight of Al, Ga, and As are 26.98, 69.72, and 74.92g respectively. Calculate the surface density of (001) plane for \(x = 0.3\).

2.5. Given a silicon wafer to you, how do you identify its crystal orientation and the type of dopant it contains?

2.6. A silicon ingot contains \(10^{16}\) boron atoms \text{cm}\textsuperscript{-3} is to be grown by the Czochraski technique. What is the concentration of born atoms should be in the molten silicon to give the required concentration in ingot? If the initial load of silicon in crucible is 60kg, how many gram of boron (atom weight is 10.8g) should be added? The density of molten silicon is given to be 3.53\text{gcm}\textsuperscript{-3}.

2.7. Describe the purpose of photolithography.

2.8. List the steps involved in fabrication of a monolithic IC.

2.9. Explain how isolation between components is obtained in an IC.

2.10. Sketch the cross section of npn bipolar junction transistor.

2.11. Criticize the missing step and re-draw the flowchart mentioned in Fig. 3.13 to reflect the actual process steps for fabrication of an npn bipolar junction transistor.
Bibliography

3 The Manufacturing Process