MODULE-5a. Integrated Circuits

SYLLABUS

Integrated Circuits Background, Evolution of ICs, CMOS Process Integration, Integration of Other Circuit Elements. (Text 1: 9.1, 9.2, 9.3.1, 9.3.3).

Text Books: 1. Ben. G. Streetman, Sanjay Kumar Banergee, "Solid State Electronic Devices", 7thEdition, Pearson Education, 2016, ISBN 978-93-325-5508-2.

Contents

1. Background

- It is important to realize the reasons, both technical and economic, for the dramatic rise of ICs to their present role in electronics.
- several main types of ICs and point out some of the applications of each. More specific fabrication techniques will be presented in later sections.

1.1. Advantages of Integration

- Small Size
- Low Cost
- Improved Performance
- High Reliability And Ruggedness
- Low Power Consumption
- Less Affected To Parameter
- Easy Troubleshooting
- Increased Operating Speed
- Less Weight, Volume
- Easy Replacement

- . Integrated circuits are used in virtually all electronic equipment today and have revolutionized the world of electronics.
- Computers, cell phones, and other digital appliances are now inextricable parts of the structure of modern societies, made possible by the low cost of production of integrated circuits.
- There are two main advantages of ICs over discrete circuits: cost and performance. Cost is low because the chips, with all their components, are printed as a unit by photolithography rather than being constructed one transistor at a time.
- Furthermore, much less material is used to construct a packaged IC die than to construct a discrete circuit.
- Performance is high because the components switch quickly and consume little power (compared to their discrete counterparts) as a result of the small size and close proximity of the components.

1.2. *Types of Integrated Circuits*

Integrated Circuitis a miniature, low cost electronic circuit consisting of active and passive components that are irreparably joined together on a single crystal chip of silicon.

1.2.1. Based on fabrication

a. Monolithic IC's

b. Hybrid IC's

a. **Monolithic IC's** : In monolithic ICs all components (active and passive) are formed simultaneously by a diffusion process. Then a metallization process is used in interconnecting these components to form the desired circuit.

b. **Hybrid IC's**: .In hybrid ICs, passive components (such as resistors and capacitors) and the interconnections between them are formed on an insulating substrate. The substrate is used as a chassis for the integrated components. Active components such as transistors and diodes as wellas monolithic integrated circuits, are then connected to form a complete circuit.

1.2.2. Based on number of components integrated on IC's

a. SSI <10 components

b. MSI <100 components

c. LSI >100 components

d. VLSI >1000 components

- **Small Scale Integration or (SSI)** - Contain up to 10 transistors or a few gates within a single package such as AND, OR, NOT gates.

- **Medium Scale Integration or (MSI)** - between 10 and 100 transistors or tens of gates within a single package and perform digital operations such as adders, decoders, counters, flip-flops and multiplexers.

- **Large Scale Integration or (LSI)** - between 100 and 1,000 transistors or hundreds of gates and perform specific digital operations such as I/O chips, memory, arithmetic and logic units.

- **Very-Large Scale Integration or (VLSI)** - between 1,000 and 10,000 transistors or thousands of gates and perform computational operations such as processors, large memory arrays and programmable logic devices.

- **Ultra-Large Scale Integration or (ULSI)** - more than 1 million transistors - the big boys that are used in computers CPUs, GPUs, video processors, micro-controllers, FPGAs and complex PICs.

1.2.3. Classification of integrated circuits based on the Application:

- **DIGITAL INTEGRATED CIRCUITS**

Digital integrated circuits, primarily used to build computer systems, also occur in cellular phones, stereos and televisions. Digital integrated circuits include microprocessors, microcontrollers and logic circuits. They perform mathematical calculations, direct the flow of data and make decisions based on Boolean logic principles.The Boolean system used centers on on two numbers: 0 and 1.

- **ANALOG INTEGRATED CIRCUITS**

Analog integrated circuits most commonly make up a part of power supplies, instruments and communications. In these applications, analog integrated circuits amplify, filter and modify electrical signals. In cellular phones, they amplify and filter the incoming signal from the phone's antenna. The sound encoded into that signal has a low amplitude level; after the circuit filters the sound signal from the incoming signal, the circuit amplifies the sound signal and sends it to the speaker in your cell phone, allowing you to hear the voice on the other end.

- **MIXED-SIGNAL INTEGRATED CIRCUITS**

Mixed-signal circuits occur in cellular phones, instrumentation, motor and industrial control applications. These circuits convert digital signals to analog signals, which in turn set the speed of motors, the brightness of lights and the temperature of heaters, for example. They also convert digital signals to

sound waveforms, allowing for the design of digital musical instruments such as electronic organs and computer keyboards capable of playing music.

- **MEMORY-INTEGRATED CIRCUITS**

Though primarily used in computer systems, memory-integrated circuits also occur in cellular phones, stereos and televisions. A computer system may include 20 to 40 memory chips, while other types of electronic systems may contain just a few.

Memory circuits store information, or data, as two numbers: 0 and 1. Digital integrated circuits will often retrieve these numbers from memory and perform calculations with them, then save the calculation result a memory chip's data storage locations. The more data it accesses---pictures, sound and text---the more memory an electronic system will require.

1.3. Evolution of Integrated Circuits

- The IC was invented in February 1959 by Jack Kilby of Texas Instruments.
- The planar version of the IC was developed independently by Robert Noyce at Fairchild in July 1959.
- One way to gauge the progress of the field is to look at the complexity of ICs as a function of time. FIG. 2 shows the number of transistors used in MOS microprocessor IC chips as a function of time.
- It is amazing that on this semilog plot, where we have plotted the log of the component count as a function of time, we get a straight line over four decades, indicating that there has been an exponential growth in the complexity of chips.

Evolution of IC with year tabulated below.

Invention of transistor	1948
Development of silicon transistor	1955-
Silicon planar technology	1959
First IC,Small Scale Integration(SSI),3 to 30	1960
Medium Scale Integration (MSI),30 to 300	1965-
Large Scale Integration(LSI),300 to 3000 gates/chip	1970-
Very Large Scale Integration(VLSI) more than 3000	1975

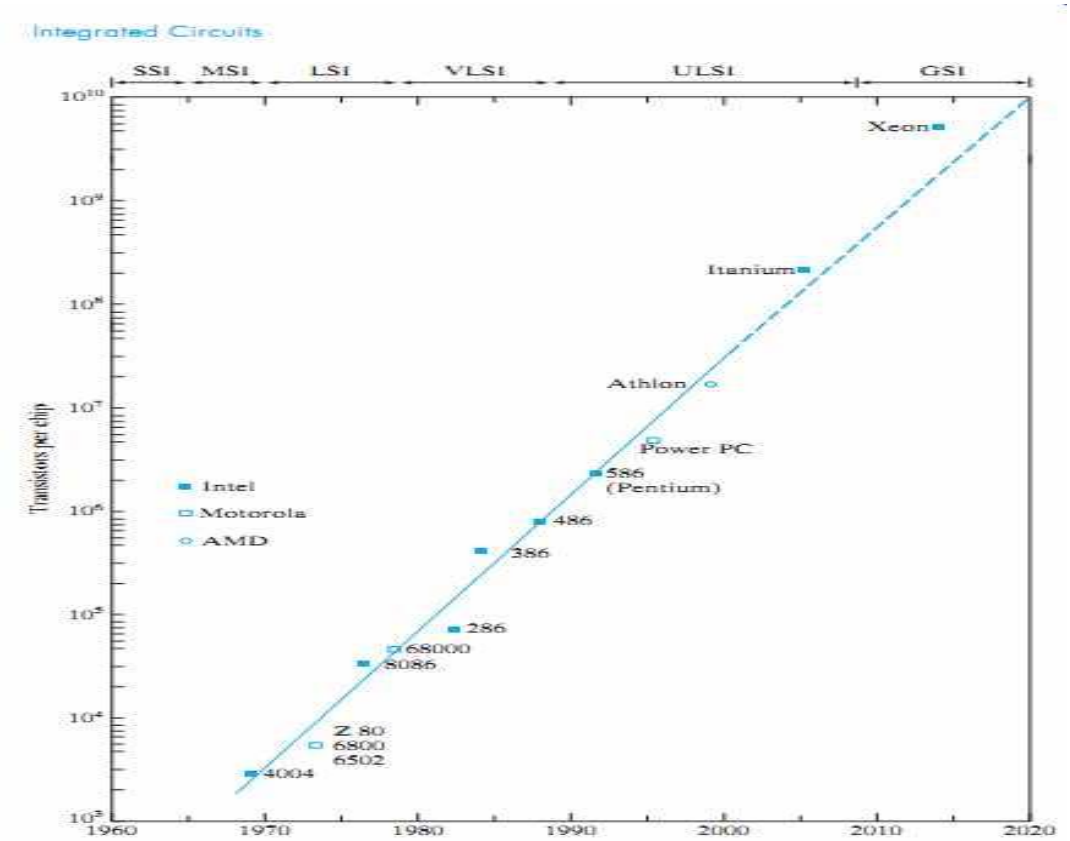


FIG. 2

Moore's law for integrated circuits:

- The component count has roughly doubled every 18 months, as was noted early on by Gordon Moore of Intel corporation. This regular doubling has become known as **Moore's law.**
- Exponential increase in transistor count as a function of time for different generations of microprocessors.
- The dashed line indicates projections based on the International Technology Roadmap for Semiconductors (ITRS). Notice that the transistor count in the future may not increase at the same rate as in the past, due to practical constraints such as economics and power dissipation.

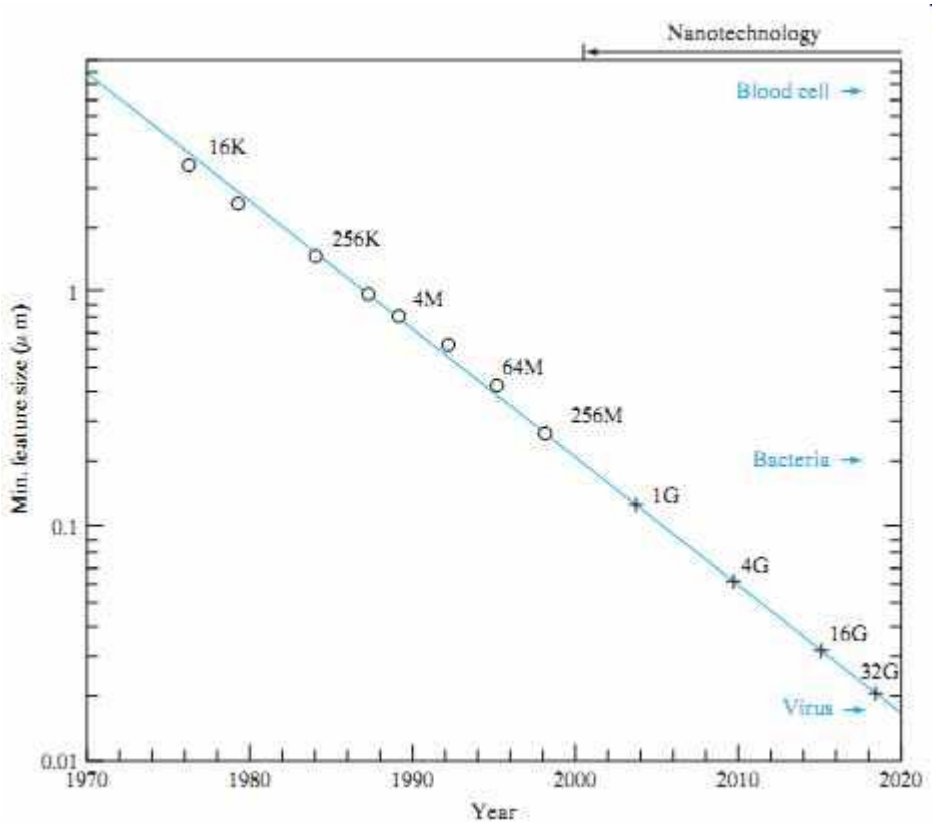


FIG. 3 Exponential decrease in typical feature size with time for different generations of dynamic random-access memories (16-kb to 32-Gb DRAMs). For reference, sizes of blood cells, bacteria, and viruses are shown on the um scale. Dimensions below 100 nm are considered to be in the realm of nanotechnology.

2. CMOS INVERTER

- A particularly useful device for digital applications is a combination of n-channel and p-channel MOS transistors on adjacent regions of the chip.
- This complementary MOS (commonly called CMOS) combination is illustrated in the basic inverter circuit of FIG. 4a.
- In this circuit the drains of the two transistors are connected together and form the output, while the input terminal is the common connection to the transistor gates.
- The p-channel device has a negative threshold voltage, and the n-channel transistor has a positive threshold voltage.
- Therefore, a zero voltage input (Vin = 0) gives zero gate voltage for the n-channel device, but the voltage between the gate and source of the p-channel device is -VDD.

- Thus the p-channel device is on, the n-channel device is off, and the full voltage VDD is measured at Vout (i.e., VDD appears across the nonconducting n-channel transistor).
- Alternatively, a positive value of Vin turns the n-channel transistor on, and the p-channel off. The output voltage measured across the "on" n-channel device is essentially zero.
- Thus, the circuit operates as an inverter-with a binary "1" at the input, the output is in the "0" state, whereas a "0" input produces a "1" output. The beauty of this circuit is that one of the devices is turned off for either condition.
- Since the devices are connected in series, no drain current flows, except for a small charging current during the switching process from one state to the other.

2.1. Applications:

- the CMOS inverter uses ultra little power, it is particularly useful in applications such as electronic watch circuits that depend on very low power consumption.
- CMOS is also advantageous in ultra large-scale integrated circuits since even small power dissipation in each transistor becomes a problem when millions of them are integrated on a chip.

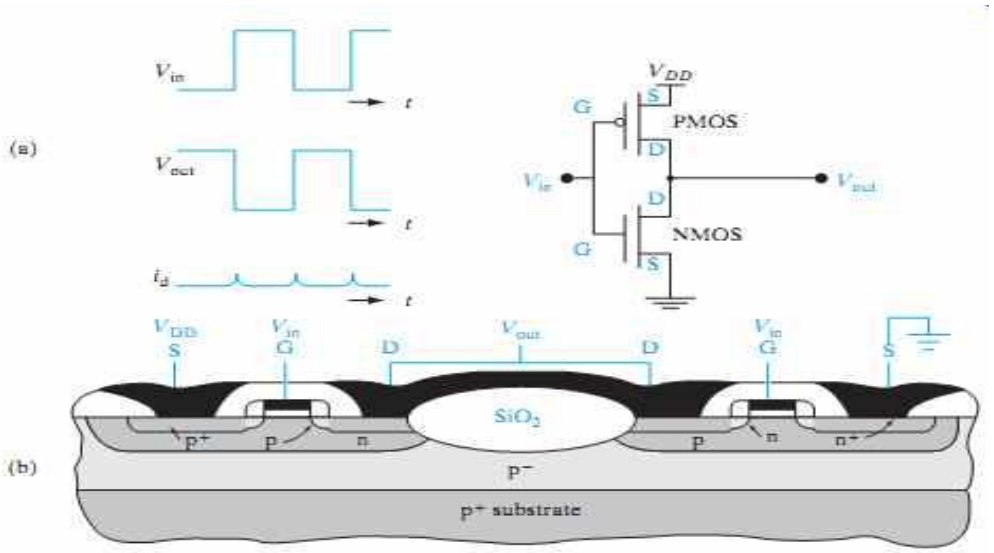


FIG. 4 Complementary MOS structure: (a) CMOS inverter; (b) formation of p-channel and n-channel devices together.

3. IC Fabrication Steps (these is for those who are not studying for only exam)

The basic IC fabrication steps will be described in the following sections. Some of these steps may be carried out many times, in different combinations and/or processing conditions during a complete fabrication run.

3.1. *Silicon Wafers*

- **Polysilicon** is produced from metallurgical grade silicon by a chemical purification process, called the Siemens process. This process involves distillation of volatile silicon compounds, and their decomposition into silicon at high temperatures. An emerging, alternative process of refinement uses a fluidized bed reactor
- **Metallurgical-grade silicon (MSG or MG-Si)** is silicon of relatively high purity in the order of 98% or higher which is used extensively in the metallurgical industry. MGS is not considered pure enough to be used for electronics and must be further purified into either extremely pure electronic-grade silicon which can be used for integrated circuit fabrication or slightly purer upgraded metallurgical-grade silicon which can be used in cheaper electronic devices such as solar cells and liquid crystal displays.
- **Electronic-grade silicon (EGS or EG-Si)** or semiconductor-grade silicon (SGS) is a highly-purified version of the metallurgical-grade silicon with extremely low impurities suitable for microelectronic device applications. Electronic-grade silicon is the raw material used for the growth of single-crystal silicon in the manufacturing of silicon wafers. From MGS to EGS, the impurities go from the order of parts per million to the low parts per billion.
- **Silicon Ingot Production: The Czochralski-technique** is a method to pull a monocrystal with the same crystallographic orientation of a small monocrystalline seed crystal out of melted silicon. electronic-grade polysilicon nuggets (e. g. from the Siemens-process) optionally together with dopants are melted in a quartz crucible at a temperature > 1400°C in an inert gas atmosphere (e. g. argon). The quartz crucible sits inside a graphite container which – due to its high heat conductivity – homogeneously transfers the heat from the surrounding heater to the quartz crucible.
- **From the Ingot to Finished Silicon Wafers (Grinding):** The ingots grown with the Czochralski technique are ground to the desired diameter and cut into shorter workable cylinders with e. g. a band saw

and ground to a certain diameter. An orientation flat is added to indicate the crystal orientation (schema right), while wafers with an 8 inch diameter and above typically use a single notch to convey wafer orientation, independent from the doping type.

3.2. Layering

3.2.1. Oxidation

- In **oxidation**, silicon reacts with oxygen to form silicon dioxide (SiO2). To speed up this chemical reaction, it is necessary to carry out the oxidation at high temperatures (e.g., 1000–1200°C) and inside ultraclean furnaces.
- To avoid the introduction of even small quantities of contaminants (which could significantly alter the electrical properties of the silicon), it is necessary to operate in a **clean room**.
- Particle filters are used to ensure that the airflow in the processing area is free from dust. All personnel must protect the clean-room environment by wearing special lint-free clothing that covers a person from head to toe.
- The oxygen used in the reaction can be introduced either as a high-purity gas (referred to as a "**dry oxidation**") or as steam (forming a "**wet oxidation**").
- In general, wet oxidation has a faster growth rate, but dry oxidation gives better electrical characteristics.
- The thermally grown oxide layer has excellent electrical insulation properties. can be used to form excellent MOS capacitors.
- Silicon dioxide can also serve as an effective mask against many impurities, allowing the introduction of dopants into the silicon only in regions that are not covered with oxide.
- Silicon dioxide is a transparent film, and the silicon surface is highly reflective. If white light is shone on an oxidized wafer, constructive and destructive interference will cause certain colors to be reflected.
- The wavelengths of the reflected light depend on the thickness of the oxide layer. In fact, by categorizing the color of the wafer surface, one can deduce the thickness of the oxide layer.

3.2.2. Nitridation

- The oxide does not grow in the regions that are protected by silicon nitride, because nitride blocks the diffusion of oxygen and water molecules (and thereby prevents oxidation of the Si substrate).
- The pad oxide that is used under the nitride has two roles: It minimizes the thermal-expansion mismatch and concomitant stress between silicon nitride and the substrate;
- it also prevents chemical bonding of the silicon nitride to the silicon substrate.

3.2.3. Chemical Vapor Deposition

- **Chemical vapor deposition** (CVD) is a process by which gases or vapors are chemically reacted, leading to the formation of solids on a substrate.
- CVD can be used to deposit various materials on a silicon substrate including SiO2, Si3N4, polysilicon, and so on.
- The properties of the CVD oxide layer are not as good as those of a thermally grown oxide, but they are sufficient to allow the layer to act as an electrical insulator.
- The advantage of a CVD layer is that the oxide deposits at a faster rate and a lower temperature (below 500°C).
- If silane gas alone is used, then a silicon layer will be deposited on the wafer. If the reaction temperature is high enough (above 1000°C), the layer deposited will be a crystalline layer (assuming that there is an exposed crystalline silicon substrate). Such a layer is called an **epitaxial** layer
- the deposition process is referred to as **epitaxy** instead of CVD. At lower temperatures, or if the substrate surface is not single-crystal silicon, the atoms will not be able to aligned along the same crystalline direction. Such a layer is called **poly crystallin silicon (poly Si),** since it consists of many small crystals of silicon aligned in random fashion.

3.2.4. Metallization

- The purpose of metallization is to interconnect the various components (transistors, capacitors, etc.) to form the desired integrated circuit.
- Metallization involves the deposition of a metal over the entire surface of the silicon. The required interconnection pattern is then selectively etched.

- The metal layer is normally deposited via a sputtering process. A pure metal disk (e.g., 99.99% aluminum target) is placed under an Ar (argon) ion gun inside a vacuum chamber.
- The wafers are also mounted inside the chamber above the target. The Ar ions will not react with the metal, since argon is a noble gas. However, the ions are made to physically bombard the target and literally knock metal atoms out of the target.
- These metal atoms will then coat all the surface inside the chamber, including the wafers.
- The thickness of the metal film can be controlled by the length of the sputtering time, which is normally in the range of 1 to 2 minutes. The metal interconnects can then be defined using photolithography and etching steps.

3.4. Patterning

3.4.1. Photolithography

- The surface patterns of the various integrated-circuit components can be defined repeatedly using photolithography. The sequence of photolithographic steps is as illustrated in Fig. A.1.
- The wafer surface is coated with a photosensitive layer called photoresist, using a spin-on technique.
- After this, a photographic plate with drawn patterns (e.g., a quartz plate with chromium layer for patterning) will be used to selectively expose the photoresist under a deep ultra-violet illumination (UV).
- The exposed areas will become softened (for positive photoresist). The exposed layer can then be removed using a chemical developer, causing the mask pattern to be duplicated on the wafer.
- Very fine surface geometries can be reproduced accurately by this technique. Furthermore, the patterns can be projected directly onto the wafer, or by using a separate photomask produced by a 10x "step and repeat" reduction technique as shown in Fig. A.2.
- The patterned photoresist layer can be used as an effective masking layer to protect materials below from wet chemical **etching** or **reactive ion etching** (RIE).
- Silicon dioxide, silicon nitride, polysilicon, and metal layers can be selectively removed using the appropriate etching methods.
- After the etching step(s), the photoresist is stripped away, leaving behind a permanent pattern of the photomask on the wafer surface.
- To make this process even more challenging, multiple masking layers (which can number more than 20 in advanced VLSI fabrication

processes) must be aligned precisely on top of Negative photo- resist previous layers.

- This must be done with even finer precision than the minimum geometry size of the masking patterns. This requirement imposes very critical mechanical and optical constraints on the photolithography equipment.

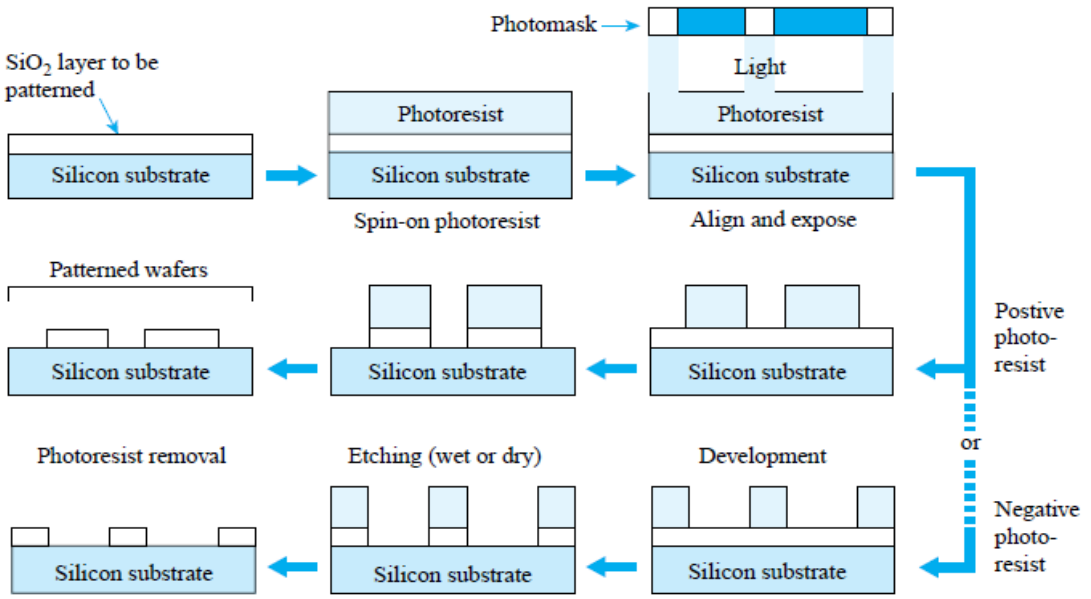


Figure A.1 Photolithography using positive or negative photoresist.

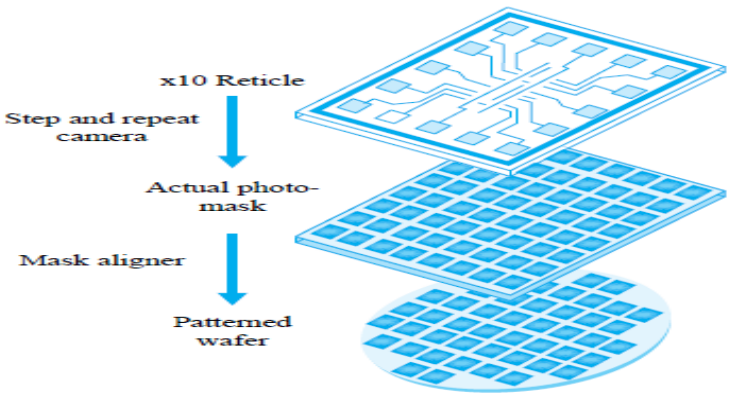


Figure A.2 Conceptual illustration of a step-and-repeat reduction technique to facilitate the mass production of integrated circuits.

3.4.2. Etching

- To permanently imprint the photographic patterns onto the wafer, chemical (**wet**) etching or RIE **dry etching** procedures can be used.
- Chemical etching is usually referred to as **wet etching**.
- Different chemical solutions can be used to remove different layers. For example, hydrofluoric (HF) acid can be used to etch SiO2, potassium hydroxide (KOH) for silicon, phosphoric acid for aluminum, and so on.
- Two most important issues in etching are **selectivity** and **anisotropy**.
- – **Selectivity** refers to the ability of an etchant to remove one material on the surface while leaving another intact.
- – **Isotropic** refers to the tendency of the etching to proceed laterally as well as downward
- The patterned photoresist layer can be used as an effective masking layer to protect materials below from wet chemical **etching** or **reactive ion etching** (RIE).
- Silicon dioxide, silicon nitride, polysilicon, and metal layers can be selectively removed using the appropriate etching methods.
- In wet etching, the chemical usually attacks the exposed regions that are not protected by the photoresist layer in all directions (**isotropic etching**).
- Depending on the thickness of the layer to be etched, a certain amount of undercut will occur. Therefore, the dimension of the actual pattern will differ slightly from the original pattern.
- If exact dimension is critical, RIE **dry etching** can be used.
- This method is essentially a directional bombardment of the exposed surface using a corrosive gas (or ions).
- The cross section of the etched layer is usually highly directional (**anisotropic etching**) and has the same dimension as the photoresist pattern. A comparison between isotropic and anisotropic etching is given in Fig. A.3.

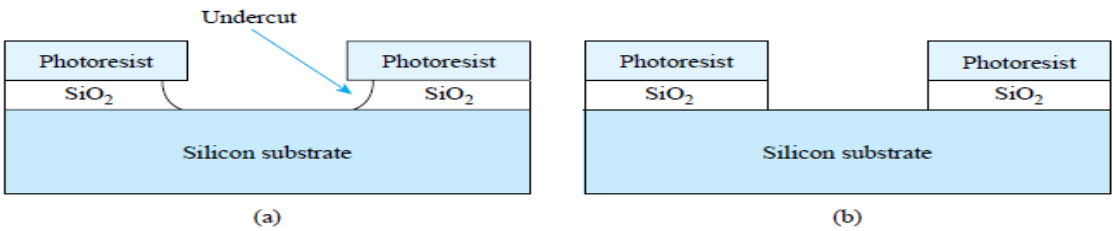


Figure A.3 (**a**) Cross-sectional view of an isotropic oxide etch with severe undercut beneath the photoresist layer. (**b**) Anisotropic etching, which usually produces a cross section with no undercut.

3.5. *Doping*

- These variables are strictly controlled during crystal growth. A specific amount of impurities can be added to the pure silicon in a process known as **doping.**
- This allows the alteration of the electrical properties of the silicon, in particular its resistivity. Depending on the types of impurity, either holes (in ***p*-type** silicon) or electrons (in ***n*-type** silicon) can be responsible for electrical conduction.
- If a large number of impurity atoms is added, the silicon will be heavily doped (e.g., concentration > ~1018 atoms/cm). When designating the relative doping concentrations in semiconductor material, it is common to use the and symbols.
- A heavily doped (low-resistivity) *n*-type silicon wafer is referred to as *n*+ material, while a lightly doped material (e.g., concentration < ~1016 atoms/cm) is referred to as *n*. Similarly, *p*+ and *p* designations refer to the heavily doped and lightly doped *p*-type regions, respectively.
- The ability to control the type of impurities and the doping concentration in the silicon permits the formation of diodes, transistors, and resistors in integrated circuits.

3.5.1. Diffusion

- **Diffusion** is a process by which atoms move from a high-concentration region to a low- concentration region.
- This is very much like a drop of ink dispersing through a glass of water except that it occurs much more slowly in solids.
- In VLSI fabrication, this is a method to introduce impurity atoms (dopants) into silicon to change its resistivity.
- The rate at which dopants diffuse in silicon is a strong function of temperature.
- Diffusion of impurities is usually carried out at high temperatures (1000–1200°C) to obtain the desired doping profile.
- When the wafer is cooled to room temperature, the impurities are essentially "frozen" in position.
- The diffusion process is performed in furnaces similar to those used for oxidation. The depth to which the impurities diffuse depends on both the temperature and the processing time.
- The most common impurities used as dopants are boron, phosphorus, and arsenic. Boron is a p-type dopant, while phosphorus and arsenic are n-type dopants.

3.5.2. Ion Implantation

- **Ion implantation** is another method used to introduce impurities into the semiconductor crystal. An ion implanter produces ions of the desired dopant, accelerates them by an electric field, and allows them to strike the semiconductor surface.
- Ion implantation can be used to form a deep region of doping using a two step procedure
- A high concentration of dopant is deposited near the surface in the **predeposition** or predep stage
- The dopant source is then removed and the wafer heated to cause redistribution of the dopant via diffusion in the **drive-in stage**

3.6. Packaging

- A finished silicon wafer may contain several hundreds of finished circuits or chips. A chip may contain from 10 to more than 108 transistors; each chip is rectangular and can be up to tens of millimeters on a side.
- The circuits are first tested electrically (while still in wafer form) using an automatic probing station. Bad circuits are marked for later identification.
- Fine gold wires are normally used to interconnect the pins of the package to the metallization pat- tern on the die. Finally, the package is sealed using plastic or epoxy under vacuum or in an inert atmosphere.

4. CMOS Process Integration (Self-aligned Twin-Well CMOS Process)

- Depending on the choice of starting material (substrate), CMOS processes can be identified as n-well, p-well, or twin-well processes.
- A modern twin-well CMOS process flow is shown in Fig. A.5. A minimum of 10 masking layers is required.
- In practice, most CMOS processes will also require additional layers such as n- and p-guards for better latchup immunity, a second polysilicon layer for capacitors, and multilayer metals for high-density interconnections.

STEP-1 well Formation (Fig. A.5a).

- The starting material for the twin-well CMOS is a p-type substrate. The process begins with the formation of the p-well and the n-well

- The n-well is required wherever p-channel MOSFETs are to be placed, while the p-well is used to house the n-channel MOS- FETs. The well-formation procedures are similar.
- A thick photoresist layer is etched to expose the regions for n-well diffusion. The unexposed regions will be protected from the n-type phosphorus impurity.
- Phosphorus implantation is usually used for deep diffusions, since it has a large diffusion coefficient and can diffuse faster than arsenic into the substrate.

STEP-2 define the active regions (Fig. A.5b).

- The second step is to define the active regions (region where transistors are to be placed) using a technique called **shallow trench isolation (STI).**
- To reduce the chance of unwanted latchup (a serious issue in CMOS technology), **dry etching** is used to produce trenches approximately 0.3 μm deep on the silicon surface.
- These trenches are then refilled using **CVD oxide**, followed by a **planarization procedure**. This results in a cross section with flat surface topology
- An alternate isolation technique is called **local oxidation of silicon (LOCOS).** This older technology uses **silicon nitride (Si3N4)** patterns to protect the penetration of oxygen during oxidation.
- This allows selective regions of the wafer surface to be oxidized. After a long **wet-oxidation** step, thick field oxide will appear in regions between transistors.
- This effectively produces an effect similar to that obtained in the STI process, but at the expense of large area overhead.

STEP-3 polysilicon gate (Fig. A.5c).

- This is one of the most critical steps in the CMOS process.
- The thin oxide layer in the active region is first removed using wet etching followed by the growth of a high-quality thin gate oxide.
- Current deep-submicron CMOS processes routinely make used of oxide thicknesses as thin as 20 Å to 50 Å (1 angstrom =108 cm).
- A polysilicon layer, usually arsenic doped (n-type), is then deposited and patterned.
- The photolithography is most demanding in this step since the finest resolution is required to pro- duce the shortest possible MOS channel length.

STEP-4 lightly doped drain (LDD) (Fig. A.5d).

The polysilicon gate is a self-aligned structure and is preferred over the older type of metal gate structure.

This is normally accompanied by the formation of lightly doped drain (LDD) regions for MOSFETs of both types to suppress the generation of hot electrons that might affect the reliability of the transistors.

A noncritical mask, together with the polysilicon gates, is used to form the self-aligned LDD regions

STEP-5 n+ diffusion (Fig. A.5e).

- Prior to the n+ and p+ drain region implant, a sidewall spacer step is performed. A thick layer of **silicon nitride is deposited** uniformly on the wafer.
- After a timed **RIE dry etch** to remove all the silicon nitride layer, pockets of silicon nitride will remain at the edge of the polysilicon gate electrode Such pockets of silicon nitride are called **sidewall spacers.**
- They are used to block subsequent n+ or p+ source/drain implants, protecting the **LDD** regions.
- A heavy **arsenic implant** can be used to form the n+ source and drain regions of the n- MOSFETs.
- The polysilicon gate also acts as a barrier for this implant to protect the channel region.

STEP-6 P+ diffusion (Fig. A.5f).

- A layer of photoresist can be used to block the regions where p-MOSFETs are to be formed.
- A reversed photolithography step can be used to protect the n-MOSFETs during the p+ boron source and drain implant for the p-MOS- FETs.
- Note that in both cases the separation between the source and drain diffusions—channel length—is defined by the polysilicon gate mask alone, hence the **self-aligned property**.

STEP-7 Contact holes (Fig. A.5g)

Before contact holes are opened, a thick layer of CVD oxide is deposited over the entire wafer. A photomask is used to define the contact window opening, followed by a wet or dry oxide etch.

STEP-7 Metallization (Fig. A.5h).

A thin aluminum layer is then evaporated or sputtered onto the wafer. A final masking and etching step is used to pattern the interconnection

STEP-8 Nitridation.

Not shown in the process flow is the final passivation step prior to packaging and wire bonding. Finally, the MOSFETs have to be properly interconnected according to the circuit layout, using the metallization level. This involves of an oxide dielectric layer doped with B and P, which is known as **boro-phospho-silicate glass (BPSG)** on the entire wafer, patterning it by means of the contact-level reticle and using RIE to open up the contact holes to the substrate

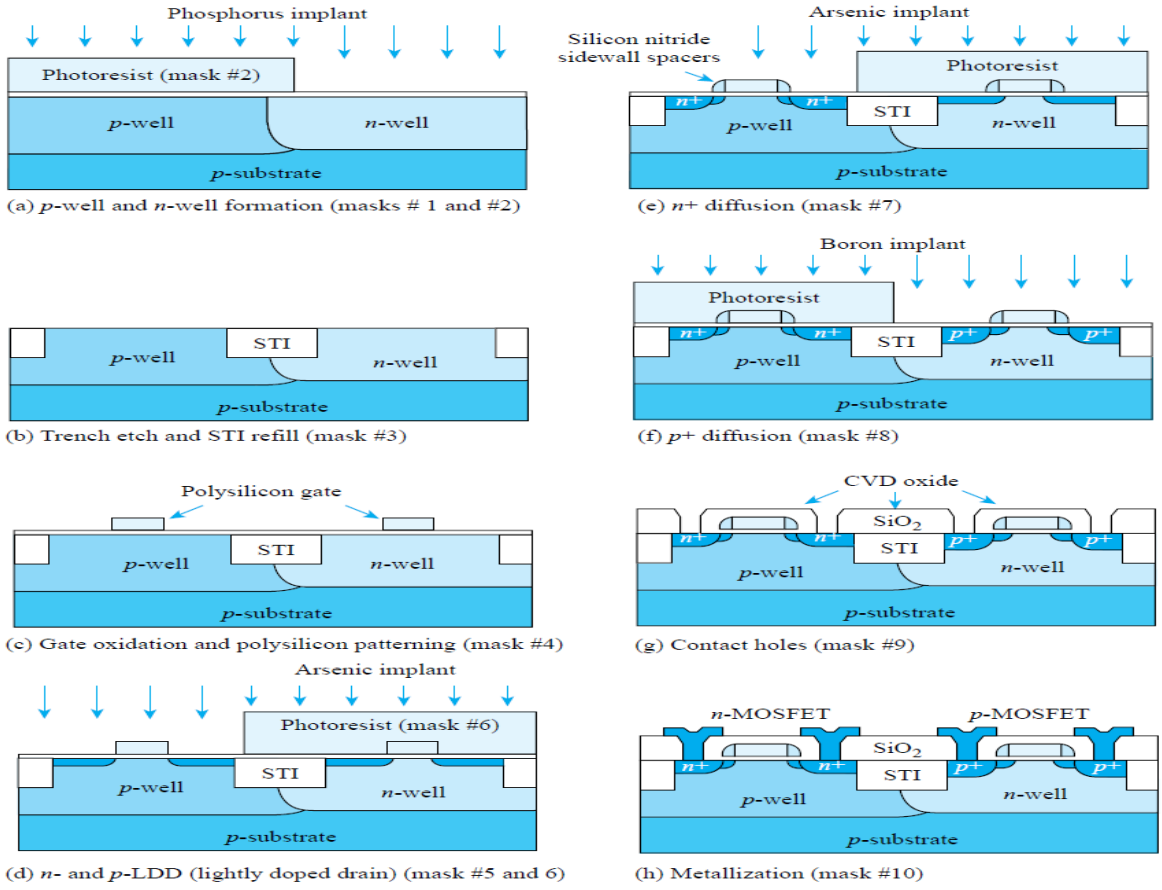


Figure A.5 A modern twin-well CMOS process flow with shallow trench isolation (STI).

5. Integration of Other Circuit elements

- One of the most revolutionary developments of integrated circuit technology is the fact that integrated transistors are cheaper to make than are more mundane elements such as resistors and capacitors.

- There are, however, numerous applications calling for diodes, resistors, capacitors, and inductors in integrated form.

5.1. Diodes.

- It is simple to build p-n junction diodes in a monolithic circuit. It is also common practice to use transistors to perform diode functions.
- Since many transistors are included in a monolithic circuit, no special diffusion step is required to fabricate the diode element.
- There are a number of ways in which a transistor can be connected as a diode.
- Perhaps the most common method is to use the emitter junction as the diode, with the collector and base shorted. This configuration is essentially the narrow base diode structure, which has high switching speed with little charge storage.
- Since all the transistors can be made simultaneously, the proper connections can be included in the metallization pattern to convert some of the transistors into diodes.

5.2. Resistors.

- Diffused or implanted resistors can be obtained in monolithic circuits by using the shallow junctions used in forming the transistor regions (FIG. 11a).
- a p region can be made during the base implant, and an n-type resistor channel can be included within the resulting p region during the emitter implant step.
- In either case, the resistance channel can be isolated from the rest of the circuit by proper biasing of the surrounding material.
- The resistance of the channel depends on its **length, width, depth** of the implant, and resistivity of the implanted material.
- Since the depth and resistivity are determined by the requirements of the base or emitter implant, the variable parameters are the length and width. Two typical resistor geometries are shown in FIG. 11b.
- In each case the resistor is long compared with its width, and a provision is made on each end for making contact to the metallization pattern.

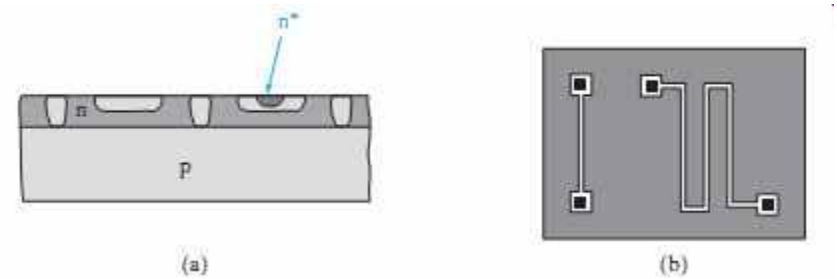


FIG. 11 Monolithic resistors: (a) cross section showing use of base and emitter diffusions for resistors; (b) top view of two resistor patterns.

- Design of diffused resistors begins with a quantity called the **sheet resistance** of the diffused layer.
- If the average resistivity of a diffused region is **ρ**, the resistance of a given length L

  R = ρL/wt,

  where w is the width and t is the thickness of the layer.
- Now if we consider one **square** of the material, such that **L = w**, we have the sheet resistance **Rs~ρ/t** in units of ohms per square.
- We notice that Rs measured for a given layer is numerically the same for any size square.
- The resistance then can be calculated from the known value of Rs and the ratio L/w (the aspect ratio) for the resistor.
- We can make the width **w as small** as possible within the requirements of **heat dissipation and photolithographic** limitations and then calculate the required length from w and Rs.
- Design criteria for diffused resistors include geometrical factors, such as the presence of high current density at the inside corner of a sharp turn. In some cases it is necessary to round corners slightly in a **folded or zigzag** resistor (FIG. 12b) to reduce this problem..
- different implant, such as the VT adjust implant, to form shallow regions having very high sheet resistance (~ 10^5 ohm /square).
- This procedure can provide a considerable saving of space on the chip. In integrated FET circuits it is common to replace load resistors with depletion-mode transistors.

5.3. Capacitors.

- One of the most important elements of an integrated circuit is the capacitor. This is particularly true in the case of memory circuits, where charge is stored in a capacitor for each bit of information.
- FIG. 12 illustrates a one-transistor DRAM cell, in which the n-channel MOS transistor provides access to the adjacent MOS capacitor.

- The top plate of the capacitor is polysilicon, and the bottom plate is an inversion charge contacted by an n+ region of the transistor. The terms bit line and word line refer to the row and column organization of the memory.
- One can also make use of the capacitance associated with p-n junctions.

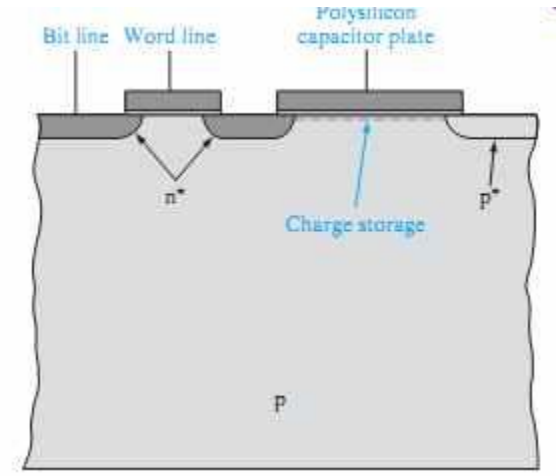


FIG. 12 Integrated capacitor for DRAM cells. A one-transistor memory cell in which the transistor stores and accesses charge in an adjacent planar MOS capacitor.

5.4. Inductors.

- Inductors have not been incorporated into ICs in the past, because it is much harder to integrate inductors than the other circuit elements.
- Also, there has not been a great need for integrating inductors. Recently, that has changed because of the growing need for rf analog ICs for portable communication electronics.
- Inductors are very important for such applications, and can be made with reasonable Q factors using spiral wound thin metal films on an IC.
- Such spiral patterns can be defined by photolithography and etching techniques compatible with IC processing, or they can be incorporated in a hybrid IC.

5.5. Contacts and Interconnections.

5.5.1. metallization

- During the **metallization step**, the various regions of each circuit element are contacted and proper interconnection of the circuit elements is made.

- Aluminum is commonly used for the top metallization, since it adheres well to Si and to SiO2 if the temperature is raised briefly to about 550°C after deposition.
- Gold is used on GaAs devices, but the adhesion properties of Au to Si and SiO2 are poor. Gold also creates deep traps in Si.
- silicide contacts and doped polysilicon conductors are commonly used in integrated circuits..
- In cases where Al is used to contact the Si surface, it is usually necessary to use Al containing about 1% Si to prevent the metal from incorporating Si from the layer being contacted, thereby causing "**spikes**" in the surface.
- Thin diffusion barriers are also used between the Al and Si layers, to prevent migration between the two.

5.5.2.multilayer metallization.

- Increased complexity and packing density in integrated circuits inevitably leads to a need for **multilayer metallization**.
- Multiple levels of Cu metallization can be incorporated with interspersing dielectrics. In general, the metals may all be Al, Cu or they may be different conductors such as poly silicon or refractory metals (depending on the heat each is subjected to in subsequent processing).
- the dielectrics may be deposited oxides, **boro-phospho-silicate glass** for reflow planarization, nitrides, and so on.
- The **planarizatio**n of the surface is extremely important to prevent breaks in the metallization, which can occur in traversing a step on the surface.
- Various approaches using reflow glass, polyimide, and other materials to achieve planarization have been used, along with CMP.

5.5.3. RC Time Constant

- The most important challenge in designing interconnects is the RC time constant, which affects the **speed** and active **power dissipation** of the chip.
- A very simplistic model of **two layers** of interconnects with an inter-metal dielectric (Figs. 14) shows that it can be regarded as a **parallel plate capacitor.**
- Regarding the interconnect as a **rectangular resistor**, its resistance is given by

$$R = \frac{\rho L}{tw} = R_s \frac{L}{w} \qquad (9\text{–}1a)$$

where R_s is the sheet resistance, and the other symbols are defined in Fig. 9–14. The capacitance is given by

$$C = \frac{\epsilon L w}{d} \qquad (9\text{–}1b)$$

The RC time constant is then

$$\left(\frac{\rho L}{wt}\right)\left(\frac{\epsilon L w}{d}\right) = \frac{\rho \epsilon L^2}{td} = \frac{R_s \epsilon L^2}{d} \qquad (9\text{–}2)$$

-
- Therefore, it does not make sense to use wider conductors for **high-speed** operation. It is also impractical to do so in terms of **packing density**.
- From Eq. (eqn. 2), it is clear we need as **thick a metal layer** (within practical limits of deposition times and etching times) and as **low a resistivity** as possible.

5.5.4. Used Material Aluminum

a).Aluminum: is very good in this regard, and thus was a mainstay for Si technology for many years. Aluminum also has other nice attributes such as good **ohmic contacts** to both n and p-type Si, and good **adhesion to oxides**.

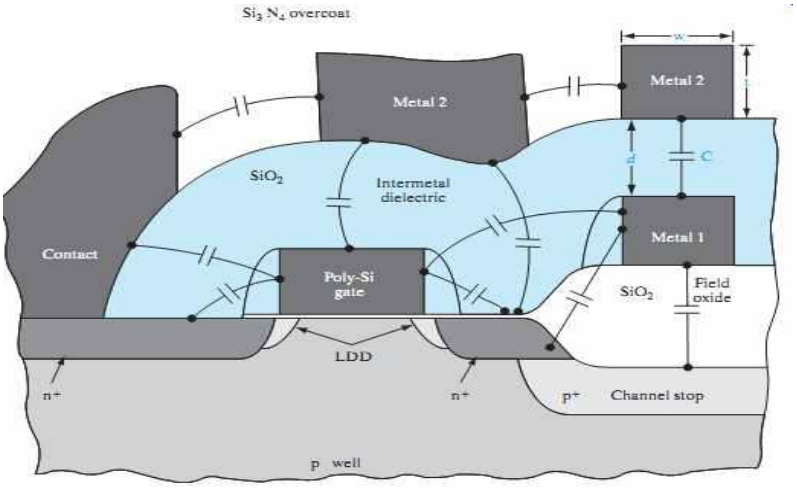


FIG. 14 Equivalent circuit illustrating the various parasitic capacitive elements associated with a multi level interconnect. On the top right hand corner of the figure, we focus on the parallel plate capacitor model referred to in Eqs. (1) and (2).

5.5.5. Used Material Copper

Copper: has even lower resistivity (1.7 u-Ohm-cm) than Al (3 u-Ohm-cm) and is about two orders of magnitude less susceptible to electromigration.

- The process breakthroughs that have made **Cu** viable for metallization include new **electrodeposition and electroplating** techniques because **CVD or sputter deposition** is not very practical for Cu.
- It is also **very difficult** to use **RIE** for Cu because the etch byproducts for Cu are not very volatile.
- Hence, instead of RIE, Cu patterning is based on the so-called **Damascene** process where grooves are **first etched** in a dielectric layer, Cu is deposited on it, and the metal layer is **chemically-mechanically polished** down, leaving inlaid metal lines in the oxide grooves.

5.5.6. Other parameters:

in Eq. (eqn. 2) that can **minimize the RC time constant** are clearly the use of a **thick inter-metal dielectric layer** (once again within the limits of practicality in terms of deposition times), and as low a dielectric constant material as possible.

- Silicon dioxide **SIO_2** has a relative **dielectric constant** of **3.9**. These include organic materials such as polyimides, or xerogels/aerogels which have air pockets or porosity purposely built in to minimize the dielectric constant.
- a short n+ region can be implanted during the source/drain step and contacted at each end by one of the conductors.
- The other conductor can then cross over the oxide layer above the n+ region.
- Usually, this can be accomplished without appreciable increase in resistance, since the n+ region is heavily doped and its length can be made small.

MODULE-5b. Fabrication

SYLLABUS

Fabrication of p-n junctions Thermal Oxidation, Diffusion, Rapid Thermal Processing, Ion implantation, chemical vapour deposition, photolithography, Etching, metallization. (Text 1: 5.1)

Text Books: 1. Ben. G. Streetman, Sanjay Kumar Banergee, "Solid State Electronic Devices", 7thEdition, Pearson Education, 2016, ISBN 978-93-325-5508-2.

Contents

1. Fabrication of p-n junctions

The basic P-N Junction fabrication steps will be described in the following sections. Some of these steps may be carried out many times, in different combinations and/or processing conditions during a complete fabrication run.

1.1. Oxidation

- In **oxidation**, silicon reacts with oxygen to form silicon dioxide (SiO2). To speed up this chemical reaction, it is necessary to carry out the oxidation at high temperatures (e.g., 1000–1200°C) and inside ultraclean furnaces.

$$Si + O_2 \rightarrow SiO_2 \text{ (dry oxidation)}$$

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2 \text{ (wet oxidation)}$$

- To avoid the introduction of even small quantities of contaminants (which could significantly alter the electrical properties of the silicon), it is necessary to operate in a **clean room**.
- Particle filters are used to ensure that the airflow in the processing area is free from dust. All personnel must protect the clean-room environment by wearing special lint-free clothing that covers a person from head to toe.
- The oxygen used in the reaction can be introduced either as a high-purity gas (referred to as a "**dry oxidation**") or as steam (forming a "**wet oxidation**").

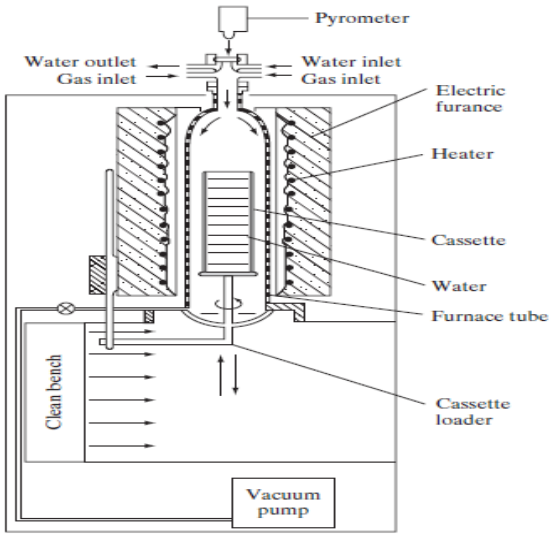


Figure 5–1b
Vertical furnace for large Si wafers. The silica wafer holder is loaded with Si wafers and moved into the furnace above for oxidation, diffusion, or deposition operations.

- In general, wet oxidation has a faster growth rate, but dry oxidation gives better electrical characteristics.
- The thermally grown oxide layer has excellent electrical insulation properties. can be used to form excellent MOS capacitors.
- Silicon dioxide can also serve as an effective mask against many impurities, allowing the introduction of dopants into the silicon only in regions that are not covered with oxide.
- Silicon dioxide is a transparent film, and the silicon surface is highly reflective. If white light is shone on an oxidized wafer, constructive and destructive interference will cause certain colors to be reflected.
- The wavelengths of the reflected light depend on the thickness of the oxide layer. In fact, by categorizing the color of the wafer surface, one can deduce the thickness of the oxide layer.

1.2. Diffusion

- **Diffusion** is a process by which atoms move from a high-concentration region to a low- concentration region.
- This is very much like a drop of ink dispersing through a glass of water except that it occurs much more slowly in solids.
- In VLSI fabrication, this is a method to introduce impurity atoms (dopants) into silicon to change its resistivity.
- The rate at which dopants diffuse in silicon is a strong function of temperature.
- Diffusion of impurities is usually carried out at high temperatures (1000–1200°C) to obtain the desired doping profile.
- When the wafer is cooled to room temperature, the impurities are essentially "frozen" in position.
- The diffusion process is performed in furnaces similar to those used for oxidation. The depth to which the impurities diffuse depends on both the temperature and the processing time.
- The most common impurities used as dopants are boron, phosphorus, and arsenic. Boron is a p-type dopant, while phosphorus and arsenic are n-type dopants.
- These dopants can be effectively masked by thin silicon dioxide layers. By diffusing boron into an n-type sub- strate, a pn junction is formed (diode).
- if the dopant atoms are supplied
- continuously, such that the concentration at the surface is maintained at a constant value, the distribution follows what is called a **complementary error function.**

- Common impurity source materials for diffusions in Si are B_2O_3, BBr_3, and BCl_3 for boron; phosphorus sources include PH_3, $P2O_5$, and $POCl_3$

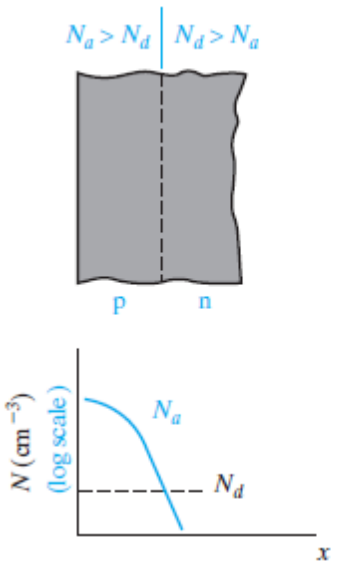


Figure 5–2 Impurity concentration profile for fabricating a p-n junction by diffusion.

-

1.3. Rapid Thermal Processing

- Rapid thermal processing (RTP) is a semiconductor manufacturing process which heats silicon wafers to high temperatures (over 1,000 °C) on a timescale of several seconds or less. During cooling, however, wafer temperatures must be brought down slowly to prevent dislocations and wafer breakage due to thermal shock.
- Such rapid heating rates are often attained by high intensity lamps or lasers. These processes are used for a wide variety of applications in semiconductor manufacturing including dopant activation, thermal oxidation, metal reflow and chemical vapor deposition

3 types of Heat Flow Mechanisms:

1.) Conduction: Flow of heat between two bodies in intimate contact.Heat flow per unit area in a solid is expressed in terms of a solids thermal conductivity

2.) Convection: Flow of heat between two bodies through an intermediate medium (a gas in our case)For a gas with effective heat transfer coefficient

3.) Radiation: Flow of heat between two bodies through radiation and absorption of light. use the spectral radiant exitance= the radiated power per area per unit wavelength,

3 Type of Process

1) rapid thermal oxidation,
2) annealing of ion implantation,
3) chemical vapor deposition,

- a single wafer is held (face down to minimize particulates) on low-thermal-mass quartz pins, surrounded by a bank of high-intensity (tens of kW) tungsten– halogen infrared lamps, with gold-plated reflectors around them.
- By turning on the lamps, the high-intensity infrared radiation shines through the quartz chamber and is absorbed by the wafer, causing its temperature to rise very rapidly (~50-100°C>s).
- The processing temperature can be reached quickly, after the gas flows have been stabilized in the chamber.
- At the of the process, the lamps are turned off, allowing the wafer temperature to drop rapidly, once again because of the much lower thermal mass of an RTP system compared to a furnace.
- In RTP, therefore, temperature is essentially used as a "switch" to start or quench the reaction. Two critical aspects of RTP are ensuring temperature uniformity across large wafers
- ,RTP operates at higher temperatures (~1000°C) but does so for only a few seconds (compared to minutes or hours in a furnace).

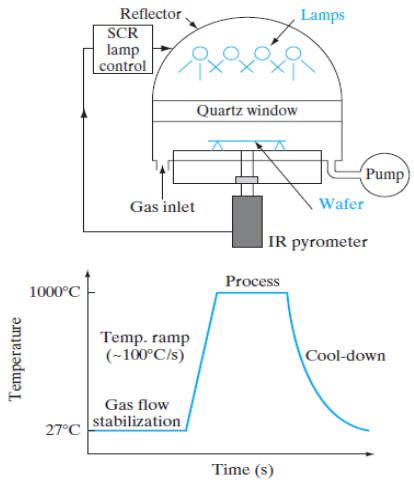


Figure 5–3 Schematic diagram of a rapid thermal processor and typical time–temperature profile.

1.4. Ion Implantation

- **Ion implantation** is another method used to introduce impurities into the semiconductor crystal. An ion implanter produces ions of the desired dopant, accelerates them by an electric field, and allows them to strike the semiconductor surface.
- Ion implantation can be used to form a deep region of doping using a **two step** procedure.
  - A high concentration of dopant is deposited near the surface in the **predeposition** or predep stage
  - The dopant source is then removed and the wafer heated to cause redistribution of the dopant via diffusion in the **drive-in stage**

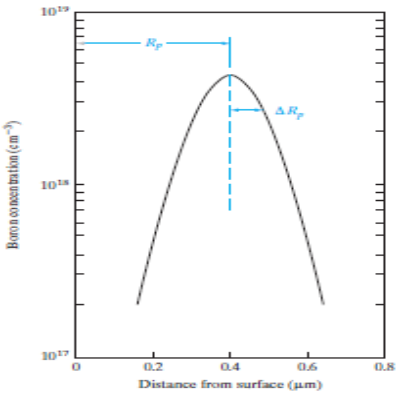


Figure 5–4 Distributions of implanted impurities: gaussian distribution of boron atoms about a projected range R_p (In this example, a dose of 10^{14} B atoms/cm^2 implanted at 140 keV).

- As the impurity atoms enter the crystal, they give up their energy to the lattice in collisions and finally come to rest at some average penetration depth, called the **projected range R_p**.
- Depending on the impurity and its implantation energy, the range in a given semiconductor may vary from a few hundred angstroms to about 1 μm.
- For most implantations the ions come to rest distributed almost evenly about the projected range R_p, as shown in Fig. 5–4. An implanted dose of h ions>cm2 is distributed approximately by a gaussian formula.

$$N(x) = \frac{\phi}{\sqrt{2\pi}\Delta R_p} \exp\left[-\frac{1}{2}\left(\frac{x - R_p}{\Delta R_p}\right)^2\right] \qquad (5\text{–}1a)$$

- where ΔRp, called the **straggle**, measures the half-width of the distribution at $e^{-1/2}$ of the peak (Fig. 5–4).
- Both Rp and ΔRp increase with increasing implantation energy. These parameters are shown as a function of energy

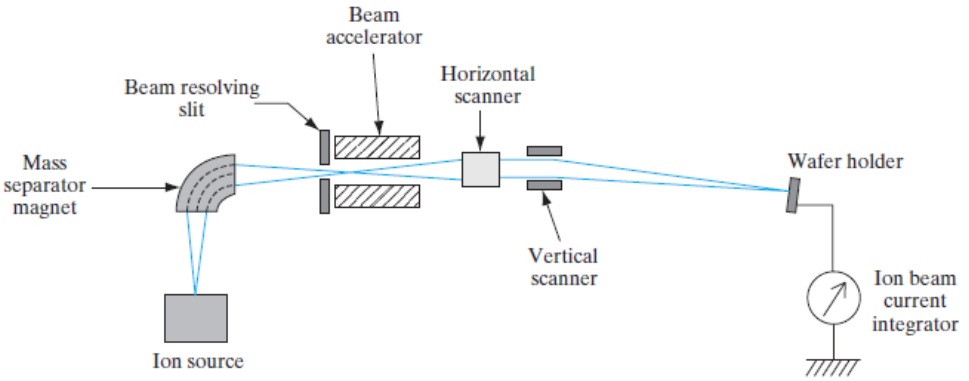


Figure 5–5
Schematic diagram of an ion implantation system.

- A gas containing the desired impurity is ionized within the source and is then extracted into the acceleration tube.
- After acceleration to the desired kinetic energy, the ions are passed through a mass separator to ensure that only the desired ion species enters the drift tube.
- The ion beam is then focused and scanned electrostatically over the surface of the wafer in the target chamber. Repetitive scanning in a raster pattern provides exceptionally uniform doping of the wafer surface.
- The target chamber commonly includes automatic wafer-handling facilities to speed up the process of implanting many wafers per hour.
- An obvious advantage of implantation is that it can be done at relatively low temperatures;
- this means that doping layers can be implanted without disturbing previously diffused regions. The ions can be blocked by metal or photoresist layers;
- therefore, the photolithographic techniques can be used to define ion-implanted doping patterns.
- Very shallow (tenths of a micron) and well-defined doping layers can be achieved by this method.
- Furthermore, it is possible to implant impurities which do not diffuse conveniently into semiconductors.

advantages

- the precise control of doping concentration it provides.
- Since the ion beam current can be measured accurately during implantation, a precise quantity of impurity can be introduced.

- This control over doping level, along with the uniformity of the implant over the wafer surface, make ion implantation particularly attractive for the fabrication of Si integrated circuits.

Disadvantage

- One problem with this doping method is the lattice damage which results from collisions between the ions and the lattice atoms.
- most of this damage can be removed in Si by heating the crystal after the implantation. This process is called **annealing**.
- Although Si can be heated to temperatures in excess of 1000°C without difficulty, GaAs and some other compounds tend to dissociate at high temperatures.
- For example, As evaporation from the surface of GaAs during annealing damages the sample. Therefore, it is common to encapsulate the GaAs with a thin layer of silicon nitride during the anneal.
- Another approach to annealing either Si or compounds is to heat the sample only briefly (e.g., 10 s) using RTP, rather than a conventional furnace. Annealing leads to some unintended diffusion of the implanted species.
- It is desirable to minimize this diffusion by optimizing the annealing time and temperature. The profile after annealing is given by

$$N(x) = \frac{\phi}{\sqrt{2\pi}(\Delta R_p^2 + 2Dt)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{(x - R_p)^2}{\Delta R_p^2 + 2Dt}\right)\right] \quad (5\text{–}1b)$$

1.5. Chemical vapor deposition (CVD)

- Chemical vapor deposition (CVD) is a process by which gases or vapors are chemically reacted, leading to the formation of solids on a substrate.
- CVD can be used to deposit various materials on a silicon substrate including SiO2, Si3N4, polysilicon, and so on.
- The properties of the CVD oxide layer are not as good as those of a thermally grown oxide, but they are sufficient to allow the layer to act as an electrical insulator.
- SiO2 films can also be formed by **low pressure chemical vapor deposition** (LPCVD) (Fig. 5–6) or **plasma enhanced CVD** (PECVD). The key differences are that thermal oxidation.

- consumes Si from the substrate, and very high temperatures are required whereas CVD of SiO2 does not consume Si from the substrate and can be done at much lower temperatures.
- The CVD process reacts a Si-containing gas such as SiH4 with an oxygen-containing precursor, causing a chemical reaction, leading to the deposition of SiO2 on the substrate.

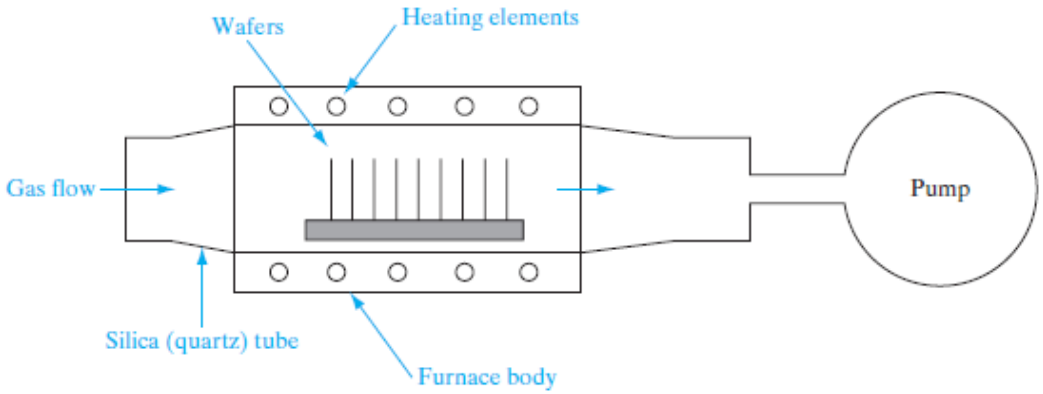


Figure 5–6
Low-pressure chemical vapor deposition (LPCVD) reactor.

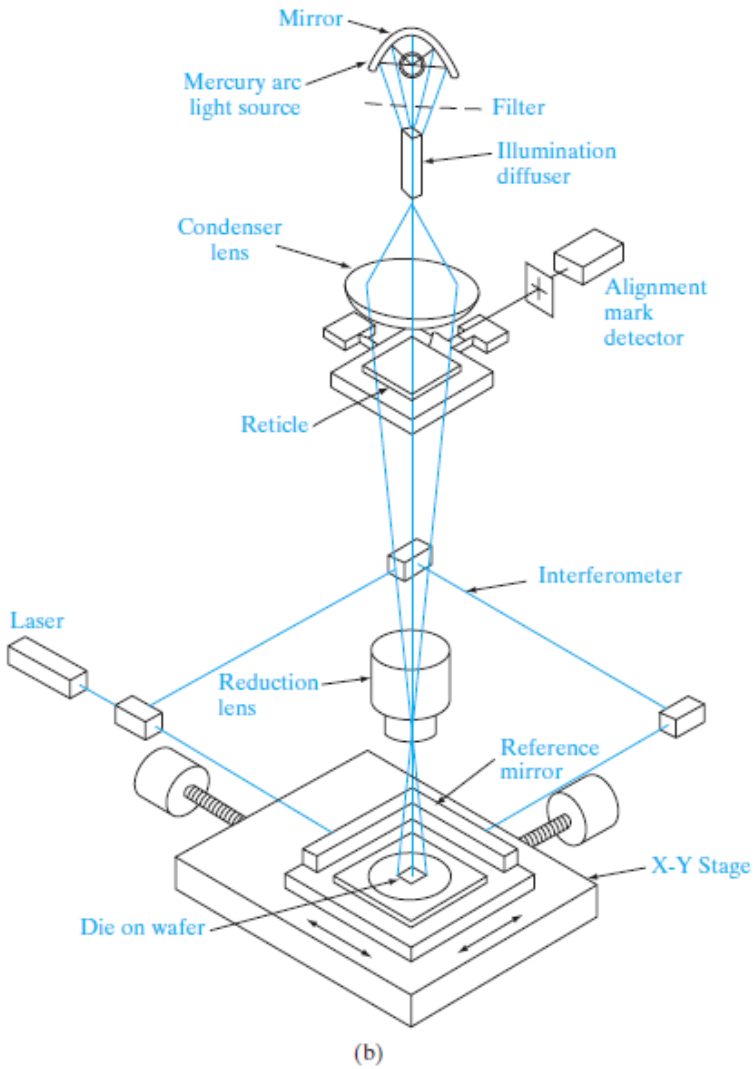
- As a complicated device structure is built up, the Si substrate may not be available for reaction, or there may be metallization on the wafer that cannot withstand very high temperatures.
- In such cases, CVD is a necessary alternative. Although we have used deposition of SiO2 as an important example, LPCVD is also widely used to deposit other dielectrics such as silicon nitride (Si3N4), and polycrystalline or amorphous Si.
- It should also be clear that the VPE of Si or MOCVD of compound semiconductors is really a special, more challenging example of CVD where not only must a film be deposited, but single-crystal growth must also be maintained.

1.6. Photolithography

- Lithography refers to the transfer of an image onto paper using a plate and ink-soluble grease.
- Photolithography is the transfer of an image using photographic techniques Photolithography transfers designer generated information (device placement and interconnections).

- **First step** in photolithography is to coat the surface with approx. 1 μm of **photoresist** (PR) PR will be the medium whereby the required image is transferred to the surface.

Figure 5–7b
Schematic diagram of an optical stepper.



•PR is often applied to the center of the wafer, which is then spun to force the PR over the entire surface
• the PR is approx. 1 μm thick while the wafer is 1000 μm thick. The PR is then **exposed to UV (ultraviolet)** radiation through a mask
• The masks generated from information about device placement and connection
• The UV radiation causes a chemical change in the PR

• The transfer of information from the mask to the surface occurs through the UV-induced chemical change - only occurs where the mask is transparent
The PR is then developed using a chemical developer
• **Two possibilities:**
– A **negative PR** is hardened against the developer by the UV radiation, and hence remains on the surface where UV shone through the mask
– A **positive PR** is the opposite, it is removed where the UV shone through the mask
• Assume a negative PR for this example, so the PR on the sides will be weakened and removed by the developer
• Once the **developer** has been washed off, the result is PR in the region corresponding to the transparent part of the mask (the mask is shown again to indicate where the final region is formed – it is not part of the final structure)
• Subsequent processing steps will use this structure to form device areas, interconnects, etc.
• Note that an optically reversed mask and a positive resist would give the same structure.
• the wave nature of light is manifested in terms of diffraction, which makes it harder to control the patterns. The diffraction-limited minimum geometry is given by

$$l_{min} = 0.8\,\lambda/NA \qquad (5\text{–}2a)$$

- where l is the wavelength of the light and NA (~0.5) is the numerical aperture or "size" of the lens used in the aligner.
- The other key parameter in lithography is the so-called **depth-of-focus(DOF)**, which is given by

$$\mathrm{DOF} = \frac{\lambda}{2(NA)^2} \qquad (5\text{–}2b)$$

- The DOF tells us the range of distances around the focal plane where the image quality is sharp. Unfortunately, this expression implies that exposure with very short wavelengths leads to poor DOF.
- We must therefore add steps in the fabrication process to planarize the surface **using chemical mechanical polishing** (CMP).

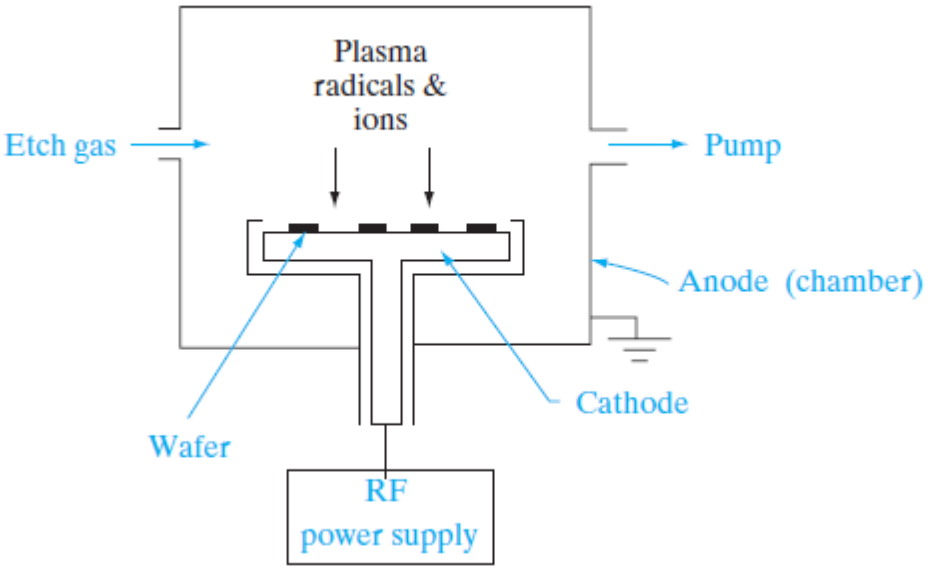
- As the name implies, the **planarizing process** is partly chemical in nature (using a basic solution), and partly mechanical grinding of the layers using an abrasive slurry.
- CMP can be achieved using a slurry of fine SiO2 particles in an NaOH solution. The expression for diffraction-limited geometry [Eq. (5–2a)] explains why there is interest in electron beam lithography. The de Broglie relation states that the wavelength of a particle varies inversely with its momentum:

$$\lambda = \frac{h}{p} \qquad (5\text{–}2c)$$

Thus, more massive or energetic particles have shorter wavelengths. Electron beams are easily generated, focused, and deflected. Since a 10-keV electron has a wavelength of about 0.1 Å, the linewidth limits become the size of the focused beam and its interaction with the photoresist layer.

1.7. Etching

- To permanently imprint the photographic patterns onto the wafer, chemical (**wet**) etching or RIE **dry etching** procedures can be used.
- Chemical etching is usually referred to as **wet etching**.
- Different chemical solutions can be used to remove different layers. For example, hydrofluoric (HF) acid can be used to etch SiO2, potassium hydroxide (KOH) for silicon, phosphoric acid for aluminum, and so on.
- Two most important issues in etching are **selectivity** and **anisotropy**.
- – **Selectivity** refers to the ability of an etchant to remove one material on the surface while leaving another intact.
- – **Isotropic** refers to the tendency of the etching to proceed laterally as well as downward
- The patterned photoresist layer can be used as an effective masking layer to protect materials below from wet chemical **etching** or **reactive ion etching** (RIE).
- Silicon dioxide, silicon nitride, polysilicon, and metal layers can be selectively removed using the appropriate etching methods.
- In wet etching, the chemical usually attacks the exposed regions that are not protected by the photoresist layer in all directions (**isotropic etching**).



-
- Depending on the thickness of the layer to be etched, a certain amount of undercut will occur. Therefore, the dimension of the actual pattern will differ slightly from the original pattern.
- If exact dimension is critical, RIE **dry etching** can be used.
- This method is essentially a directional bombardment of the exposed surface using a corrosive gas (or ions).
- The cross section of the etched layer is usually highly directional (**anisotropic etching**) and has the same dimension as the photoresist pattern. A comparison between isotropic and anisotropic etching is given in Fig. A.3.

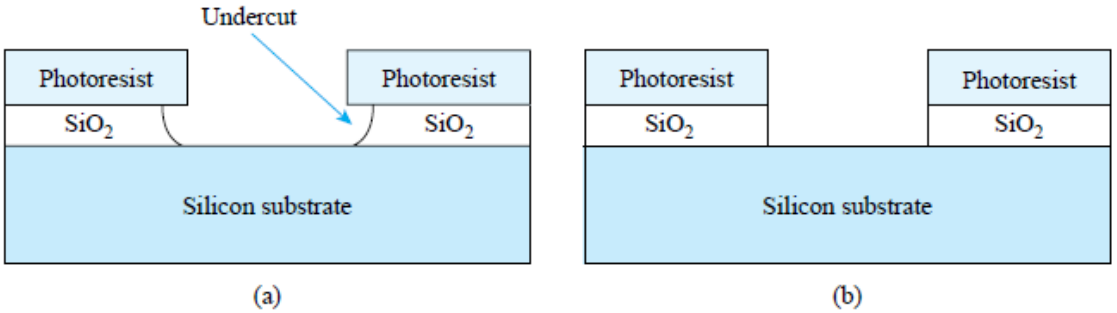


Figure A.3 (**a**) Cross-sectional view of an isotropic oxide etch with severe undercut beneath the photoresist layer. (**b**) Anisotropic etching, which usually produces a cross section with no undercut.

1.8. Metallization

- The purpose of metallization is to interconnect the various components (transistors, capacitors, etc.) to form the desired integrated circuit.
- Metallization involves the deposition of a metal over the entire surface of the silicon. The required interconnection pattern is then selectively etched.
- The metal layer is normally deposited via a sputtering process. A pure metal disk (e.g., 99.99% aluminum target) is placed under an Ar (argon) ion gun inside a vacuum chamber.
- The wafers are also mounted inside the chamber above the target. The Ar ions will not react with the metal, since argon is a noble gas. However, the ions are made to physically bombard the target and literally knock metal atoms out of the target.
- These metal atoms will then coat all the surface inside the chamber, including the wafers.
- The thickness of the metal film can be controlled by the length of the sputtering time, which is normally in the range of 1 to 2 minutes. The metal interconnects can then be defined using photolithography and etching steps.

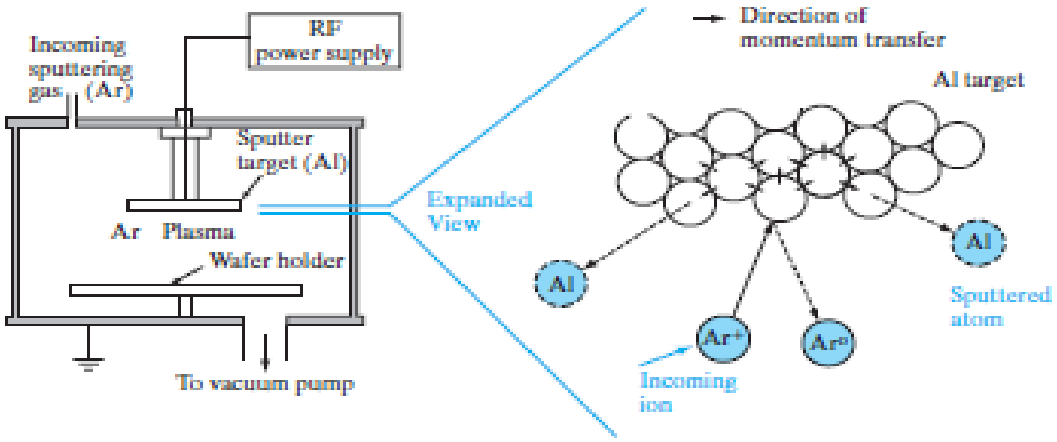


Figure 5–9
Aluminum sputtering by Ar^+ ions. The Ar^+ ions with energies of ~1–3 keV physically dislodge Al atoms which end up depositing on the Si wafers held in close proximity. The chamber pressures are kept low such that the mean free path of the ejected Al atoms is long compared to the target-to-wafer separation.