Semiconductor Manufacturing Technology

Michael Quirk
Julian Serda

SEMICONDUCTOR MANUFACTURING TECHNOLOGY

MICHAEL QUIRK
STAKTEK CORPORATION

JULIAN SERDA
ADVANCED MICRO DEVICES

Upper Saddle River, New Jersey
Columbus, Ohio

Library of Congress Cataloging-in-Publication Data
Quirk, Michael
Semiconductor manufacturing technology / Michael Quirk, Julian Serda.
p. cm.
ISBN 0-13-081520-9
1. Electronic industries. 2. Semiconductors—Design and construction. I. Serda, Julian.
II. Title.

TK7836.Q27 2001

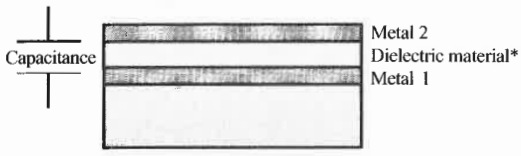
00-041653

Vice President and Publisher: Dave Garza
Editor in Chief: Stephen Helba
Assistant Vice President and Publisher: Charles E. Stewart, Jr.
Production Editor: Alexandrina Benedicto Wolf
Production Coordination: Lithokraft II
Design Coordinator: Robin G. Chukes
Cover Designer: Linda Fares
Cover Image: John Foxx
Production Manager: Matthew Ottenweller
Marketing Manager: Barbara Rose

This book was set in Times Roman by Lithokraft II. It was printed and bound by R. R. Donnelley & Sons Company. The cover was printed by Phoenix Color Corp.

10 9 8 7 6 5 4 3 2
ISBN 0-13-081520-9

CONTENTS



* *Low* k *dielectric reduces capacitance between metal layers.*

FIGURE 2.17 Low-*k* Dielectric Material

action occurs when a semiconductor is heated, thus improving conductivity with an increase in temperature (which is the opposite condition for a conductor).

The most important semiconductor material in wafer fabrication is silicon, used as the semiconductor wafer substrate in over 85% of all chips.[3] This text will only address silicon because of its predominant use in the semiconductor industry. A short review of some of the other semiconductor materials and their applications is given at the end of this chapter.

SILICON

Silicon is an elemental semiconductor material because it has four valence-shell electrons, along with the other elements located in Group IVA of the periodic table (see Figure 2.18 on page 34). The number of valence-shell electrons in silicon places it midway between the valence-shell condition of a good conductor (one valence electron) and an insulator (eight valence electrons).

Silicon is not found pure in nature. It must be refined and purified to become pure for semiconductor manufacturing (see Chapter 4). It usually occurs as silica (silicon dioxide, or SiO_2) and other silicates. Silica is sand and is a primary ingredient of glass. Other forms of SiO_2 are rock crystal, quartz, agate, and opal.

The melting point of silicon is 1412°C. Silicon is a hard and brittle material that fractures easily if deformed, much like a piece of glass. It can be polished to a mirrorlike finish. Silicon exhibits many of the same characteristics as metals and at the same time exhibits characteristics of nonmetals. This is why silicon is classified as a semiconductor, midway between the conductors (metals) and insulators (nonmetals) on the periodic table.

Pure Silicon

Pure silicon is referred to as *intrinsic silicon,* with no contaminants or impurities from other substances. The silicon atoms in pure silicon bond together through covalent bonds to share electrons and complete their valence shells (see Figure 2.19 on page 34).

Many of silicon's properties result from this strong covalent bonding. The covalent bonds in pure silicon hold the atoms together to form a solid, electrically stable material that is an insulator. Pure silicon is a poor conductor because all valence shells are fully occupied in covalent bonds. In this pure form, silicon is not useful as a semiconductor.

The solid material formed when two or more silicon atoms bond together in this set, repeatable pattern, as the one shown in Figure 2.19 on page 34, is referred to as a *crystal.* A crystal is a smooth, glassy solid that forms a lattice structure with a three-dimensional shape. An example of a crystal material is window glass. We will study more about crystals in Chapter 4 when we discuss silicon wafer preparation.

Why Silicon?

Germanium was the first material used for semiconductors in the 1940s and early 1950s, but as we learned in Chapter 1, it was soon replaced by silicon. Why was silicon chosen as the predominant semiconductor material? There are four main reasons for the selection of silicon as the primary semiconductor material:

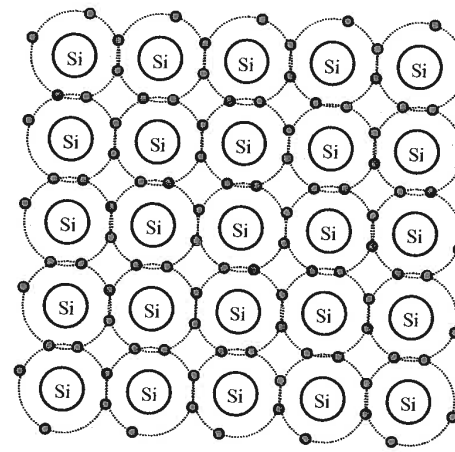
- Abundance of silicon
- Higher melting temperature for wider processing range

- Wider temperature range of operation
- Natural growth of silicon dioxide

Semiconductors

Group IVA	
C, Carbon	6
Si, Silicon	14
Ge, Germanium	32
Sn, Tin	50
Pb, Lead	82

FIGURE 2.18 Group IVA Elemental Semiconductors



Silicon atoms share valence electrons to form insulator-like bonds.

FIGURE 2.19 Covalent Bonding of Pure Silicon

Silicon is the second most abundant element on earth and makes up about 25% of the earth's crust. If processed properly, silicon can be refined into ample quantities of the very pure form necessary for semiconductor fabrication which leads to lower costs. Silicon's melting temperature of 1412°C is much higher than the melting temperature for germanium which is 937°C. This higher melting temperature permits silicon to withstand high-temperature processing. Another advantage to using silicon is that a semiconductor device made from silicon can function over a wider temperature range than germanium, which increases the semiconductor's applications and reliability.

Finally, an important reason for using silicon as a semiconductor material is the ability to naturally grow *silicon dioxide* (SiO_2) on the silicon surface (see Figure 2.20). SiO_2 is a high-quality, stable electrical insulator material that also serves as a good chemical barrier to protect silicon from external contaminants.[4] Electrical stability is important to avoid leakage between adjacent conductors in an IC. The ability to grow stable, thin SiO_2 material is fundamental to the fabrication of high-performance metal-oxide semiconductor (MOS) devices. SiO_2 has mechanical properties similar to silicon, which allows high temperature processing without excessive wafer warpage.

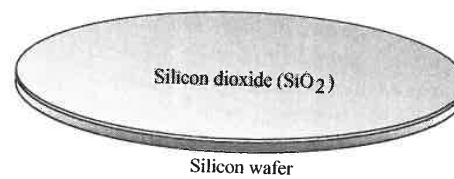


FIGURE 2.20 SiO_2 on Silicon Wafer

Doped Silicon

Silicon in its pure state has little use in semiconductor technology. However, the structure of silicon can be altered to greatly enhance its conductivity by adding small amounts of other elements to the material through a process known as doping. *Doping* is the process of adding certain elements to pure silicon to improve the conductivity of the semiconductor (see Figure 2.21). For example, the electrical resistivity (ρ) of pure silicon is approximately 2.5×10^5 ohms-cm. If only one in

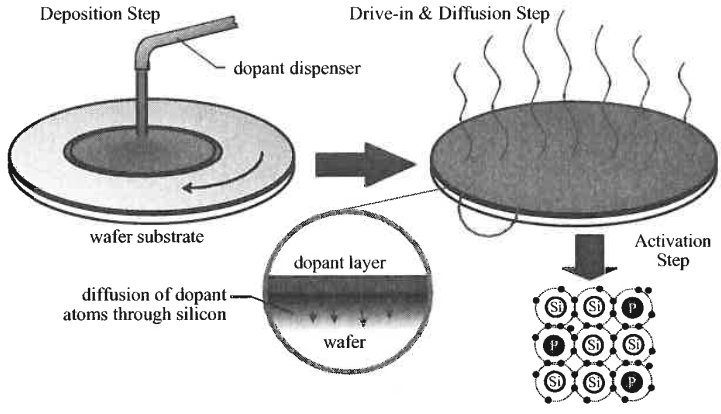


FIGURE 2.21 Doping of Silicon
(Used with permission from Advanced Micro Devices)

every one million silicon atoms is replaced by one atom of arsenic, resistivity will drop to 0.2 ohms-cm.[5] This is an improvement in conductivity of 1,250,000 times.

The elements added during doping are referred to as *dopants* or *impurities* because the silicon is no longer pure. In other words, we dope the silicon with an impurity so that it will conduct electricity. The more impurity added then the higher the conductivity (or the lower the resistivity). Note that we use the term impurity to indicate that we have added another element to the silicon. We intentionally add the impurity to increase the conductivity of silicon. Doped silicon is also known as *extrinsic silicon.*

The concept of doping the silicon with an impurity to improve electrical conductivity is a critical aspect of semiconductor fabrication. If we can introduce impurities that alter the silicon's resistivity and then control when the silicon performs as an insulator or as a conductor, then we have the essence of solid-state technology.

Dopant Materials ■ Silicon is located in Group IVA of the periodic table and has four valence electrons. Elements from the two adjacent groups are commonly used for doping: Group IIIA and Group VA (see Figure 2.22). Group IIIA elements are referred to as *trivalent* because of three valence electrons, whereas elements from group VA are *pentavalent* because of five valence electrons. The trivalent dopant increases the number of holes (positive doping or p-type), whereas the pentavalent will increase the number of free electrons (negative doping or n-type).

Acceptor Impurities		Semiconductor		Donor Impurities	
Group III (p-type)		**Group IV**		**Group V (n-type)**	
Boron	**5**	Carbon	6	Nitrogen	7
Aluminum	13	**Silicon**	**14**	**Phosphorus**	**15**
Gallium	31	Germanium	32	**Arsenic**	**33**
Indium	49	Tin	50	**Antimony**	**51**

* *Elements underlined are the most commonly used in silicon-based IC manufacturing.*

FIGURE 2.22 Silicon Dopants

When trivalent dopant atoms are added to silicon, the resulting material is called a *p-type silicon.* The trivalent dopants are referred to as *acceptors* (they accept an extra mobile electron), with the most common acceptor element being boron. When a pentavalent element is added to pure silicon, the resulting material is referred to as an *n-type silicon.* Pentavalent dopants are referred to as

donors (they donate an extra mobile electron), and are typically either phosphorus, arsenic, or antimony. The different doping elements are listed in Table 2.2.

TABLE 2.2 Dopant Elements Commonly Used in Semiconductor Manufacturing

Trivalent Dopants (p-type, positive doping, acceptor)	Pentavalent Dopants (n-type, negative doping, donor)
Boron (B)	Phosphorus (P)
	Arsenic (As)
	Antimony (Sb)

n-Type Silicon. For an n-type silicon, there are more electrons than valence-band holes. This is shown in Figure 2.23 for silicon with a pentavalent phosphorous dopant atom.

The silicon atoms will form covalent bonds with the donor phosphorous atom, each sharing one phosphorous electron. However, the fifth phosphorous electron is not bound to any of the surrounding silicon atoms. Because of this, the fifth phosphorous electron needs little energy to break away and enter the conduction band. For n-type silicon, conduction-band free electrons are the majority carriers, of which the material has an abundance. There are also many fewer minority carrier valence-band holes.

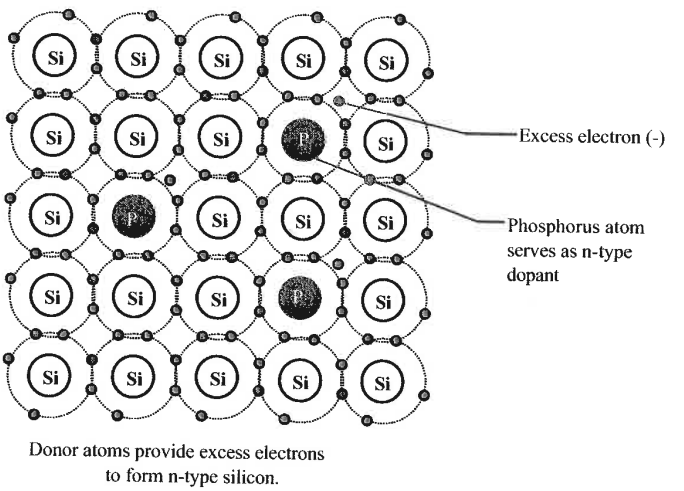


FIGURE 2.23 Electrons in n-Type Silicon with Phosphorous Dopant

One electron in the conduction band is not significant for conduction. However, when we dope silicon with a dopant, we add literally millions of dopant atoms, producing many electrons that are not part of the covalent bonds. There is a high amount of movement between electrons and holes. Negative electrons are attracted to the positive holes. Electrons enter the conduction band relatively easily. If a voltage is applied to this material, the electrons can be made to flow through the material as electric current (see Figure 2.24).

Note that the doped silicon is still electrically neutral (this is true for n-type or p-type silicon). In the case of n-type silicon, this is because each phosphorous atom still has the same number of protons as electrons, as do the silicon atoms. Thus the overall number of protons and electrons in the semiconductor is still equal and the result is a net charge of zero. What is not equal is that there are many more conduction-band electrons (majority carriers) than there are valence-band holes (minority carriers).

p-Type Silicon. In the p-type silicon shown in Figure 2.25, the boron atom is a p-type acceptor dopant that forms a covalent bond with four adjacent silicon atoms. The boron acceptor atoms provide a deficiency of electrons caused by the lack of a fourth electron in the boron

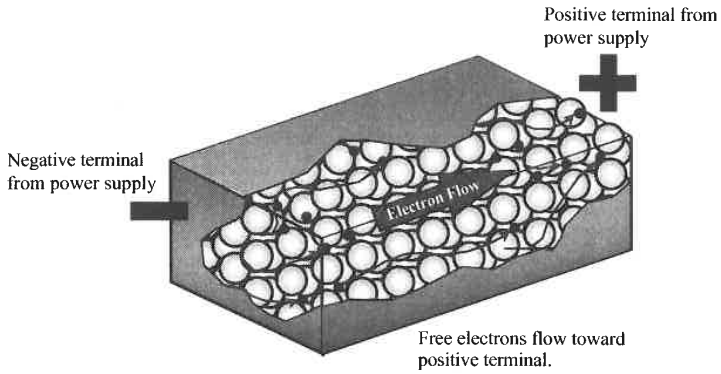


FIGURE 2.24 Flow of Free Electrons in n-Type Silicon

atom, thus creating the p-type silicon. There is an excess of holes (absence of electrons) in the valence shell.

Since there are many more valence-band holes than conduction-band electrons, the holes are the major current carriers in p-type material. The holes are referred to as *majority carriers* and electrons are the *minority carriers.* If a dc voltage is applied across a p-type semiconductor, then the large number of holes in the material will attract electrons from the negative terminal of the voltage source into the p-type semiconductor. This is current flow in a p-type material (see Figure 2.26 on page 38). The holes appear to move because each time an electron moves into a hole it serves to create a hole in its previous position. The holes seemingly move in the opposite direction as the electrons.

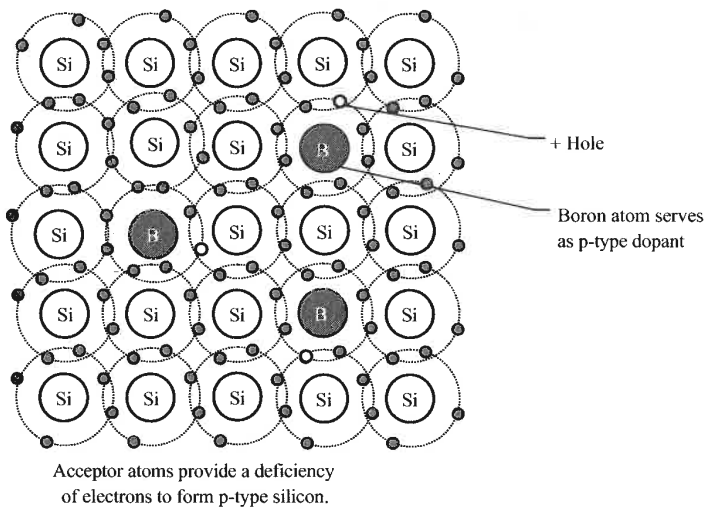


FIGURE 2.25 Holes in p-Type Silicon with Boron Dopant

Resistivity of Doped Silicon. We obtain precise control of silicon's resistivity by introducing dopants into the silicon crystal structure. The concentration of the dopant atom placed in the silicon crystal defines how well the material will conduct electrical current. Pure silicon has a resistivity of about 250,000 Ω-cm and is an insulator. Compare this to copper, a good conductor, with a resistivity of 1.7 μΩ-cm (0.0000017 Ω-cm). By adding the proper type and concentration of dopants to pure silicon, the doped silicon's resistivity is decreased and conductivity is improved (see Figure 2.27 on page 38). For a given resistivity there is less concentration of n-type dopants than p-type dopants. This is because it takes less energy to move an electron than to move a hole.

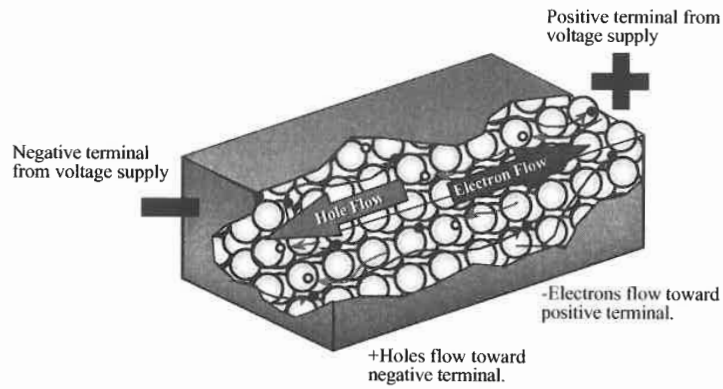


FIGURE 2.26 Flow of +Holes in p-Type Silicon

It only takes a small amount of dopant (from as little as 0.000001% to 0.1%) to make silicon a useful conductor. This is important for the fabrication of semiconductor devices on wafers. The amount of dopant in the silicon, or concentration, is carefully controlled during semiconductor fabrication to attain a precise resistivity. The ion implantation process for adding dopants to silicon will be covered in Chapter 17.

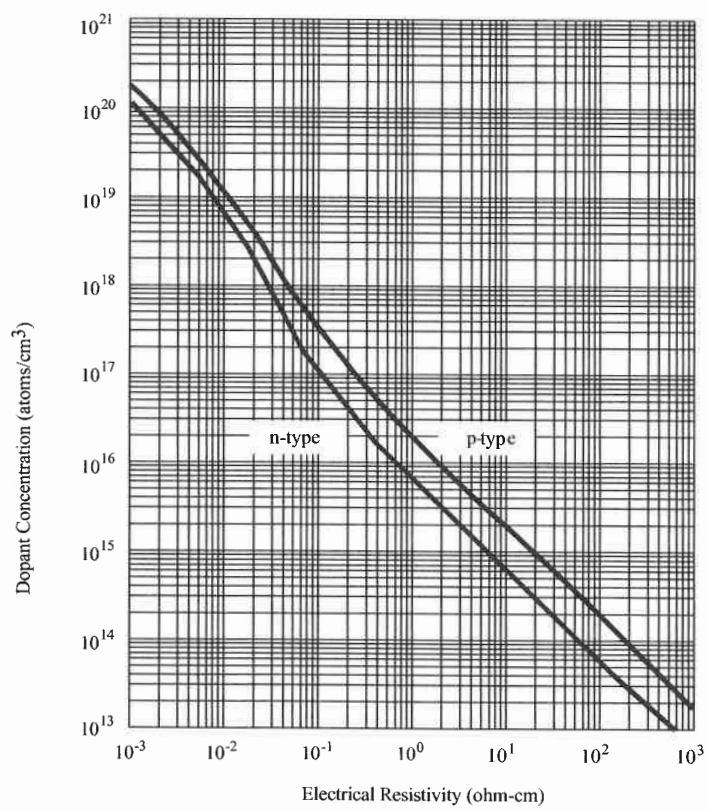


FIGURE 2.27 Silicon Resistivity Versus Dopant Concentration
Redrawn from S. Ghandi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, John Wiley & Sons, 1994.

pn Junctions

We dope pure silicon with pentavalent or trivalent elements to obtain either an n-type or p-type semiconductor. The type and concentration of the dopant element determines whether conduction occurs by hole or by electron and determines the final value of the silicon's resistivity. The number of carriers, and therefore the resistivity, is determined by the net donor and acceptor atom concentration in the silicon crystal. Because of this, an n-dopant can be implanted or diffused into a p-region to convert that region into an n-region (or vice versa). To achieve this, the concentration of the n-type donor dopant must exceed that of the p-type acceptor dopants in the region.

The ability to incorporate both n-type and p-type regions in the silicon crystal is beneficial because semiconductor devices require both of these regions in order to function as a useful electronic device. It is the junction between the n-type and p-type regions that is important and creates the useful characteristics of silicon as a semiconductor. This junction is referred to as a *pn junction* (see Figure 2.28).

The pn junction is the essence of solid-state electronics and is the basis for how semiconductor wafers can achieve their unique ability to act as an insulator and a conductor, depending on the bias voltage applied to the junction. The specifics of how a pn junction functions as a useful electronic device are discussed in Chapter 3. The creation of a pn junction is almost universally done in wafer fabrication with ion implantation (see Chapter 17).

Remember that a pn junction is formed between two pieces of essentially the same material. The p-type and the n-type materials scarcely differ except for a minute amount of a dopant. The n-type material has extra mobile electrons from the donor impurity, whereas the p-type material has extra mobile holes. It is deceptive to speak of one material meeting the other.[6] The junction is so intimate that the n-type and p-type materials are formed from one continuous solid. The silicon crystal with a pn junction continues to look like and in most ways behave like the solid crystal material that it is.

The depth and definition of this junction is critical in semiconductor manufacturing. As device critical dimensions are reduced, the ability to precisely control both the pn junction in silicon (e.g., junction depth) and the concentration of dopants is becoming a major challenge for semiconductor chip fabrication.

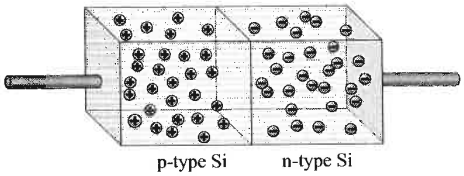


FIGURE 2.28 Cross Section of Planar pn Junction

ALTERNATIVE SEMICONDUCTOR MATERIALS

Germanium and silicon are the two elemental semiconductor materials from Group IVA with four valence electrons. We have learned that germanium was the first semiconductor material used for transistor fabrication and was replaced by silicon in the 1950s for process and performance reasons.

There are alternative semiconductor materials used for specific market applications. These are primarily the *compound semiconductors*. A major class of compound semiconductors are formed from Group IIIA and Group VA of the periodic table (often referred to as III-V compounds). An example is gallium arsenide (GaAs). In addition, other semiconductor compounds come from elements found in Groups IIA and VIA, and are referred to as II-VI compounds.

The need for alternative semiconductor materials is driven by the growth of markets that require IC performance beyond that capable by silicon semiconductors. The key IC performance factor is speed. Wireless and high-speed digital communications, space applications, and consumer markets such as the automotive industry, are developing special niche semiconductor markets that

Oxide growth is a natural phenomenon that occurs by exposing the silicon wafer to oxygen at an elevated temperature. The ability to grow an oxide film on silicon is often listed as a significant reason for the use of silicon as the most common semiconductor substrate material (another major reason is silicon's relatively high melting temperature). We use the term "grow" to indicate that temperature is used to cause the oxide to grow out of the silicon semiconductor material, actually consuming silicon in the process.

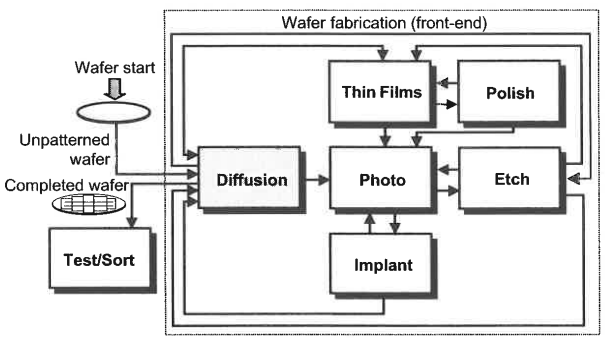


FIGURE 10.1 Diffusion Area of Wafer Fabrication
(Used with permission from Advanced Micro Devices)

The amount of thermal exposure (i.e., temperature multiplied by time) for a wafer during processing is referred to as the *thermal budget.* The thermal budget requirement for wafer fabrication is rapidly dropping.[2] A goal throughout semiconductor processing is to minimize the amount of thermal (heat) exposure to the wafer. As the critical dimension on device structures decreases to 0.18 μm and beyond, scaling requirements dictate shallower junction depths. To minimize the unacceptable diffusion of dopants out of a shallow junction region, the thermal budget must decrease accordingly. Another problem caused by excessive thermal budget is the increase in ohmic contact resistance of metal-layer interconnects (see Chapter 12), which increases the overall resistance of the conductive paths. A factor in determining the process conditions for many wafer fabrication process steps is the ability to minimize the thermal budget, either by reduced temperature or by minimizing the time at temperature.

OXIDE FILM

An oxide layer is grown on silicon typically at thermal oxidation temperatures between 750°C to 1100°C. An oxide layer grown on silicon is referred to as *thermal oxide* or *thermal silicon dioxide* (SiO_2). Since silicon dioxide is an oxide material, the two terms are used interchangeably. Another term for silicon dioxide is *glass*. Silicon dioxide is a dielectric material and will not conduct electricity.

Nature of Oxide Film

When a silicon surface is exposed to oxygen, an amorphous silicon dioxide film grows immediately. Although undoped silicon is a semiconductor material, SiO_2 is an insulator. The atomic structure of this SiO_2 film consists of a silicon atom surrounded by four oxygen atoms (see Figure 10.2). We can refer to this atomic structure as a silicon dioxide *tetrahedron cell.* Amorphous SiO_2 has no long-range periodic crystal order at the atomic level. Long-range order is absent because the tetrahedra are not arranged in a regular, three-dimensional array as in a crystal.[3]

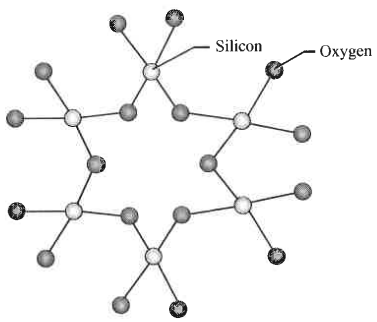


FIGURE 10.2 Atomic Structure of Silicon Dioxide
(Used with permission from International SEMATECH)

Silicon dioxide is a form of intrinsic (pure) glass with a melting temperature of 1732°C.[4] Thermally grown SiO_2 has strong adhesion to silicon and exhibits excellent dielectric properties. A layer of SiO_2 always covers the wafer surface because silicon oxidizes immediately upon exposure to air with a few monolayers of native oxide (a monolayer is one layer of atoms on a surface). Over time, the native oxide layer will thicken to an upper limit of about 40 Å, even at 25°C room temperature.[5] This type of oxide is nonuniform and usually considered a contaminant, with little use in semiconductor fabrication. There are some applications for use of native oxide in wafer fabrication, such as part of a dielectric film stack in memory cells. Native oxide grown from air is contaminated, but native oxide grown in chemical baths using high-purity chemicals is very clean.

Uses of Oxide Film

Oxide is important for silicon semiconductor fabrication because of its ease of formation and its excellent interface with the underlying silicon material. It is the most common layer used in semiconductor fabrication. Different ways in which an oxide layer is used to fabricate a microchip are:

- Device scratch protection and contaminant isolation
- Field isolation to confine charged carriers (termed surface passivation)
- Dielectric material in the gate oxide or memory cell structures
- Impurity-mask barrier during doping
- Dielectric layer between metal conductor layers

Device Protection and Isolation ■ Silicon dioxide grown on the surface of the wafer serves as an effective barrier to isolate and protect sensitive devices in the silicon. SiO_2 physically protects devices because it is a very hard and nonporous (dense) material that effectively insulates active devices in the silicon surface. The hard SiO_2 layer will protect the silicon from scratches and processing damage that might occur during fabrication. Transistors have traditionally been electrically isolated by thermally growing thick SiO_2 in the region between them using the LOCOS process (discussed later in this chapter). However, this process is unacceptable for sub-0.25 μm technology and has been replaced by shallow trench isolation (STI). The STI process uses deposited oxide as the main dielectric (see Chapters 9 and 11).

Surface Passivation ■ A major benefit from thermally grown SiO_2 is the reduction in the surface-state density of silicon that occurs by tying up dangling silicon bonds. This action is referred to as *surface passivation,* which prevents the deterioration of the electronic properties and reduces current-leakage paths of the semiconductor caused by moisture, ions or other external contaminants.[6] The hard SiO_2 layer will protect the silicon from scratches and processing damage that might occur during backend final fabrication. SiO_2 growth on the silicon surface can serve to tie up electrically active contaminants (mobile ionic contaminants) in the oxide layer away from the active surface of the silicon. Passivation is also important for controlling the leakage current of

junction devices and growing a stable gate oxide.[7] For the oxide layer to function as a good passivation layer, it has a quality requirement to be a uniform thickness and not have pinholes or voids in the material.

Another factor in using the oxide layer to passify the wafer surface is the oxide thickness. Adequate oxide thickness is necessary to prevent electrical charging in a metal layer caused by a charge buildup in the silicon wafer surface, much like charge storage and breakdown characteristics of ordinary capacitors. This charging would lead to electrical shorting and other undesirable electrical effects. A thick layer of oxide that inhibits charge buildup from metal layers is the *field oxide layer,* which typically is between 2,500 Å to 15,000 Å thick (see Figure 10.3).

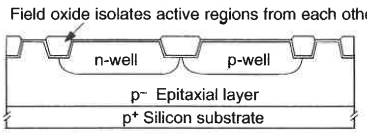


FIGURE 10.3 Field Oxide Layer

Silicon dioxide also has a coefficient of thermal expansion very similar to that of silicon. During high-temperature processing, the wafer will expand and then contract during cooling. The SiO_2 will expand and contract at close to the same rate as silicon, thus minimizing wafer warpage during the thermal excursions. This action also prevents film stress from separating the oxide film from the silicon substrate.

Gate Oxide Dielectric ■ An extremely thin layer of oxide is used as the dielectric material for the important *gate oxide structure* common in MOS technology (see Figure 10.4). The gate oxide is thermally grown because of its high quality and stability with the underlying semiconductor silicon. SiO_2 has a high dielectric strength of 10^7 volts/cm and high resistivity of approximately 10^{17} ohms-cm.[8]

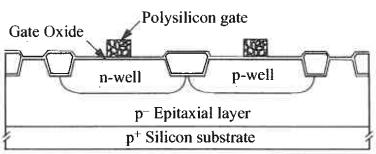


FIGURE 10.4 Gate Oxide Dielectric

The thickness of the gate oxide is chosen specifically for the scaling requirements of the device technology. For 0.18-μm generation technology, a typical gate oxide thickness is 20 ± 1.5 Å. The gate oxide has a specific thickness so that it scales properly with the entire gate structure to permit induction of a charge in the silicon wafer under the oxide. The widespread use of MOS technology in the ULSI era has made the formation of the gate oxide a primary concern in process development. A critical aspect of device reliability is *gate oxide integrity.* The gate structure in a MOS device is the part that controls the flow of electrical current through the device. Because this oxide is fundamental to the proper function of microchips based on field-effect technology, it is essential that it have high quality, with excellent film-thickness uniformity and no contaminants. Any contamination that degrades the proper functioning (integrity) of the gate oxide structure must be scrupulously controlled.

Dopant Barrier ■ Silicon dioxide can be used as an effective barrier to selectively introduce dopants into the silicon surface (see Figure 10.5). Once an oxide layer is grown on the silicon surface, then mask openings are etched into the SiO_2 to create windows where dopant materials can enter the wafer. The oxide protects the surface of silicon from impurity diffusion where there are

no openings and therefore allows selective impurity doping. Dopants have a slow rate of movement through SiO_2 when compared to silicon, requiring only a thin oxide layer to block dopants (note that this rate is temperature dependent).

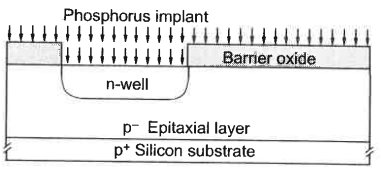


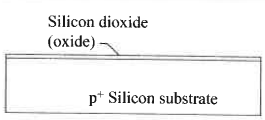
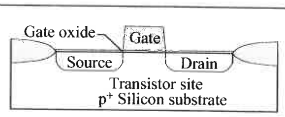
FIGURE 10.5 Oxide Layer Dopant Barrier

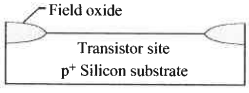
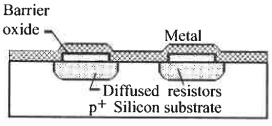
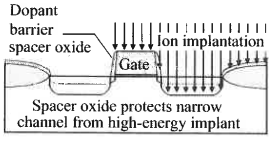
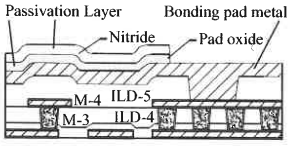
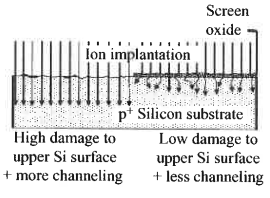
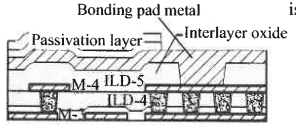
A thin oxide layer (e.g., 150 Å) is also grown in areas where ion implant occurs. This oxide screen serves to reduce damage to the silicon wafer and obtain better control over the depth that the dopant is implanted into the silicon by reducing channeling effects (see Chapter 17). Once implantation is done, HF (hydrofluoric acid) can be used to selectively remove the oxide, resulting in a planar silicon surface again.

Dielectric Between Metal Layers ■ Under normal conditions, silicon dioxide will not conduct electricity. For this reason, SiO_2 is an effective insulator between metal layers on the microchip. Silicon dioxide prevents shorting of an upper metal layer to the underlying metal layer, just as the insulator material on an electrical cord prevents an electrical short. The oxide quality is critical with no holes or voids allowed. This oxide is often doped to produce more effective flow properties and to better minimize the diffusion of contaminants (i.e., it acts as a getter). It is typically deposited using chemical vapor deposition (not thermally grown) and will be covered in more detail in Chapter 11.

A summary of different types of oxides and their uses in semiconductor fabrication is shown in Table 10.1.[9]

TABLE 10.1 Oxide Applications

Application	Purpose	Structure	Comments
Native Oxide	This oxide is a contaminant and is generally undesirable. Sometimes used in memory storage or film passivation.		Growth rate at room temperature is 15 Å per hour up to a maximum thickness of about 40 Å.
Gate Oxide	Serves as a dielectric between the gate and source-drain parts of MOS transistor.		Common gate oxide film thickness range from about 20 Å to several hundred Å. Dry thermal oxidation is the preferred growth method.

Application	Purpose	Structure	Comments
Field Oxide	Serves as an isolation barrier between individual transistors to isolate them from each other.		Common field oxide film thickness ranges from 2,500 Å to 15,000 Å. Wet oxidation is the preferred growth method.
Barrier Oxide	Protects active devices and silicon from follow-on processing.		Thermally grown to several hundred angstroms thickness.
Dopant Barrier	Masking material for depositing or implanting dopants into wafer.		Dopants diffuse into unmasked areas of silicon by selective diffusion.
Pad Oxide	Provides stress reduction for silicon nitride (Si_3N_4).		Thermally grown and very thin.
Implant Screen Oxide	Used to reduce implant channeling and damage.		Thermally grown.
Insulating Barrier Between Metal Layers	Serves as a protective layer between metal lines.		This oxide is not thermally grown, but is deposited.

CHAPTER 17

ION IMPLANT

Intrinsic silicon crystal structure is formed by silicon covalent bonds. As explained in Chapter 2, intrinsic silicon is a poor conductor. Silicon only becomes useful as a semiconductor when its structure and conductivity are altered by adding small amounts of impurities called dopants. This process is referred to as doping. The doping of silicon is fundamental to pn junction fabrication in semiconductor devices. Ion implantation is the primary method for doping wafers.

Doping is widely used throughout wafer fabrication. Dopants are precisely and uniformly placed into the device to alter its electrical performance. Silicon chips require doping with Group IIIA and Group VA dopants for many different reasons, which will be discussed in this chapter.

The continual shrinking of chip critical dimensions and the subsequent increase in chip packing density force the scaling of all device elements. An example of this is the reduction of the polysilicon gate length to obtain a narrower channel region. A shorter channel length requires shallower junction depths for the source and drain doped regions. Junction depths are ultrashallow for chips designed at the 0.18 μm and below technology nodes. The junction depth requirements decrease to 30 ± 10 nm for devices with a CD of 0.1 μm.[1]

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Explain the purpose and applications for doping in wafer fabrication.
2. Discuss the principles and process of dopant diffusion.
3. Provide an overview of ion implantation, including its advantages and disadvantages.
4. Discuss the importance of dose and range in ion implant.
5. List and describe the five major subsystems for an ion implanter.
6. Explain annealing and channeling in ion implantation.
7. Describe different applications of ion implantation.

INTRODUCTION

Doping is the introduction of a dopant into the crystal structure of a semiconductor material to modify its electronic properties (e.g., electrical resistivity). As discussed in Chapter 2, certain Group IIIA and VA elements are used as dopants in wafer fabrication. For review, Table 17.1 on page 476 highlights the four principal dopants used in semiconductor manufacturing.

Dopant species are also referred to as impurities but should not be confused with contaminating impurities. Dopants are introduced into the silicon and other fabrication materials for a variety of reasons. For instance, dopants such as boron and phosphorus are used to form the majority carriers in silicon devices, to form conductive layers in the wafer, or to alter material properties (e.g., doping SiO_2 to form borophosphosilicate glass, or BPSG). The polysilicon gate electrode is doped for conductivity.

There are two techniques in wafer fabrication for introducing dopant elements to the wafer materials: thermal diffusion and ion implantation. *Thermal diffusion* uses high temperature to move the dopant through the silicon lattice structure. This method is dependent on time and temperature. *Ion implantation* introduces dopants into the substrate through a high-voltage ion bombardment. The dopants are implanted into the wafer through high-energy collisions at the atomic level. In the beginning of semiconductor manufacturing, thermal diffusion was the primary method for doping the wafer. However, due to shrinking critical dimensions and the corresponding scaling requirements, nearly all dopant processes in modern wafer fabrication are now done

TABLE 17.1 Common Dopants Used in Semiconductor Manufacturing

Acceptor Dopant Group IIIA (p-Type)		Semiconductor Group IVA		Donor Dopant Group VA (n-Type)	
Element	**Atomic Number**	**Element**	**Atomic Number**	**Element**	**Atomic Number**
Boron (B)	5	Carbon	6	Nitrogen	7
Aluminum	13	Silicon (Si)	14	Phosphorus (P)	15
Gallium	31	Germanium	32	Arsenic (As)	33
Indium	49	Tin	50	Antimony	51

Note: Common dopants are shaded.

with ion implantation (see Figure 17.1). Table 17.2 outlines some critical processes commonly doped with ion implantation versus the limited doping done by diffusion. Figure 17.1 shows these doped regions on a cross section drawing of a CMOS inverter. Note the dopant type (p or n) and concentration (− or +) are also shown. The different doped regions labeled A to O are defined in Table 17.2.

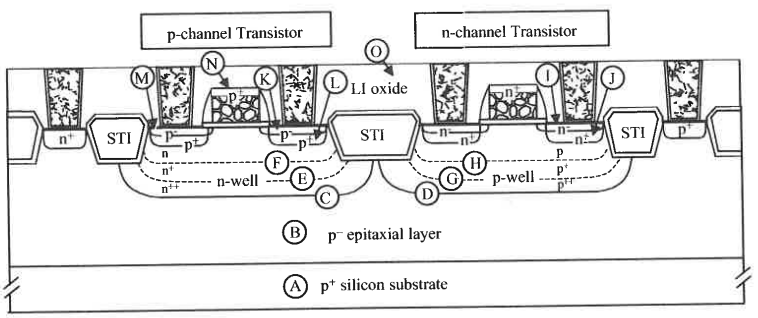


FIGURE 17.1 CMOS Structure with Doped Regions

TABLE 17.2* Common Dopant Processes CMOS Fabrication

Process Step	Common Dopant Species	Ion Implant or Diffusion	Remarks
A. p^+ Silicon Substrate	B	Diffusion	The monosilicon crystal is doped during the crystal growing process.
B. p^- Epitaxial Layer	B	Diffusion	The epitaxial layer is doped by diffusion during epitaxial silicon growth.
C. Retrograde n-well	P	Ion Implant	Retrograde well has dopant profile with peak concentration buried at a certain depth. The dopant concentration decreases as it approaches the surface.
D. Retrograde p-well	B	Ion Implant	Retrograde well has dopant profile with peak concentration buried at a certain depth.
E. p-Channel Punchthrough	P	Ion Implant	Phosphorus implanted more laterally than source/drain implants to keep electric field of drain from punching through p-channel and reaching source region.

Process Step	Common Dopant Species	Ion Implant or Diffusion	Remarks
F. p-Channel Threshold Voltage (V_T) Adjust	P	Ion Implant	Implant phosphorus to adjust MOS V_T threshold voltage.
G. n-Channel Punchthrough	B	Ion Implant	Boron is implanted more laterally than source/drain implants to keep electric field of drain from punching through n-channel and reaching source region.
H. n-Channel V_T Adjust	B	Ion Implant	Implant boron through oxide layer to adjust MOS threshold voltage.
I. n-Channel Lightly Doped Drain (LDD)	As	Ion Implant	Low-dose implant of arsenic in region adjacent to the n-channel to improve electrical performance by reducing the peak electrical field and therefore hot carriers, which minimizes interface charges trapped in the gate oxide.
J. n-Channel Source/Drain (S/D)	As	Ion Implant	Higher implant dose of arsenic to form n-channel source and drain.
K. p-Channel LDD	BF_2	Ion Implant	Low-dose implant of boron in region adjacent to the p-channel to improve electrical performance between drain and channel region.
L. p-Channel S/D	BF_2	Ion Implant	Higher implant dose of boron to form p-channel source and drain.
M. Silicon	Si	Ion Implant	Implant with nondopant atom to amorphize silicon to reduce transient enhanced diffusion (TED) and channeling effects.
N. Doped Polysilicon	P or B	Ion Implant or Diffusion	Doping of polysilicon gate electrode to make conductive.
O. Doped SiO_2	P or B	Ion Implant or Diffusion	Doped oxide to obtain material benefits (e.g., create better flow and serve as getter).

*Adapted from E. Rimini, *Ion Implantation: Basics to Device Fabrication,* (Boston: Kluwer Academic Publishers, 1995).

Ion implantation typically follows photolithography in the general process flow (see Figure 17.2 on page 478); although, there may be some processes that call for the use of an etched-oxide layer to serve as a mask during either diffusion doping or ion implant. In this case, wafers could flow from etch to diffusion or ion implant. Ion implantation plays a crucial role in chip performance improvement because of the need for smaller dimensions. Transistor performance is increasingly dependent on precise doping profiles in the silicon that are achievable only with ion implant.

Doped Regions

Recall from Chapter 4 that a wafer is uniformly doped with dopant atoms during the crystal growth process to create a p-type or n-type wafer. The type of dopant depends on the manufacturer's specification. During wafer fabrication, dopants are selectively introduced to create the devices on the individual wafers. The selective introduction of dopants into masked openings on the wafer forms a *doped region* where dopants reside within the silicon crystal structure (see Figure 17.3 on page 478). The doped region is characterized by its *dopant profile,* which determines the amount of dopant actually added in the doped region as a function of depth.

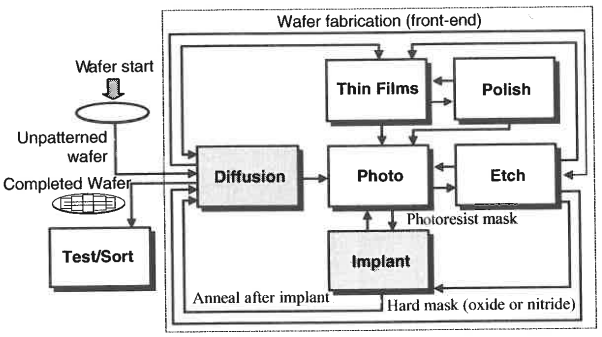


FIGURE 17.2 Ion Implant in Wafer Process Flow
(Used with permission from Advanced Micro Devices)

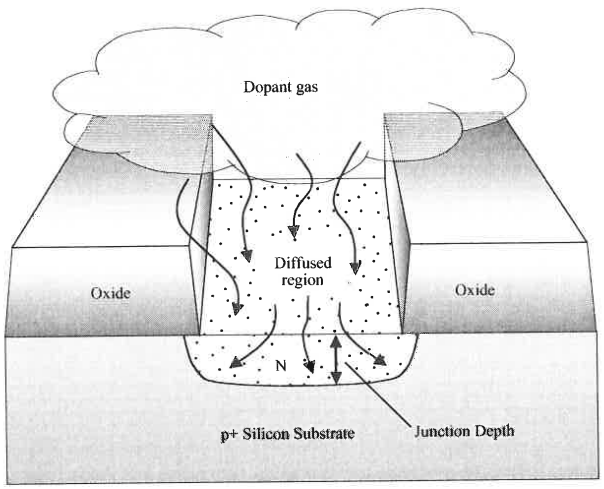


FIGURE 17.3 Doped Region in a Silicon Wafer

The dopant atoms enter the solid silicon through an oxide mask opening. In ion implantation, this mask is often photoresist, but a variety of masks can be used since it is a low-temperature process. Doping by diffusion requires an oxide or nitride mask because this technique is done in a high-temperature furnace.

The doped region can be of the opposite type as the silicon substrate, such as an n-type doped region in a p-type wafer, or it can be the same type with a different dopant concentration (e.g., a heavily doped p^+ region in a p-type wafer). The location where the conductivity of the doped region makes a change from p-type to n-type, or vice versa, is the pn junction. The exact depth in the wafer where the p-type dopant meets the n-type dopant is the *junction depth*, x_j. At the precise location of the junction depth, the concentration of electrons equals the concentration of holes. Another way of looking at this feature is the net dopant concentration at x_j is zero.

The ultimate interest in wafer doping is related to the individual dopant profiles for the devices after all processing is complete. A wafer undergoes many thermal excursions during processing, such as oxide growth, CVD, and so on, all contributing to undesirable dopant diffusion through the silicon. Dopant diffusion can alter critical parameters (e.g., junction depth and concentration) of the initial doped region and affect device performance. This is an important reason for minimizing the wafer thermal budget.

The main focus of this chapter will be doping by ion implantation. We will first review diffusion to analyze basic concepts and explain how intentional and unintentional diffusion occurs in wafer fabrication.

DIFFUSION

Diffusion is a basic property of matter that describes the movement of one material through another. Diffusion occurs from a region of relatively higher concentration into a region of lower concentration, resulting from the motion of atoms, molecules, or ions. In semiconductor fabrication, high-temperature diffusion is the mechanism used to move a dopant through the silicon crystal lattice. Diffusion can occur as a gas-state, liquid-state or solid-state. An example of liquid-state diffusion is a drop of food colorant in a glass of water. The highly concentrated drop of colorant immediately starts to diffuse into the glass of water and will continue until the entire volume of water is the same color as the drop of colorant. The same type of diffusion occurs in the solid-state when a high concentration of dopant material is introduced to a wafer. The dopant material diffuses through the silicon atoms of the solid crystal lattice.

Diffusion Principles

Thermal diffusion of a solid-state dopant in silicon requires three steps: predeposition, drive-in, and activation. During *predeposition,* or *predep,* wafers are loaded into a high-temperature diffusion furnace and a quantity of dopant atoms is transferred from the source cabinet to the diffusion furnace. The furnace temperature setting is typically 800 to 1100°C for 10 to 30 minutes. The dopants are introduced into a very thin layer of the wafer while maintaining a constant concentration of dopants at the surface. A thin oxide layer, referred to as a *cap oxide,* is grown on the surface to prevent dopant atoms from diffusing out of the silicon. The total number of dopant atoms deposited is referred to as *Q*. Note that for ion implantation (discussed later in this chapter), Q is referred to as the dose and also represents the number of dopant atoms implanted.

The predeposition step establishes the concentration gradient for the entire diffusion process. The concentration of dopants is highest at the surface and reduces when moving away from the surface, forming a gradient (or slope). This slope establishes the dopant profile, which can be mapped using four point probes (see Chapter 7). *Fick's laws,* first postulated in 1855, mathematically describe the diffusion process. These laws state in part that the number of particles moving through a cross-sectional area is proportional to the concentration gradient.[2] Fick's laws are used to predict the dopant concentration at some distance, *x*, from the surface.

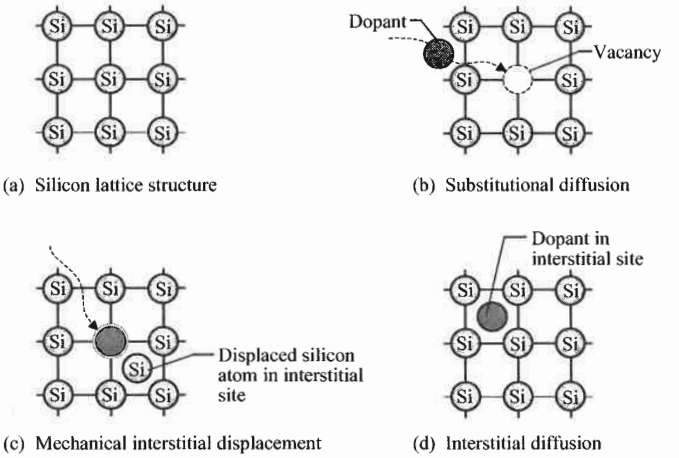
The second part of thermal diffusion is the *drive-in* step. This is a high-temperature process (1000 to 1250°C) used to move the deposited dopants through the silicon crystal to the desired junction depth in the wafer. No additional dopants are added to the wafer. However, the wafer surface oxidizes in the high-temperature environment, which affects the diffusion of the dopant atoms during drive-in. Some dopants, such as boron, tend to move into the growing oxide layer while others, such as phosphorus, are pushed away from the SiO_2. This modification of dopant concentration due to the silicon surface oxidation is termed *redistribution.*

The third part of thermal diffusion is the *activation* step. Here the temperature of the furnace is raised slightly higher to enable the dopant atoms to bond with the silicon atoms in the lattice structure. This action activates the dopant atoms, which changes the conductivity of silicon.

Dopant Movement ■ Each dopant has a particular *diffusivity* in silicon, which represents the rate (or speed) that the dopant moves in the wafer. The higher the diffusivity, the faster the dopant moves. Diffusivity increases with temperature, which is reflected in a term known as the *diffusion coefficient, D,* of the dopant being diffused. Elements have different diffusion coefficients within the temperature range of the drive-in step that reflect whether they are fast or slow diffusers in silicon. The diffusion coefficient is an equation variable in Fick's laws used to predict the final dopant concentration in silicon.

Within the wafer, dopant atoms move by two different mechanisms: interstitional and substitutional (see Figure 17.4 on page 480). Dopants with high diffusivity, such as gold (Au), copper (Cu), and nickel (Ni), use primarily *interstitional movement* to move between the interstitial space

between regular crystal sites of the silicon lattice. Slower moving dopants, such as arsenic and phosphorus that are common in semiconductor doping, typically use *substitutional movement* where atoms fill empty crystal positions in the lattice.



(a) Silicon lattice structure
(b) Substitutional diffusion
(c) Mechanical interstitial displacement
(d) Interstitial diffusion

FIGURE 17.4 Dopant Diffusion in Silicon

Dopants are useful to form semiconductor silicon only if they are *activated dopants,* meaning they are part of the silicon lattice structure. An activated dopant can act as a donor or accepter of electrons and is an n-type or p-type dopant with respect to silicon. If the dopant occupies interstitial space, then it is not activated and is ineffective as a dopant. Thermal energy moves dopants into the regular crystal sites, a process known as *crystal activation.* Activation is a regular part of diffusion because it occurs at a high temperature. For ion implant, crystal activation is done with an anneal step.

Solid Solubility ■ At a given temperature, there is a limit to how much dopant can be absorbed by the silicon. This is referred to as the *solid solubility limit.* This is true for most materials, such as pouring salt into a container of water and watching it dissolve. If salt is continually added, at some point the solubility of the salt in the water reaches a limit where it no longer dissolves. Each particular dopant has a solid solubility limit (see Table 17.3). Note that only a fraction of the dopants in silicon are actually activated and contribute electrons or holes for conduction (about 3 to 5%).[3] Most dopants remain in interstitial sites and are electrically inactive.

TABLE 17.3* Solid Solubility Limits in Silicon at 1100°C

Dopant	Solubility Limit (atoms/cm^3)
Arsenic (As)	1.7×10^{21}
Phosphorus (P)	1.1×10^{21}
Boron (B)	2.2×10^{20}
Antimony (Sb)	5.0×10^{19}
Aluminum (Al)	1.8×10^{19}

*SEMATECH "Diffusion Processes," *Furnace Processes and Related Topics* (Austin, TX: SEMATECH, 1994), p. 15.

Lateral Diffusion ■ The diffusion mask is either silicon dioxide (SiO_2) or silicon nitride (Si_3N_4) since a photoresist mask could not withstand the high-temperature process. When atoms diffuse into the wafer, they move in all directions: down into the silicon, laterally, and back out of the wafer. *Lateral diffusion* occurs when dopant atoms travel in a direction along the surface of the wafer. Typically, lateral diffusion during a thermal diffusion process is 75 to 85% of the vertical junction

depth.[4] Lateral diffusion is undesirable in advanced MOS circuits because it can cause a reduction in the channel length, affecting device density and performance.

Diffusion Process

The objective of the diffusion process is to bring the diffusing impurity in contact with a wafer and maintain the specified time and temperature for diffusion to occur. Diffusion takes place in high-temperature diffusion furnaces, discussed in Chapter 10. Typically, one furnace is used for the predeposition, drive-in, and activation steps. Quartz and other furnace parts used in diffusion processes should be segregated from other furnace processes to avoid cross-contamination.

Diffusion should produce reproducible results from run-to-run and wafer-to-wafer. The eight steps required to properly perform diffusion in wafer fabrication are:

1. Run qualification test to ensure the tool meets production quality criteria.
2. Verify the wafer properties using a lot control system.
3. Download the process recipe with the desired diffusion parameters.
4. Set up the furnace, including a temperature profile.
5. Clean the wafers and dip them in hydrofluoric (HF) to remove native oxide.
6. Perform predeposition: load wafers into the deposition furnace and diffuse the dopant.
7. Perform drive-in: increase the furnace temperature to drive-in and activate the dopant bonds, then unload wafers.
8. Measure, evaluate, and record the junction depth and sheet resistivity.

The first six steps are done to prepare wafers for diffusion in the drive-in operation of step 7. A qualification test (or "qual") is performed on the furnace. This test determines the tool's production worthiness in terms of particles and other specified criteria. Once the tool is proven production worthy, the appropriate wafer lots are identified and the proper process recipe is downloaded to the furnace. A thermal profile should always be done to verify furnace temperatures are correct for the process to be performed. In a production environment, a given furnace is dedicated to a specific process step and is maintained at the precise temperature profile.

Wafer Cleaning ■ Wafer cleaning is critical because contaminants can block the diffusion of the dopant atoms into the wafer. Wafers should be cleaned immediately prior to loading the wafers into the furnace to minimize wafer contamination (e.g., native oxide). Cleaning typically includes an immersion in an acid and oxidizer, followed by an etch in an HF solution to remove any remaining oxides (see Chapter 6 for wet cleaning). The use of a dilute HF solution is referred to as a *deglaze* step to remove the poor quality native oxide on the wafer surface.

Dopant Sources ■ Although they may have been used in the early days of the semiconductor industry, it is impractical to use pure elements as a dopant source for submicron IC manufacturing. For instance, boron and phosphorus are solids at room temperature with low vapor pressures, making them difficult to melt or vaporize. Dopant concentration is also difficult to control when using a solid dopant source. Dopants are typically supplied as a gaseous or liquid source in the form of a compound. Some of the most common dopants are listed in Table 17.4 on page 482.

A liquid source (bubbler) has a carrier gas (e.g., nitrogen) bubbled through it and is delivered to the furnace tube as a vapor. Oxygen is also supplied for the dopant source to react and form oxides. Using boron as an example, the boric oxide undergoes a second reaction with silicon to form a boron-rich silicon dioxide layer on top of the wafer. This layer serves as the local source of boron for the predeposition step. The following two reaction equations illustrate how the diborane source is converted into a boron dopant:[5]

1st Reaction:	B_2H_6 (gas) *diborane*	+	$3O_2$ (gas) *oxygen*	→	B_2O_3 (solid) *boric oxide*	+	$3H_2O$ (liquid) *water*
2nd Reaction:	$2B_2O_3$ (solid) *boric oxide*	+	3Si (solid) *silicon*	→	4B (solid) *boron (dopant)*	+	$3SiO_2$ (solid) *silicon dioxide*

TABLE 17.4* Typical Dopant Sources for Diffusion

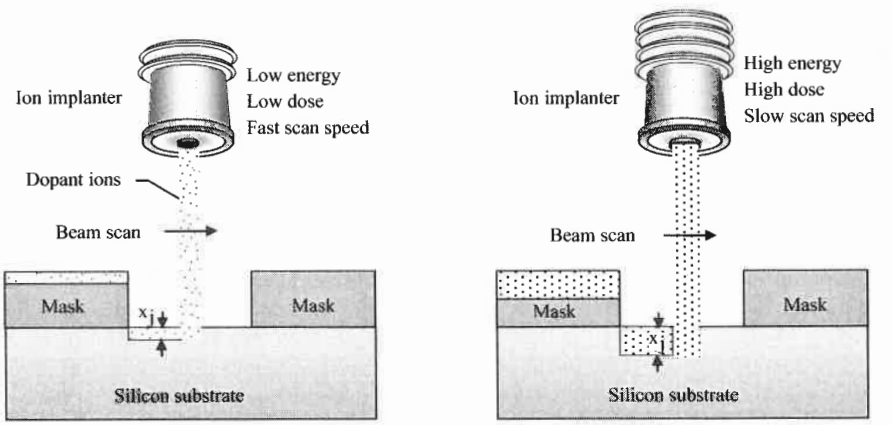
Dopant	Formula of Source	Chemical Name
Arsenic (As)	AsH_3	Arsine (gas)
Phosphorus (P)	PH_3	Phosphine (gas)
Phosphorus (P)	$POCl_3$	Phosphorus oxychloride (liquid)
Boron (B)	B_2H_6	Diborane (gas)
Boron (B)	BF_3	Boron tri-fluoride (gas)
Boron (B)	BBr_3	Boron tri-bromide (liquid)
Antimony (Sb)	$SbCl_5$	Antimony pentachloride (solid)

*SEMATECH "Diffusion Processes," *Furnace Processes and Related Topics,* (Austin, TX: SEMATECH, 1994), p. 7.

Gaseous sources are delivered to the furnace directly from the gas cylinder through a mass flow controller. The dopant is diluted in the cylinder with an inert gas to help prevent system corrosion and control the flow of the gas. A point-of-use filter should always be used to prevent fine particulates from getting into the furnace. Some diffusion sources are highly toxic, especially the gaseous sources. The most toxic of the commonly used dopant sources are arsine (AsH_3) and diborane (B_2H_6). Proper storage and usage is required at all times, especially to avoid a gas leak.

ION IMPLANTATION

Ion implantation is a method for introducing controlled amounts of dopants into the silicon substrate to change its electronic properties. It is a physical process—there are no chemical reactions. A number of applications use ion implantation in modern wafer fabrication (see Table 17.2 on page 476). The main application of ion implantation is the doping of semiconductor materials. Each doping application has specific requirements for the concentration of the dopant material and its depth. Ion implantation is preferred over diffusion in nearly all applications because of its ability to repeatedly control dopant concentration and depth. It has become a critical process to meet the challenges of wafer fabrication with sub-quarter micron feature sizes and larger diameter wafers (see Figure 17.5).



(a) Low dopant concentration (n^-, p^-) and shallow junction (x_j)

(b) High dopant concentration (n^+, p^+) and deep junction (x_j)

FIGURE 17.5 Controlling Dopant Concentration and Depth

until reaching high vacuum. At this point, an isolation valve opens and a robotic mechanism moves the cassette into the main target chamber. When the cassette is in the process chamber, a robotic wafer handler system moves the wafers from the cassette and places them on the scanning disk. The wafer usually undergoes initial alignments using the wafer flat or notch to ensure proper wafer orientation in the disk.

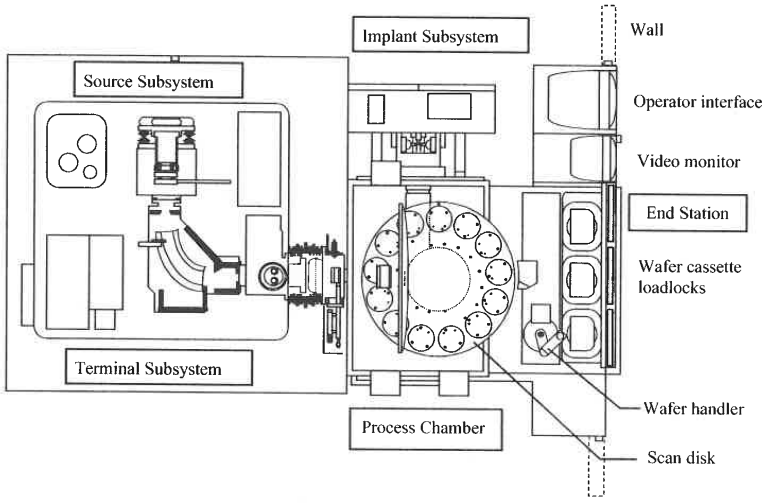


FIGURE 17.25 Wafer Handler for an Implant Process Chamber
(Used with permission from Varian Semiconductor Equipment, VIISion 200 Ion Implanter)

All subsystems are accessed by the technician through a central computer system. System status and information are obtained through the computer, including diagnostics and maintenance information.

Dose Control ■ In ion implanters, real-time dose control is done by measuring the ion beam impinging upon the wafers. The beam current is measured using a sensor called the *Faraday cup.* In a simple Faraday system, a current sensing target is placed in the ion beam path to measure the beam current. However, this placement poses a problem because the ion beam interacts with the target and creates secondary electrons, which give a false beam current reading. The Faraday system uses a design that suppresses the secondary electrons using either an electrical bias (see Figure 17.26 on page 500) or magnetic field to obtain a true ion beam current reading. The measured beam current from the Faraday cup is input into the electronic *dose controller,* which functions as a current integrator (it continuously sums the measured ion beam current). The dose controller correlates the summed beam current to the elapsed implant time and calculates the total time needed to reach a set dose value.

Annealing

Ion implantation damages the silicon lattice by knocking atoms out of the lattice structure. With high doses, the implanted layer actually becomes amorphous. Furthermore, the implanted ions rarely enter the lattice structure sites of the silicon, instead stopping in interstitial sites outside the lattice. These interstitial dopants are electrically inactive until activated by a high-temperature *annealing* step. Annealing heats the implanted silicon substrate to repair crystalline damage and electrically activate dopants by moving the atoms into crystal lattice sites (see Figure 17.27 on page 500). Crystal damage repair is done at about 500°C and dopant activation occurs at about 950°C. Electrical activation of dopants occurs as a function of time and temperature, with longer

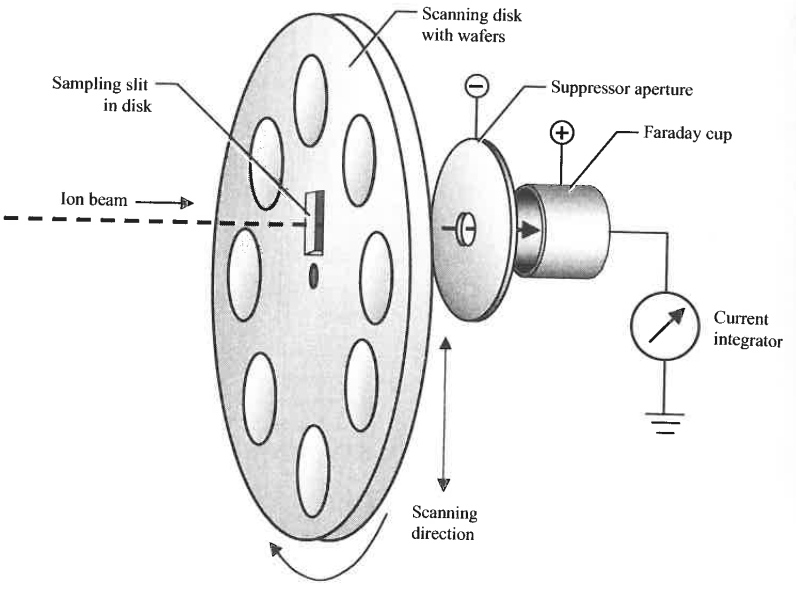
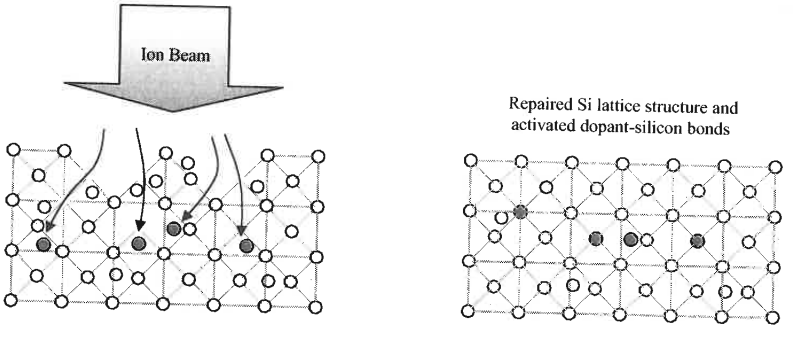


FIGURE 17.26 Faraday Cup Beam Current Measurement
Redrawn from S. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, 2nd ed., (New York: Wiley, 1994), p. 417.



(a) Damaged Si lattice during implant

(b) Si lattice after annealing

FIGURE 17.27 Annealing of Silicon Crystal

times and higher temperatures increasing the dopant activation. There are two basic methods for heating the implanted wafer for anneal:

- Furnace anneal
- Rapid thermal anneal (RTA)

Furnace Anneal ■ *Furnace anneal* is the traditional method of annealing by heating the wafer in a hot-wall furnace in the range of 800 to 1000°C for approximately 30 minutes. At this temperature, the silicon atoms move back into the lattice sites and the dopant atoms enter substitutional sites in the lattice. However, annealing cycles at this time and temperature can cause extensive dopant diffusion and is undesirable for advanced IC wafer fabrication.

Rapid Thermal Anneal (RTA) ■ *Rapid thermal anneal (RTA)* anneals the wafer by using an extremely fast ramp and short dwell time at the target temperature (typically 1000°C). Annealing of

implanted wafers is usually done in a single-wafer rapid thermal processor (RTP) with argon (Ar) or nitrogen (N_2) flowing into the chamber. A fast ramp rate and short dwell time is optimized to anneal the wafer to restore lattice damage and electrically activate the dopant while minimizing dopant diffusion in the silicon. RTA also minimizes a phenomenon known as transient enhanced diffusion (see the following section). RTA is the optimal anneal method to achieve acceptable junction depth control in shallow implants.

Investigations have been done into spike anneals for RTA. The ramp-up rate is very fast at up to 150°C/sec. This speed permits a soak time reduction to about 1 second. Uniform ramp rates during a spike anneal are critical because much of the thermal exposure occurs during the ramp.[22]

The RTA equipment design is similar to the rapid thermal processor previously discussed in Chapter 10. There are different types of heating, with a common method being tungsten halogen lamps positioned on both sides of the wafer. Multiple optical pyrometers are often used to measure temperature across the wafer to ensure uniformity.

Transient Enhanced Diffusion. There is a recent discovery of a dopant diffusion referred to as *transient enhanced diffusion* (*TED*) that occurs in the anneal step after implantation. It is caused by extra interstitial atoms from the dopant ion implanted in the silicon and is not directly related to silicon damage. It has significance when forming ultrashallow junctions because of the need to minimize all dopant diffusion.

Channeling

Single-crystal silicon atoms exhibit long range order, lying in rows and planes throughout the crystal. *Channeling* occurs when the implanted ions are not slowed by collisions with silicon atoms and instead pass through the interstitial areas of the crystal (see Figure 17.28). The undesirable aspect of channeling is that the expected projected range has a very large spread, especially for shallow implants at low ion beam energies. Dopant ions passing through lattice channels do not encounter electronic or nuclear stopping (i.e., have little energy loss) and therefore penetrate deeper than those ions that are stopped by the lattice atoms. There are four ways that channeling is controlled during implant: (1) wafer tilt, (2) screen oxide layer, (3) preamorphization of the silicon, and (4) using dopants with greater amu's (atomic mass units).

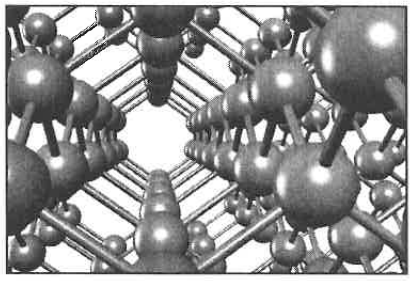


FIGURE 17.28 Silicon Lattice Viewed Along <110> Axis
(Used with permission from Edgard Torres Designs)

Wafer Tilt. The most widely used method to reduce channeling during ion implantation is *wafer tilt.* This method orients the wafer at an angle relative to the incident ion beam. A common angle for (100) wafers is 7° from the perpendicular and ensures that a collision occurs within a short distance of the dopant ion entering the silicon. The tilt angle is set during the scanning process. In this manner, the crystal lattice presents a dense orientation to the incident beam to obtain better control of the projected range on the implanted ion (see Figure 17.29 on page 502). Note that channeling behavior is different at low energies (<1 keV) for ultrashallow junctions, where tilt has little effect on reducing channeling.[23] Furthermore, tilt angles increase shadowing effects and may lead to asymmetric device behavior. It is also necessary sometimes to twist (or rotate in the vertical plane) the wafer surface. For the <110> crystal orientation, twisting the wafer 15° to 35° can help reduce channeling.