UNIT I - INTRODUCTION

INTRODUCTION TO INTEGRATED CIRCUIT TECHNOLOGY

There is no doubt that our daily lives are significantly affected by electronic engineering technology. This is true on the domestic scene, in our professional disciplines, in the workplace, and in leisure activities. Indeed, even at school, tomorrow's adults are exposed to and are coming to terms with quite sophisticated electronic devices and systems. There is no doubt that revolutionary changes have taken place in a relatively short time and it is also certain that even more dramatic advances will be made in the next decade.

Electronics as we know it today is characterized by reliability, low power dissipation, extremely low weight and volume, and low cost, coupled with an ability to cope easily with a high degree of sophistication and complexity. Electronics, and in particular the integrated circuit, has made possible the design of powerful and flexible processors which provide highly intelligent and adaptable devices for the user. Integrated circuit memories have provided the essential elements to complement these processors and, together with a wide range of logic and analog integrated circuitry, they have provided the system designer with components of considerable capability and extensive application. Furthermore, the revolutionary advances in technology have not yet by any means run their full course and the potential for future developments is exciting to say the least.

Up until the 1950s electronic active device technology was dominated by the vacuum tube and, although a measure of miniaturization and circuit integration did take place, the technology did not lend itself to miniaturization as we have come to accept it today. Thus the vast majority of present-day electronics is the result of the invention of the transistor in 1947.

The invention of the transistor by William B. Shockley, Walter H. Brattain and John Bardeen of Bell Telephone Laboratories was followed by the development of the Integrated Circuit (IC). The very first IC emerged at the beginning of 1960 and since that time there have already been four generations of ICs: SSI (small scale integration), MSI (medium scale integration), LSI (large scale integration), and VLSI (very large scale integration). Now we are beginning to see the emergence of the fifth generation, ULSI (ultra large scale integration) which is characterized by complexities in excess of 3 million devices on a single IC chip. Further miniaturization is still to come and more revolutionary advances in the application of this technology must inevitably occur.

Over the past several years, Silicon CMOS technology has become the dominant fabrication process for relatively high performance and cost effective VLSI circuits. The revolutionary nature of this development is indicated by the way in which the number of transistors integrated in circuits on a single chip has grown as indicated in Figure 1.1. Such progress is highlighted by recent products such as RISC chips in which it is possible to process some

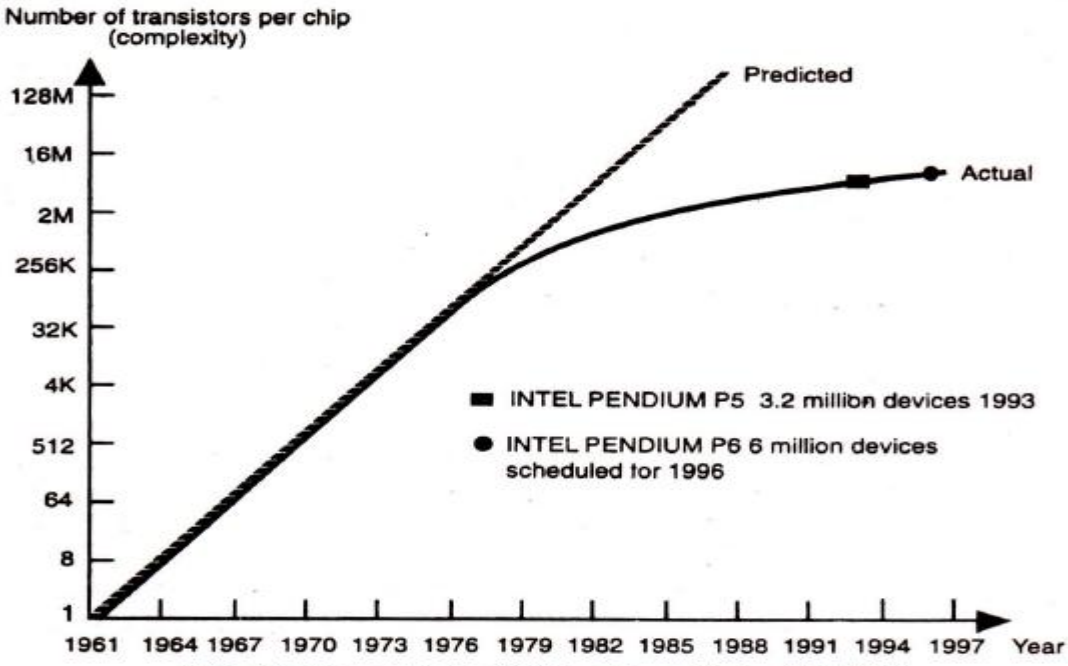


FIGURE 1.1 Moore's first law: Transistors integrated on a single chip (commercial products).

35 million instructions per second. In order to improve on this throughput rate it will be necessary to improve the technology, both in terms of scaling and processing, and through the incorporation of other enhancements such as BiCMOS. The implications of this approach is that existing silicon technology could effectively facilitate the tripling of rate. Beyond this, i.e., above 100 million instructions per second, one must look to other technologies. In particular, the emerging Gallium Arsenide(GaAs) based technology will be most significant in this area of ultra high speed logic/fast digital processors. GaAs also has further potential as a result of its photo-electronic properties, both as a receiver and as a transmitter of light. GaAs in combination with silicon will provide the designer with some very exciting possibilities.

It is most informative in assessing the role of the currently available technologies to review their speed and power performance domains. This has been set out as Figure 1.2 and the potential presented by each may be readily assessed.

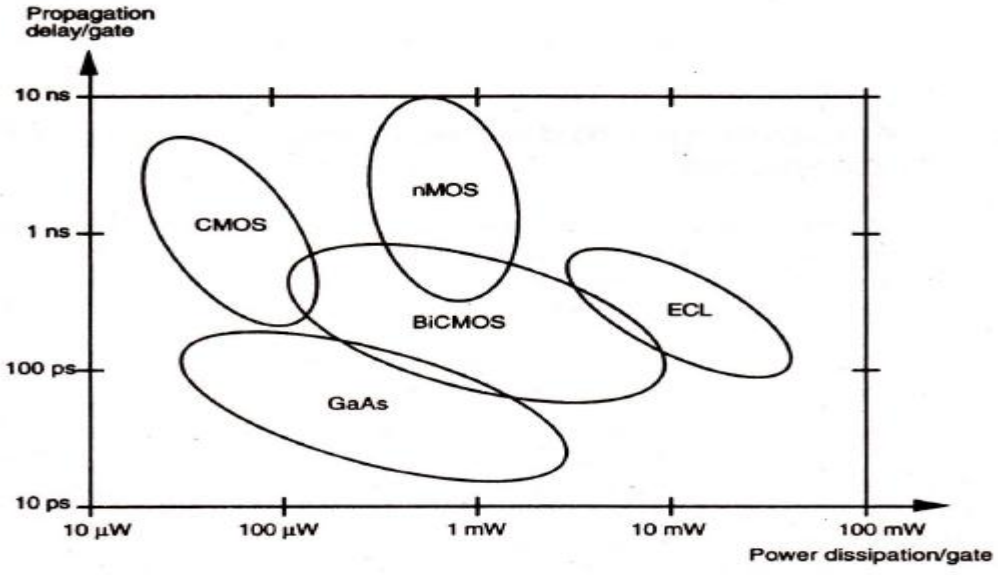


FIGURE 1.2 Speed/power performance of available technologies.

THE INTEGRATED CIRCUIT (IC) ERA

Such has been the potential of the silicon integrated circuit that there has been an extremely rapid growth in the number of transistors (as a measure of complexity) being integrated into circuits on a single silicon chip. In less than three decades, this number has risen from tens to millions as can be seen in Figure 1.1. The figure sets out what has become known as "Moore's first law" after predictions made by Gordon Moore (of Intel) in the 1960s. It may be seen that his predictions have largely come true except for an increasing divergence between "predicted" and "actual" over the last few years due to problems associated with the complexities involved in designing and testing such very large circuits.

Such has been the impact of this revolutionary growth that IC technology now affects almost every aspect of our lives. More is still to come since we have not yet reached the limits of miniaturization and there is no doubt that tens of millions of transistors will be readily integrated onto a single chip in the future. This evolutionary process is reflected in Table 1.1.

Truly the 1970s, the 1980s and now the 1990s may well be described as the integrated circuit era.

METAL-OXIDE-SEMICONDUCTOR (MOS) AND RELATED VLSI TECHNOLOGY

Within the bounds of MOS technology, the possible circuit realizations may be based on pMOS, nMOS, CMOS and now BiCMOS devices.

However, this text will deal with nMOS, then with CMOS (which includes pMOS transistors) and BICMOS, and finally with GaAs technology, all of which may be classed as leading Integrated circuit technologies.

Although CMOS is the dominant technology, some of the examples used to illustrate the design processes will be presented in nMOS form. The reasons for this are as follows:

- For nMOS technology, the design methodology and the design rules are easily learned, thus providing a simple but excellent introduction to structured design for VLSI.
- nMOS technology and design processes provide an excellent background for other technologies. In particular, some familiarity with nMOS allows a relatively easy transition to CMOS technology and design.
- For GaAs technology some arrangements in relation to logic design are similar to those employed in nMOS technology. Therefore, understanding the basics of nMOS design will assist in the layout of GaAs circuits.

Not only is VLSI technology providing the user with a new and more complex range of 'off the shelf' circuits, but VLSI design processes are such that system designers can readily design their own special circuits of considerable complexity. This provides a new degree of freedom for designers and it is probable that some very significant advances will result. Couple this with the fact that integration density is increasing rapidly, as advances in technology shrink the feature size for circuits integrated in silicon. Typical manufacturers'

commercial IC products have shown this trend quite clearly as shown in Figure 1.3 and, simultaneously, the effectiveness of the circuits produced has increased with scaling down. A common measure of effectiveness is the speed power product of the basic logic gate circuit of the technology (for nMOS, the *Nor* gate, with *Nand* and *Nor* gates for CMOS). Speed power product is measured in picojoules (pJ) and is the product of the gate switching delay in nanoseconds and the gate power dissipation in milliwatts. Typical figures are implied in Figure 1.2.

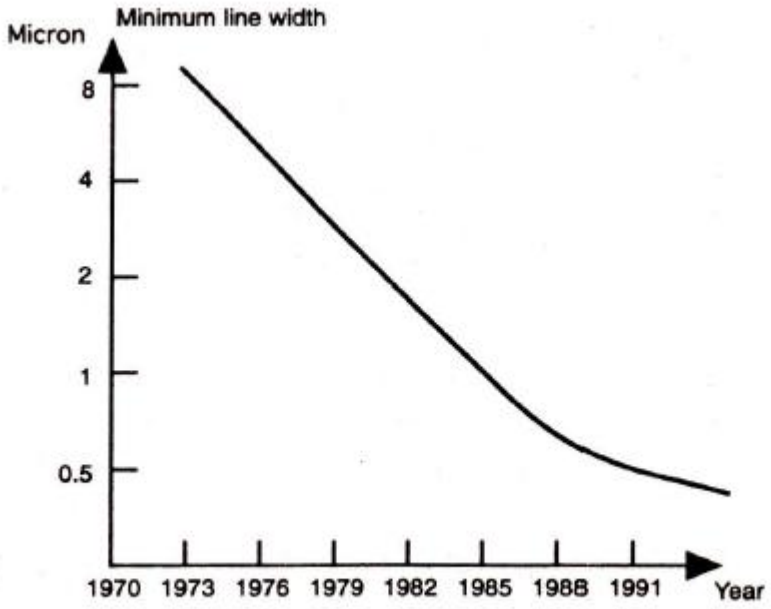


FIGURE 1.3 Approximate minimum line width of commercial products versus year.

TABLE 1.1 Microelectronics evolution

Year	*1947*	*1950*	*1961*	*1966*	*1971*	*1980*	*1990*	*2000*
Technology	*Invention of the transistor*	*Discrete components*	*SSI*	*MSI*	*LSI*	*VLSI*	*ULSI**	*GSI†*
Approximate numbers of transistors per chip in commercial products	1	1	10	100–1000	1000–20,000	20,000–1,000,000	1,000,000 10,000,000	>10,000,000
Typical products	—	Junction Transistor and diode	Planar devices Logic gates Flip-flops	Counters Multiplexers Adders	8 bit micro-processors ROM RAM	16 and 32 bit micro-processors Sophisticated peripherals GHM Dram	Special processors, Virtual reality machines, smart sensors	

* Ultra large-scale integration
† Giant-scale integration

BASIC MOS TRANSISTORS

Having now established some background, let us turn our attention to basic MOS processes and devices. In particular, let us examine the basic nMOS enhancement and depletion mode transistors as shown in Figures 1.4(a) and (b).

nMOS devices are formed in a p-type substrate of moderate doping level. The source and drain regions are formed by diffusing n-type impurities through suitable masks into these areas to give the desired n-impurity concentration and give rise to depletion regions which extend mainly in the more lightly doped p-region as shown. Thus, source and drain are isolated from one another by two diodes. Connections to the source and drain are made by a deposited metal layer. In order to make a useful device, there must be the capability for establishing and controlling a current between source and drain, and this is commonly achieved in one of two ways, giving rise to the enhancement mode and depletion mode transistors.

Consider the enhancement mode device first, shown in Figure 1.4(a). A polysilicon gate is deposited on a layer of insulation over the region between source and drain. Figure 1.4(a) shows a basic enhancement mode device in which the channel is not established and the device is in a non-conducting condition, $V_D = V_S = V_{gs} = 0$. If this gate is connected to a suitable positive voltage with respect to the source, then the electric field established between the gate and the substrate gives rise to a charge inversion region in the substrate under the gate insulation and a conducting path or channel is formed between source and drain.

The channel may also be established so that it is present under the condition $V_{gs} = 0$ by implanting suitable impurities in the region between source and drain during manufacture and prior to depositing the insulation and the gate. This arrangement is shown in Figure 1.4(b). Under these circumstances, source and drain are connected by a conducting channel, but the channel may now be closed by applying a suitable negative voltage to the gate.

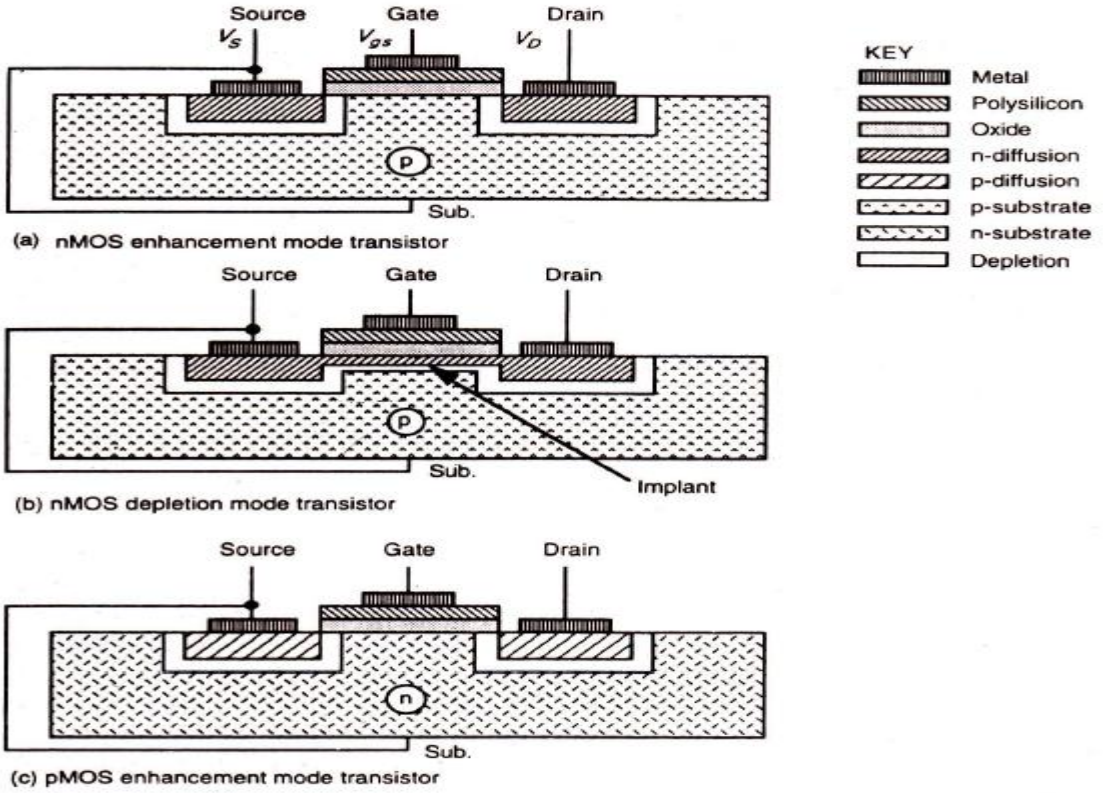


FIGURE 1.4 MOS transistors (V_D = 0 V. Source gate and substrate to 0 V).

In both cases, variations of the gate voltage allow control of any current flow between source and drain.

Figure 1.4(c) shows the basic pMOS transistor structure for an enhancement mode device. In this case the substrate is of n-type material and the source and drain diffusions are

consequently p-type. In the figure, the conditions shown are those for an unbiased device; however, the application of a *negative* voltage of suitable magnitude ($> |V_t|$) between gate and source will give rise to the formation of a channel (p-type) between the source and drain and current may then flow if the drain is made negative with respect to the source. In this case the current is carried by holes as opposed to electrons (as is the case for nMOS devices). In consequence, pMOS transistors are inherently slower than nMOS, since hole mobility μ_p is less, by a factor of approximately 2.5, than electron mobility μ_n. However, bearing these differences in mind, the discussions of nMOS transistors which follow relate equally well to pMOS transistors.

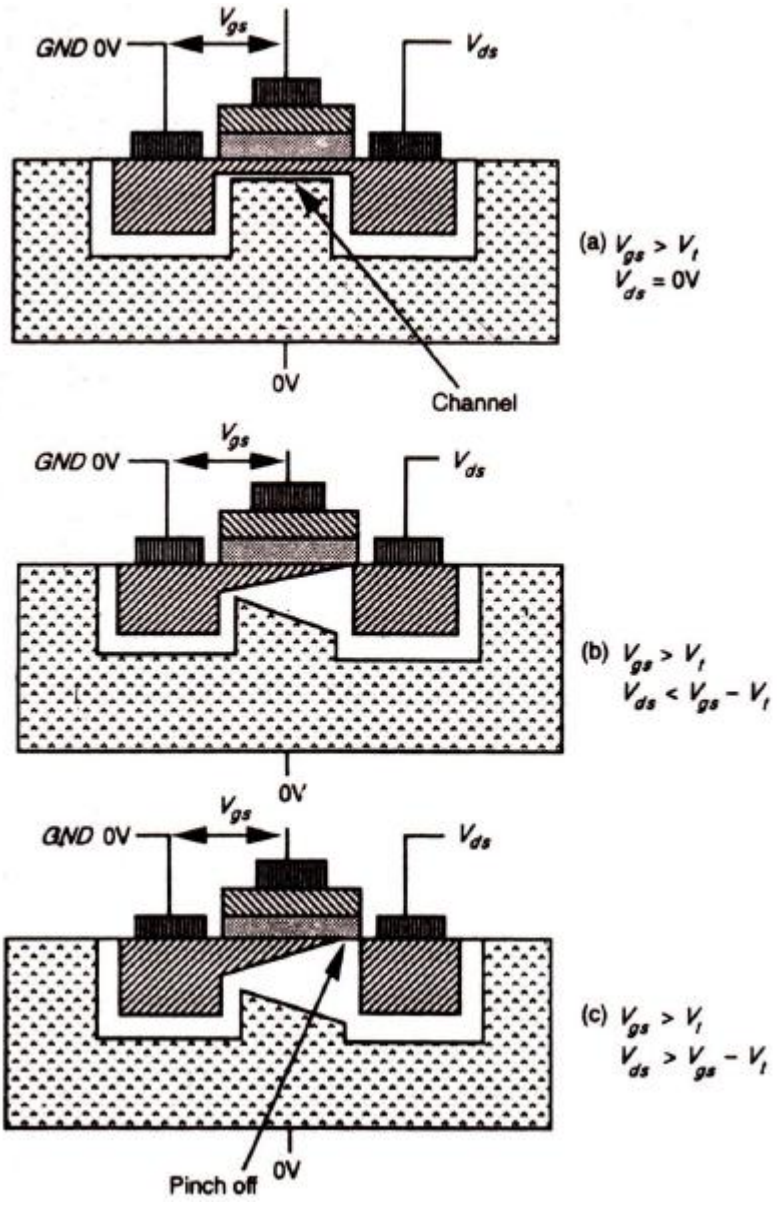
ENHANCEMENT MODE TRANSISTOR ACTION

To gain some understanding of this mechanism, let us further consider the enhancement mode device, as in Figure 1.5, under three sets of conditions. It must first be recognized that in order to establish the channel in the first place a minimum voltage level of *threshold voltage* V_t must be established between gate and source (and of course between gate and substrate as a result). Figure 1.5(a) then indicates the conditions prevailing with the channel established but no current flowing between source and drain ($V_{ds} = 0$). Now consider the conditions prevailing when current flows in the channel by applying a voltage V_{ds} between drain and source. There must, of course, be a corresponding IR drop = V_{ds} along the channel. This results in the voltage between gate and channel varying with distance along the channel with the voltage being a maximum of V_{gs} at the source end. Since the effective gate voltage is $V_g = V_{gs} - V_t$, (no current flows when $V_{gs} < V_t$) there will be voltage available to invert the channel at the drain end so long as $V_{gs} - V_t \geq V_{ds}$. The limiting condition comes when $V_{ds} = V_{gs} - V_t$. For all voltages $V_{ds} < V_{gs} - V_t$, the device is in the non-saturated region of operation which is the condition shown in Figure 1.5(b).

Consider now what happens when V_{ds} is increased to a level greater than $V_{gs} - V_t$. In this case, an IR drop = $V_{gs} - V_t$ takes place over less than the whole length of the channel so that over part of the channel, near the drain, there is insufficient electric field available to give rise to an inversion layer to create the channel. The channel is, therefore, 'pinched off' as indicated in Figure 1.5(c). Diffusion current completes the path from source to drain in this case, causing the channel to exhibit a high resistance and behave as a constant current source. This region, known as *saturation,* is characterized by almost constant current for increase of V_{ds} above $V_{ds} = V_{gs} - V_t$. In all cases, the channel will cease to exist and no current will flow when $V_{gs} < V_t$. Typically, for enhancement mode devices, $V_t = 1$ volt for $V_{DD} = 5$ V or, in general terms, $V_t = 0.2\ V_{DD}$.

DEPLETION MODE TRANSISTOR ACTION

For depletion mode devices the channel is established, due to the implant, even when $V_{gs} = 0$, and to cause the channel to cease to exist a negative voltage V_{td} must be applied between gate and source.



Note: V_{ds} is the drain-to-source voltage. Substrate assumed connected to 0 V.

FIGURE 1.5 Enhancement mode transistor for particular values of V_{ds} with ($V_{gs} > V_t$).

V_{td} is typically $< -0.8\ V_{DD}$, depending on the implant and substrate bias, but, threshold voltage differences apart, the action is similar to that of the enhancement mode transistor.

Commonly used symbols for nMOS and pMOS transistors are set out in Figure 1.6.

nMOS FABRICATION

A brief introduction to the general aspects of the polysilicon gate self-aligning nMOS fabrication process will now be given. As well as being relevant in their own right, the fabrication processes used for nMOS are relevant to CMOS and BiCMOS which may be viewed as involving additional fabrication steps. Also, it is clear that an appreciation of the fabrication processes will give an insight into the way in which design information must be presented and into the reasons for certain performance characteristics and limitations. An nMOS process is illustrated in Figure 1.7 and may be outlined as follows:

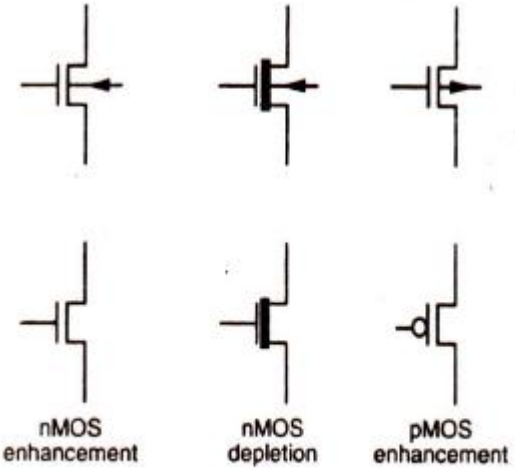


FIGURE 1.6 Transistor circuit symbols.

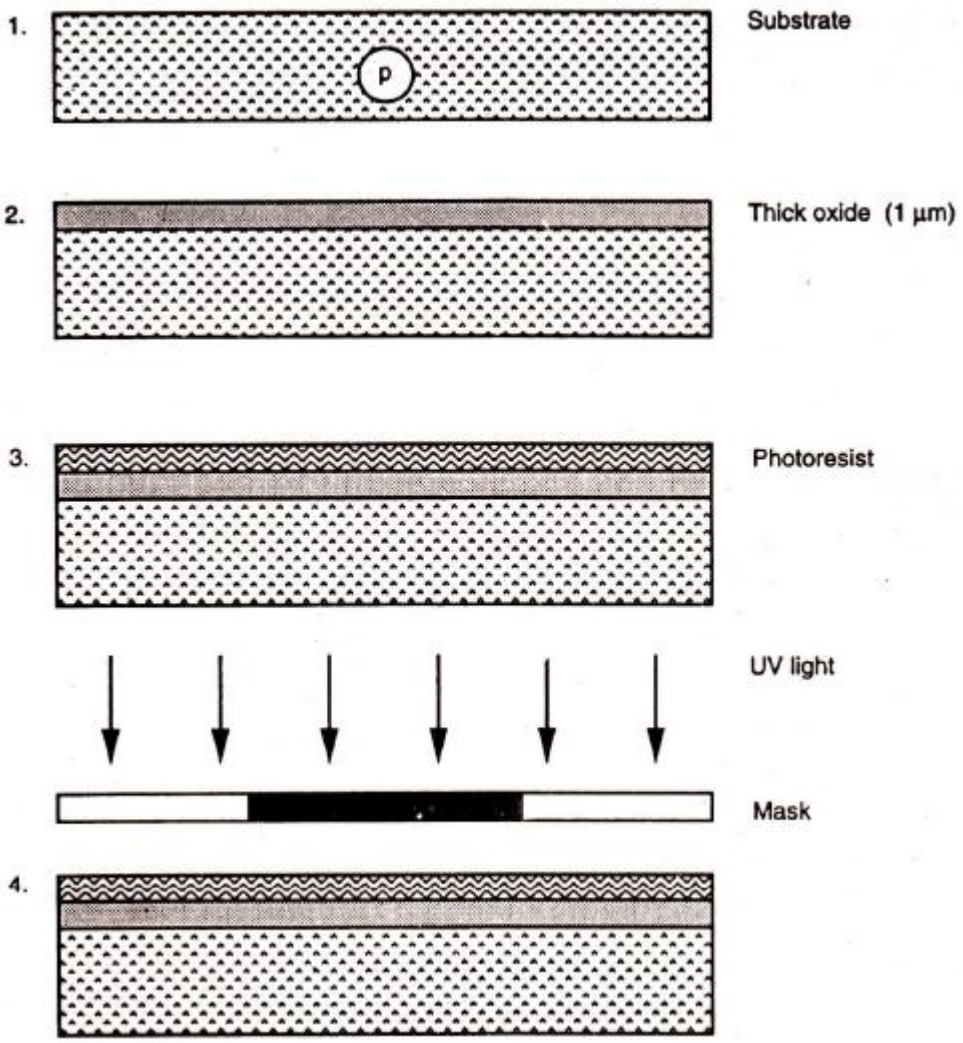
1. Processing is carried out on a thin wafer cut from a single crystal of silicon of high purity into which the required p-impurities are introduced as the crystal is grown. Such wafers are typically 75 to 150 mm in diameter and 0.4 mm thick and are doped with, say, boron to impurity concentrations of $10^{15}/cm^3$ to $10^{16}/cm^3$, giving resistivity in the approximate range 25 ohm cm to 2 ohm cm.
2. A layer of silicon dioxide (SiO_2), typically 1 μm thick, is grown all over the surface of the wafer to protect the surface, act as a barrier to dopants during processing, and provide a generally insulating substrate on to which other layers may be deposited and patterned.
3. The surface is now covered with a photoresist which is deposited onto the wafer and spun to achieve an even distribution of the required thickness.
4. The photoresist layer is then exposed to ultraviolet light through a mask which defines those regions into which diffusion is to take place together with transistor channels. Assume, for example, that those areas exposed to ultraviolet radiation are polymerized (hardened), but that the areas required for diffusion are shielded by the mask and remain unaffected.
5. These areas are subsequently readily etched away together with the underlying silicon dioxide so that the wafer surface is exposed in the window defined by the mask.
6. The remaining photoresist is removed and a thin layer of SiO_2 (0.1 μm typical) is grown over the entire chip surface and then polysilicon is deposited on top of this to form the gate structure. The polysilicon layer consists of heavily doped polysilicon deposited by chemical vapor deposition (CVD). In the fabrication of fine pattern devices, precise control of thickness, impurity concentration, and resistivity is necessary.
7. Further photoresist coating and masking allows the polysilicon to be patterned (as shown in Step 6) and then the thin oxide is removed to expose areas into which n-type impurities are to be diffused to form the source and drain as shown. Diffusion is achieved by heating the wafer to a high temperature and passing a gas containing the desired n-type impurity (for example, phosphorus) over the surface as indicated in Figure 1.8. Note that the polysilicon with underlying thin oxide act as masks during diffusion—the process is self-aligning.
8. Thick oxide (SiO_2) is grown over all again and is then masked with photoresist and etched to expose selected areas of the polysilicon gate and the drain and source areas where connections (i.e. contact cuts) are to be made.

9. The whole chip then has metal (aluminum) deposited over its surface to a thickness typically of 1 μm. This metal layer is then masked and etched to form the required interconnection pattern.

It will be seen that the process revolves around the formation or deposition and patterning of three layers, separated by silicon dioxide insulation. The layers are diffusion within the substrate, polysilicon on oxide on the substrate, and metal insulated again by oxide.

To form depletion mode devices it is only necessary to introduce a masked ion implantation step between steps 5 and 6 or 6 and 7 in Figure 1.7. Again, the thick oxide acts as a mask and this process stage is also self-aligning.

Consideration of the processing steps will reveal that relatively few masks are needed and the self-aligning aspects of the masking processes greatly ease the problems of mask registration. In practice, some extra process steps are necessary, including the overglassing of the whole wafer, except where contacts to the outside world are required. However, the process is basically straightforward to envisage and circuit design eventually comes down to the business of delineating the masks for each stage of the process.



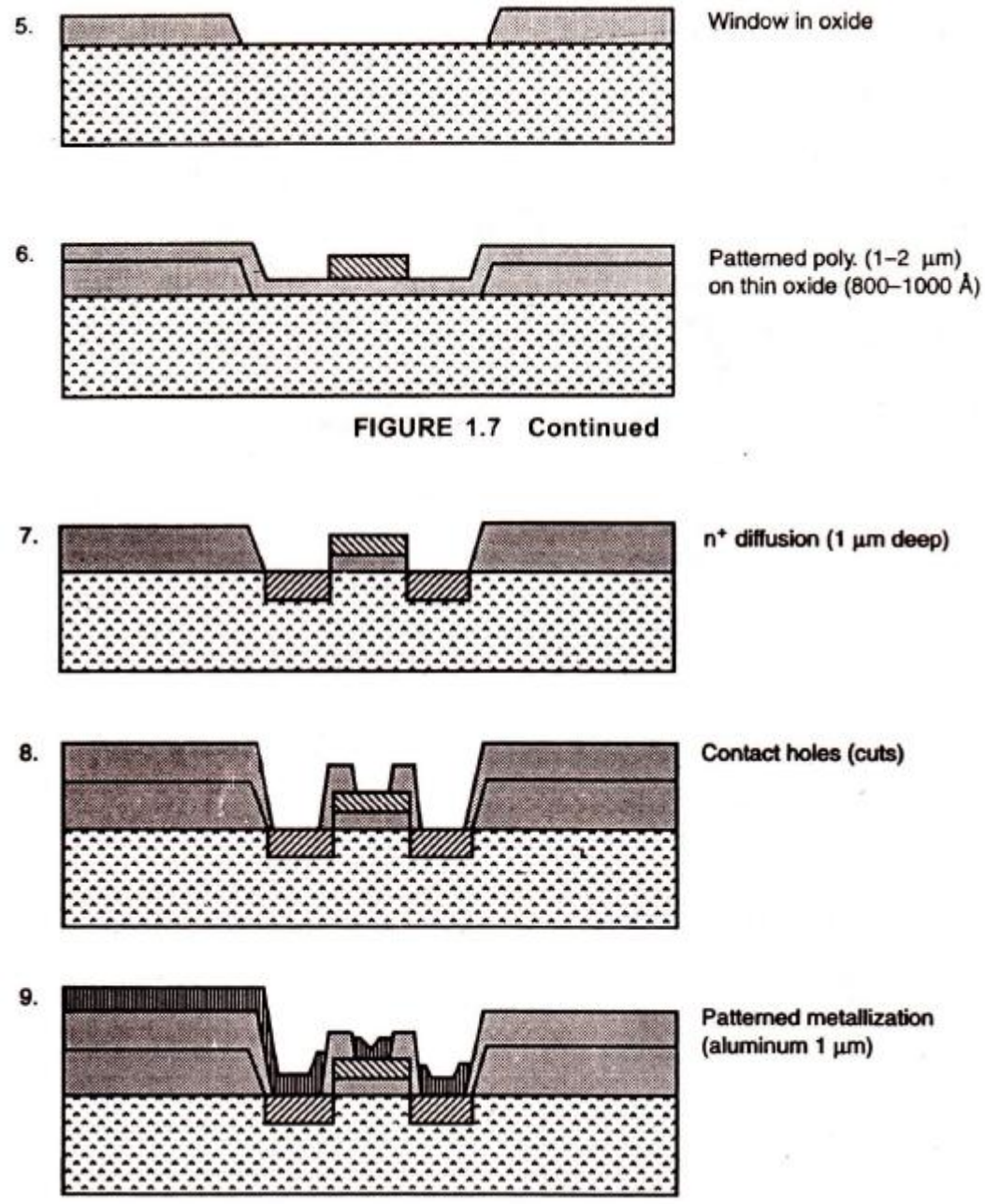


FIGURE 1.7 Continued

FIGURE 1.7 nMOS fabrication process.

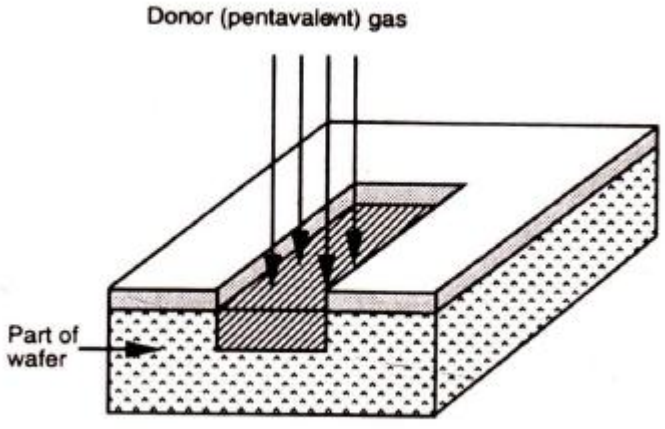


FIGURE 1.8 Diffusion process.

Summary of An nMOS Process

- Processing takes place on a p-doped silicon crystal wafer on which is grown a 'thick' layer of SiO_2.
- *Mask* 1—Pattern SiO_2 to expose the silicon surface in areas where paths in the diffusion layer or gate areas of transistors are required. Deposit thin oxide over all. For this reason, this mask is often known as the *'thinox' mask* but some texts refer to it as the *diffusion mask.*
- *Mask* 2—Pattern the ion implantation within the thinox region where depletion mode devices are to be produced—*self-aligning.*
- *Mask* 3—Deposit polysilicon over all (1.5 μm thick typically), then pattern using Mask 3. Using the same mask, remove thin oxide layer where it is not covered by polysilicon.
- Diffuse n^+ regions into areas where thin oxide has been removed. Transistor drains and sources are thus self-aligning with respect to the gate structures.
- *Mask* 4—Grow thick oxide over all and then etch for contact cuts.
- *Mask* 5—Deposit metal and pattern with Mask 5.
- *Mask* 6—Would be required for the overglassing process step.

CMOS FABRICATION

There are a number of approaches to CMOS fabrication, including the p-well, the n-well, the twin-tub, and the silicon-on-insulator processes. In order to introduce the reader to CMOS design we will be concerned mainly with well-based circuits. The p-well process is widely used in practice and the n-well process is also popular, particularly as it is an easy retrofit to existing nMOS lines, so we will also discuss it briefly.

For the lambda-based rules set out later, we will assume a p-well process.

The p-well Process

A brief overview of the fabrication steps may be obtained with reference to Figure 1.9, noting that the basic processing steps are of the same nature as those used for nMOS.

In primitive terms, the structure consists of an n-type substrate in which p-devices may be formed by suitable masking and diffusion and, in order to accommodate n-type devices, a deep p-well is diffused into the n-type substrate as shown.

This diffusion must be carried out with special care since the p-well doping concentration and depth will affect the threshold voltages as well as the breakdown voltages of the n-transistors. To achieve low threshold voltages (0.6 to 1.0 V) we need either deep-well diffusion or high-well resistivity. However, deep wells require larger spacing between the n- and p-type transistors and wires due to lateral diffusion and therefore a larger chip area.

The p-wells act as substrates for the n-devices within the parent n-substrate, and, provided that voltage polarity restrictions are observed, the two areas are electrically isolated. However, since there are now in effect two substrates, two substrate connections (V_{DD} and V_{SS}) are required, as shown in Figure 1.10.

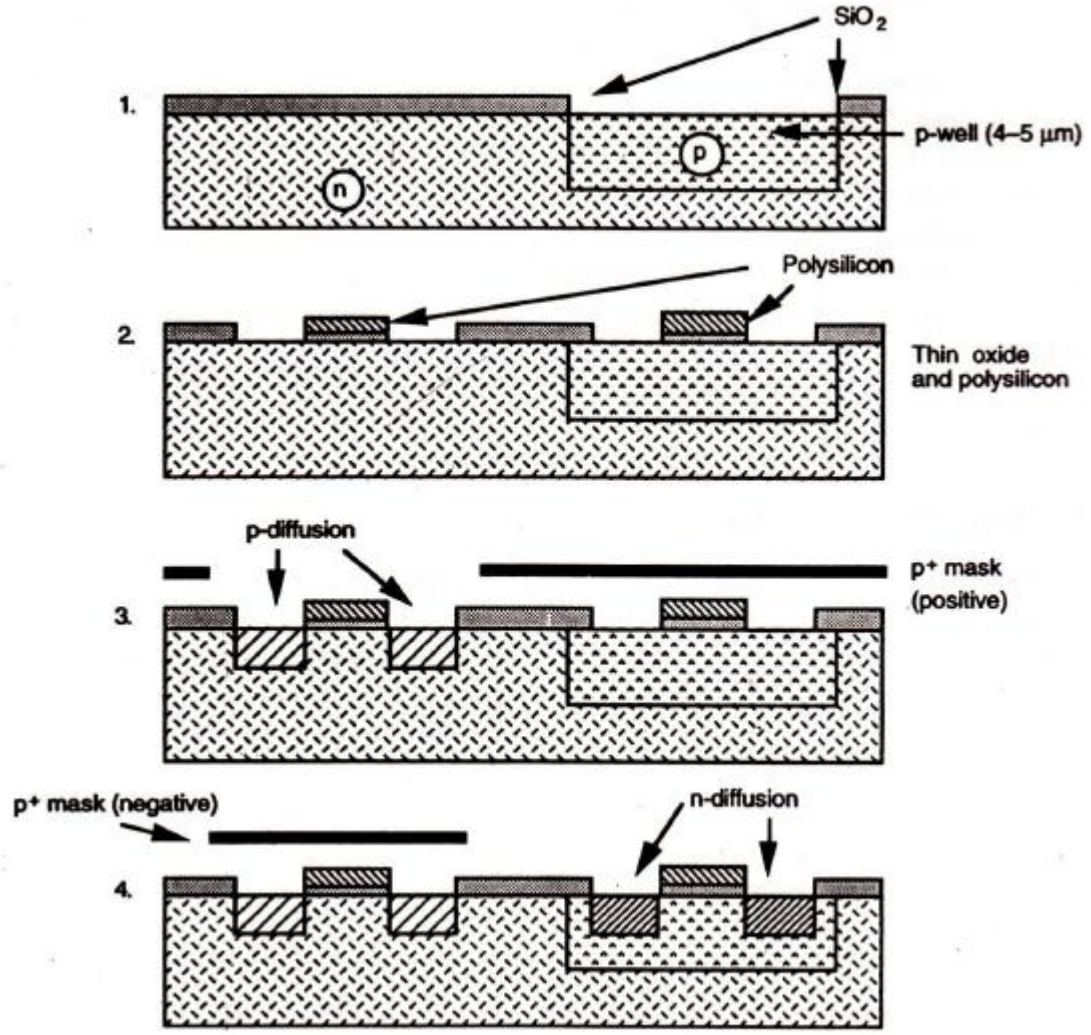


FIGURE 1.9 CMOS p-well process steps.

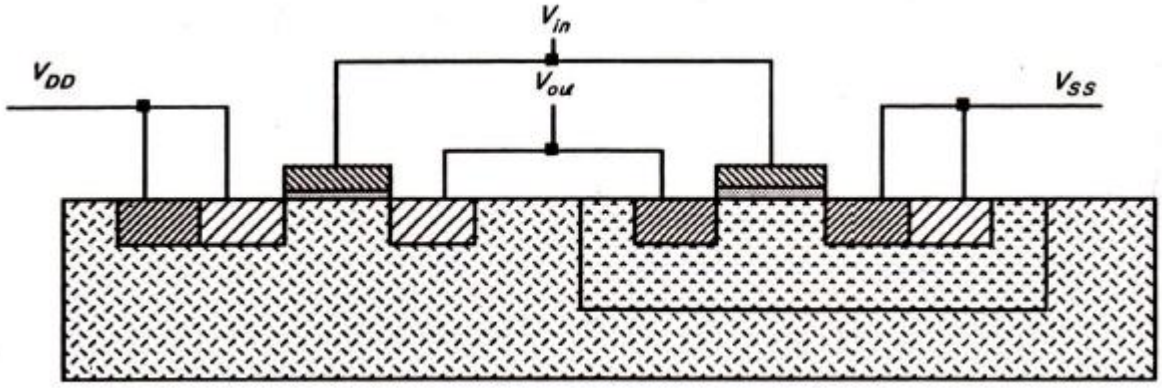


FIGURE 1.10 CMOS p-well inverter showing V_{DD} and V_{SS} substrate connections.

In all other respects—masking, patterning, and diffusion—the process is similar to nMOS fabrication. In summary, typical processing steps are:

- *Mask* 1 — defines the areas in which the deep p-well diffusions are to take place.
- *Mask 2* — defines the thinox regions, namely those areas where the thick oxide is to be stripped and thin oxide grown to accommodate p- and n-transistors and wires.
- *Mask* 3 — used to pattern the polysilicon layer which is deposited after the thin oxide.
- *Mask* 4 — A p-plus mask is now used (to be in effect "Anded" with Mask 2) to define all areas where p-diffusion is to take place.
- *Mask* 5 — This is usually performed using the negative form of the p-plus mask and defines those areas where n-type diffusion is to take place.
- *Mask* 6 — Contact cuts are now defined.
- *Mask* 7 — The metal layer pattern is defined by this mask.
- *Mask* 8 — An overall passivation (overglass) layer is now applied and Mask 8 is needed to define the openings for access to bonding pads.

The n-well Process

As indicated earlier, although the p-well process is widely used, n-well fabrication has also gained wide acceptance, initially as a retrofit to nMOS lines.

N-well CMOS circuits are also superior to p-well because of the lower substrate bias effects on transistor threshold voltage and inherently lower parasitic capacitances associated with source and drain regions.

Typical n-well fabrication steps are illustrated in Figure 1.11. The first mask defines the n-well regions. This is followed by a low dose phosphorus implant driven in by a high temperature diffusion step to form the n-wells. The well depth is optimized to ensure against p-substrate to p^+ diffusion breakdown without compromising the n-well to n^+ mask separation. The next steps are to define the devices and diffusion paths, grow field oxide, deposit and pattern the polysilicon, carry out the diffusions, make contact cuts, and finally metalize as before.

It will be seen that an n^+ mask and its complement may be used to define the n- and p-diffusion regions respectively. These same masks also include the V_{DD} and V_{SS} contacts (respectively). It should be noted that, alternatively, we could have used a p^+ mask and its complement, since the n^+ and p^+ masks are generally complementary.

By way of illustration, Figure 1.12 shows an inverter circuit fabricated by the n-well process, and this may be directly compared with Figure 1.10.

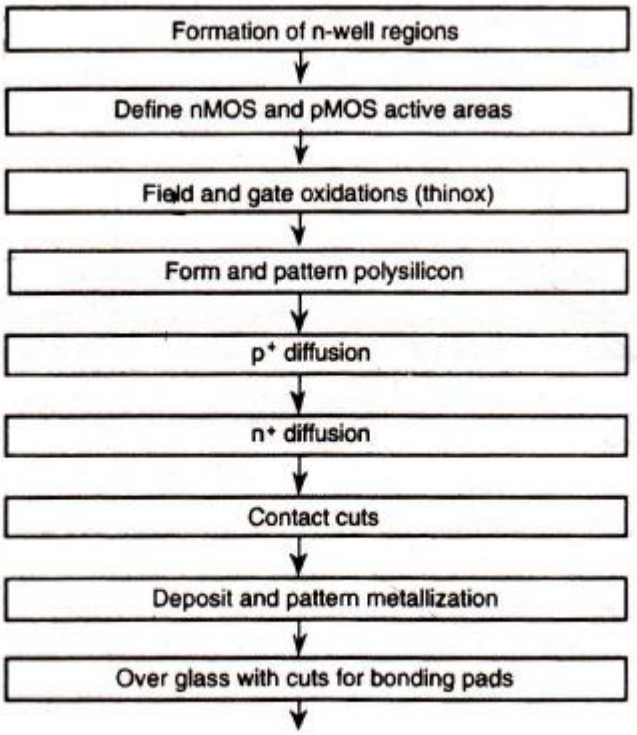


FIGURE 1.11 Main steps in a typical n-well process.

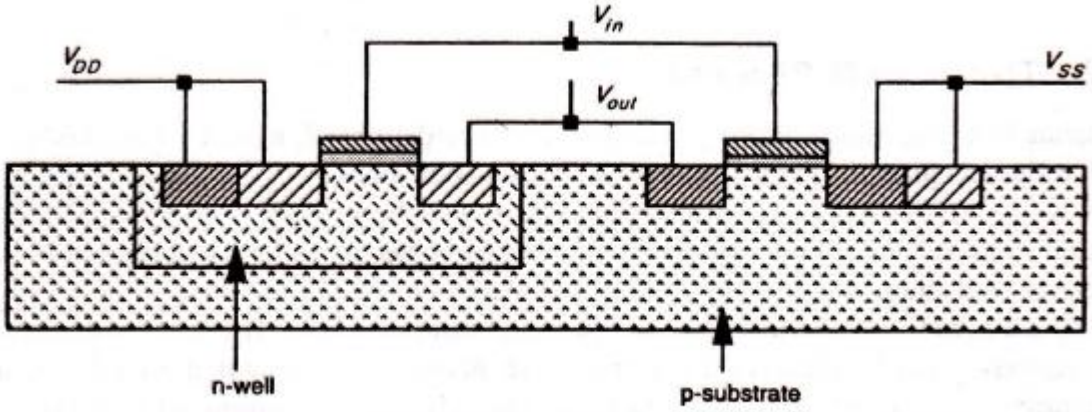


FIGURE 1.12 Cross-sectional view of n-well CMOS inverter.

Owing to differences in charge carrier mobilities, the n-well process creates non-optimum p-channel characteristics. However, in many CMOS designs (such as domino-logic and dynamic-logic structures), this is relatively unimportant since they contain a preponderance of n-channel devices. Thus the n-channel transistors are mainly those used to form logic elements, providing speed and high density of elements.

Latch-up problems can be considerably reduced by using a low-resistivity epitaxial p-type substrate as the starting material, which can subsequently act as a very low resistance ground-plane to collect substrate currents.

However, a factor of the n-well process is that the performance of the already poorly performing p-transistor is even further degraded. Modern process lines have come to grips with these problems, and good device performance may be achieved for both p-well and n-well fabrication.

The Berkeley n-well process

There are a number of p-well and n-well fabrication processes and, in order to look more closely at typical fabrication steps, we will use the Berkeley n-well process as an example. This process is illustrated in Figure 1.13.

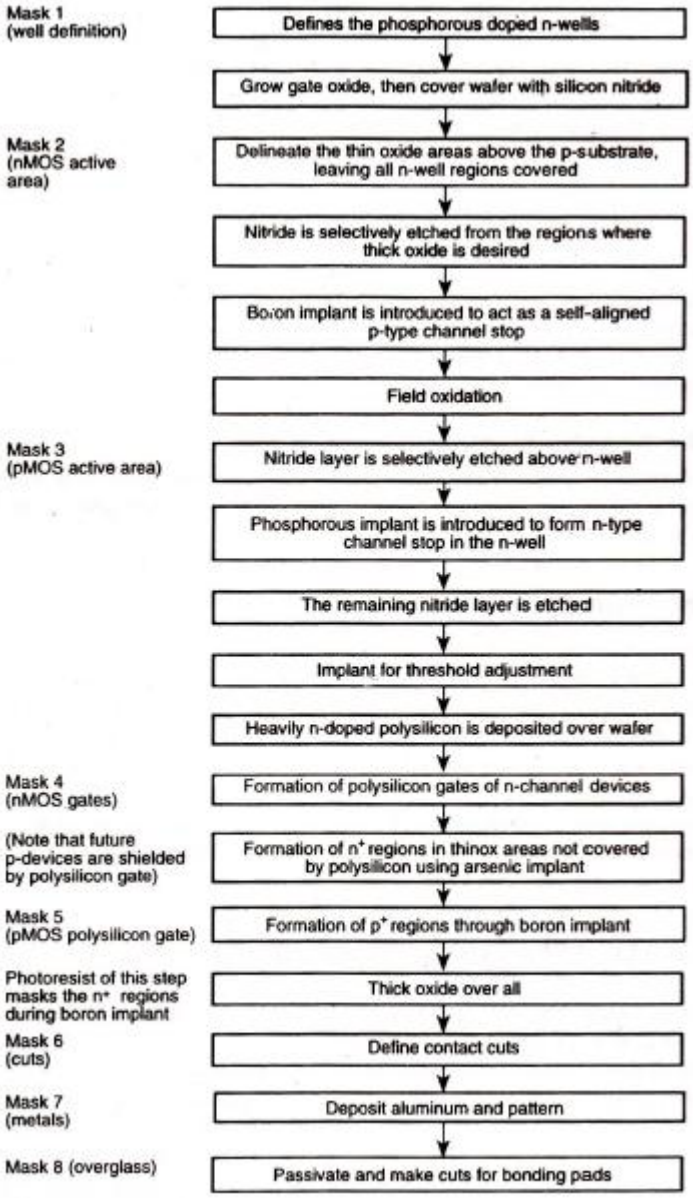


FIGURE 1.13 Flow diagram of Berkeley n-well fabrication.

The Twin-Tub Process

A logical extension of the p-well and n-well approaches is the twin-tub fabrication process.

Here we start with a substrate of high resistivity n-type material and then create both n-well and p-well regions. Through this process it is possible to preserve the performance of n-transistors without compromising the p-transistors. Doping control is more readily achieved and some relaxation in manufacturing tolerances results. This is particularly important as far as latch-up is concerned.

In general, the twin-tub process allows separate optimization of the n- and p-transistors. The arrangement of an inverter is illustrated in Figure 1.14, which may, in turn, be compared with Figures 1.10 and 1.12.

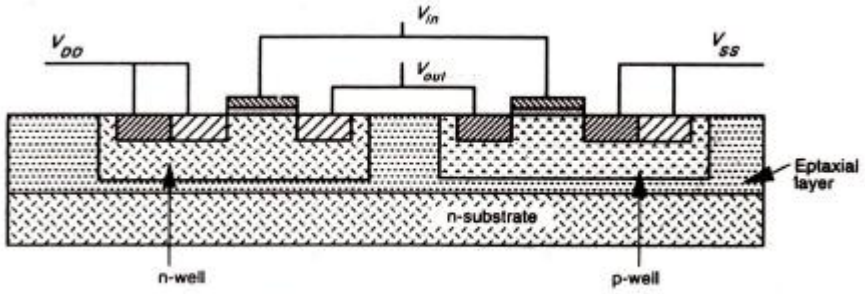


FIGURE 1.14 Twin-tub structure.

THERMAL ASPECTS OF PROCESSING

The processes involved in making nMOS and CMOS devices have differing high temperature sequences as indicated in Figure 1.15.

The CMOS p-well process, for example, has a high temperature p-well diffusion process (1100 to 1250°C), the nMOS process having no such requirement. Because of the simplicity, ease of fabrication, and high density per unit area of nMOS circuits, many of the earlier IC designs, still in current use, have been fabricated using nMOS technology and it is likely that nMOS and CMOS system designs will continue to co-exist for some time to come.

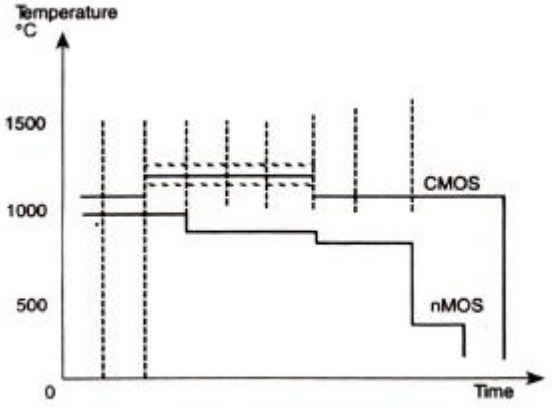


FIGURE 1.15 Thermal sequence difference between nMOS and CMOS processes.

BiCMOS TECHNOLOGY

A known deficiency of MOS technology lies in the limited load driving capabilities of MOS transistors. This is due to the limited current sourcing and current sinking abilities associated with both p- and n-transistors and although it is possible, for example, to design so called super-buffers using MOS transistors alone, such arrangements do not always compare well with the capabilities of bipolar transistors. Bipolar transistors also provide higher gain and have generally better noise and high frequency characteristics than MOS transistors and it may be seen (Figure 1.2) that BiCMOS gates could be an effective way of speeding up VLSI circuits. However, the application of BiCMOS in sub-systems such as ALU, ROM, a register-file, or, for that matter, a barrel shifter, is not always an effective way of improving speed. This is because most gates in such structures do not have to drive large capacitive loads so that the BiCMOS arrangements give no speed advantage. To take advantage of BiCMOS, the whole functional entity, not just the logic gates, must be considered. A comparison between the characteristics of CMOS and bipolar circuits is set out in Table 1.2 and the

differences are self-evident. BiCMOS technology goes some way towards combining the virtues of both technologies.

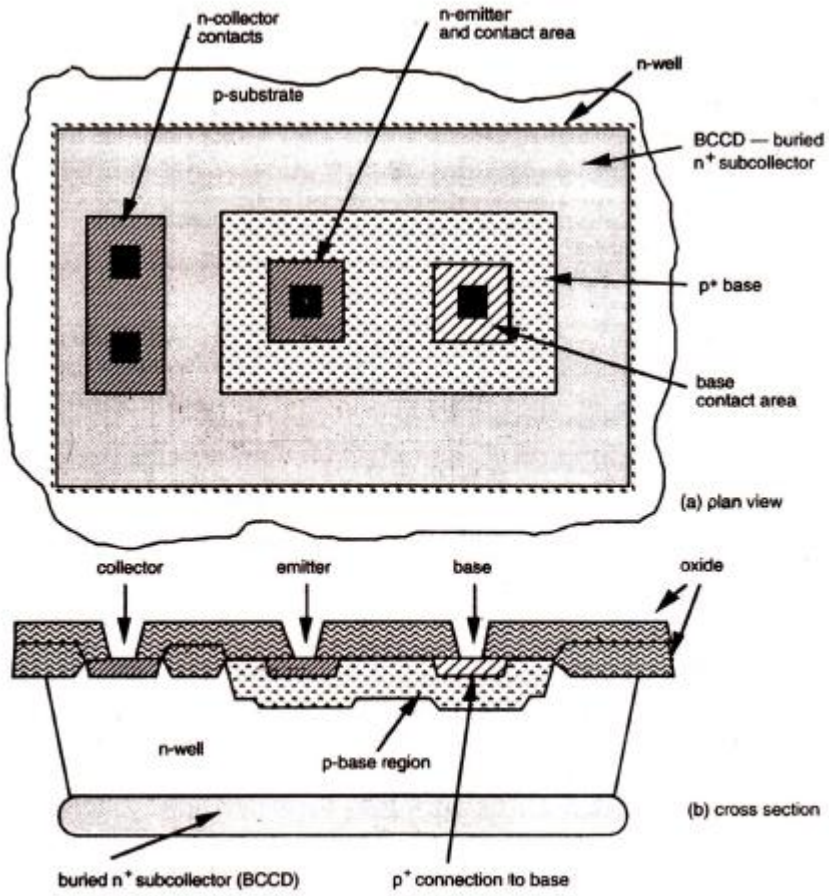
When considering CMOS technology, it becomes apparent that theoretically there should be little difficulty in extending the fabrication processes to include bipolar as well as MOS transistors. Indeed, a problem of p-well and n-well CMOS processing is that parasitic bipolar transistors are inadvertantly formed as part of the outcome of fabrication. The production of npn bipolar transistors with good performance characteristics can be achieved, for example, by extending the standard n-well CMOS processing to include further masks to add two additional layers—the n^+ subcollector and p^+ base layers. The npn transistor is formed in an n-well and the additional p^+ base region is located in the well to form the p-base region of the transistor. The second additional layer—the buried n^+ subcollector (BCCD) is added to reduce the n-well (collector) resistance and thus improve the quality of the bipolar transistor. The simplified general arrangement of such a bipolar npn transistor may be appreciated with regard to Figure 1.16.

TABLE 1.2 Comparison between CMOS and bipolar technologies

CMOS technology	*Bipolar technology*
• Low static power dissipation	• High power dissipation
• High input impedance (low drive current)	• Low input impedance (high drive current)
• Scalable threshold voltage	
• High noise margin	• Low voltage swing logic
• High packing density	• Low packing density
• High delay sensitivity to load (fan-out limitations)	• Low delay sensitivity to load
• Low output drive current	• High output drive current
• Low g_m ($g_m \propto V_{in}$)	• High g_m ($g_m \propto e^{V_{in}}$)
	• High f_t at low currents
• Bidirectional capability (drain and source are interchangeable)	• Essentially unidirectional
• A near ideal switching device	

BiCMOS Fabrication in an n-well Process

The basic process steps used are those already outlined for CMOS but with additional process steps and additional masks defining: (i) the p^+ base region; (ii) n^+ collector area; and (iii) the buried subcollector (BCCD).



Note: For clarity, the layers have not been drawn transparent but BCCD underlies the entire area and the p^+ base underlies all within its boundary.

FIGURE 1.16 Arrangement of BiCMOS npn transistor (Orbit 2 μm CMOS).

Some Aspects of Bipolar and CMOS Devices

Clearly there are relative advantages and disadvantages when comparing bipolar technology with CMOS technology. A readily assimilated comparison of some key features was set out as Table 1.2.

TABLE 1.3 n-well BiCMOS fabrication process steps

Single poly. single metal CMOS	*Additional steps for bipolar devices*
• Form n-well	
	• Form buried n^+ layer (BCCD)
• Delineate active areas	
• Channel stop	
	• Form deep n^+ collector
• Threshold V_t adjustment	
• Delineate poly./gate areas	
• Form n^+ active areas	
• Form p^+ active areas	
	• Form p^+ base for bipolars
• Define contacts	
• Delineate the metal areas	

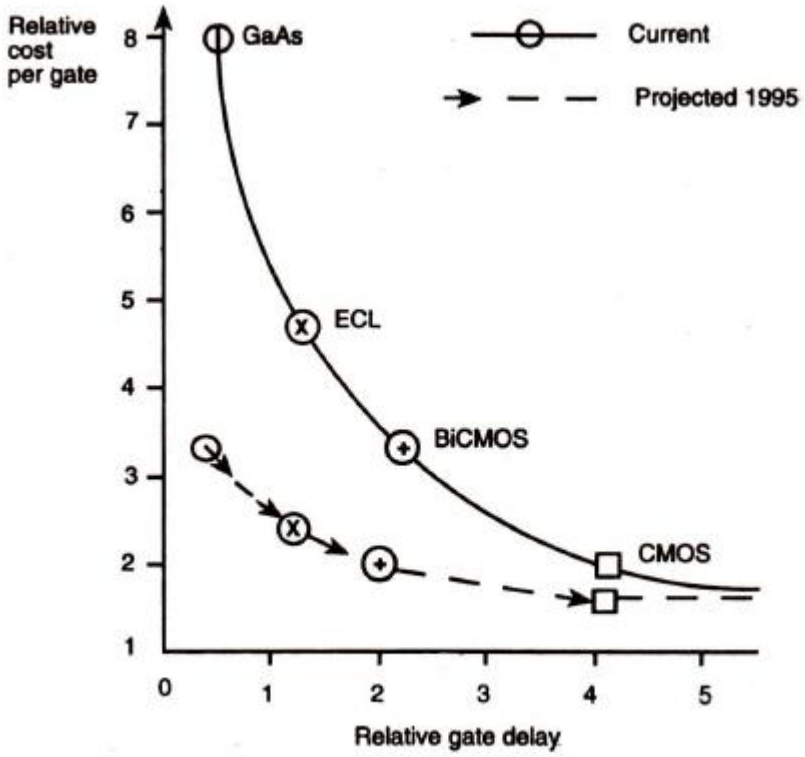


FIGURE 1.17 Cost versus delay for logic gate.

It will be seen that there are several advantages if the properties of CMOS and bipolar technologies could be combined. This is achieved to a significant extent in the BiCMOS technology. As in all things, there is a penalty which, in this case, arises from the additional process steps, some loss of packing density and thus higher cost.

A cost comparison of all current high speed technologies may be assessed from Figure 1.17.

A further advantage which arises from BiCMOS technology is that analog amplifier design is facilitated and improved. High impedance CMOS transistors may be used for the input circuitry while the remaining stages and ouput drivers are realised using Bipolar transistors.

To take maximum advantage of available Silicon technologies one might envisage the following mix of technologies in a silicon system.

CMOS	for logic
BiCMOS	for I/O and driver circuits
ECL	for Critical high speed parts of the system

However, in this text we will not be dealing with the ECL technology.

PRODUCTION OF E-BEAM MASKS

All the processes discussed have made use of masks at various stages of fabrication. In many processes, the masks are produced by standard optical techniques and much has been written on the photolithographic processes involved. However, as geometric dimensions shrink and also to allow for the processing of a number of different chip designs on a single wafer, other techniques are evolving. One popular process used for this purpose uses an E-beam machine. A rough outline of this type of mask making follows:

1. The starting material consists of chrome-plated glass plates which are coated with an E-beam sensitive resist.
2. The E-beam machine is loaded with the mask description data (MEBES).
3. Plates are loaded into the E-beam machine, where they are exposed with the patterns specified by the customer's mask data.
4. After exposure to the E-beam, the plates are introduced into a developer to bring out the patterns left by the E-beam in the resist coating.
5. The cycle is followed by a bake cycle and a plasma de-summing, which removes the resist residue.
6. The chrome is then etched and the plate is stripped of the remaining E-beam resist.

The advantages of E-beam masks are:

- tighter layer to layer registration;
- smaller feature sizes.

There are two approaches to the design of E-beam machines:

- raster scanning;
- vector scanning.

In the first case, the electron beam scans all possible locations (in a similar fashion to a television display) and a bit map is used to turn the E-beam on and off depending on whether the particular location being scanned is to be exposed or not.

For vector scanning, the beam is directed only to those locations which are to be exposed. Although this is inherently faster, the data handling involved is more complex.

INTRODUCTION

We are going through a period of micro-electronic revolution. For a common person, the role of electronics is limited to audio-visual gadgets like radio and television, but the truth is, today the growth of any industry like communication, control, instrumentation or computer, is dependent upon electronics to a great extent. And integrated circuits are electronics.

The integrated circuit or IC is a miniature, low cost electronic circuit consisting of active and passive components that are irreparably joined together on a single crystal chip of silicon. Most of the components used in ICs are not similar to conventional components in appearance although they perform similar electrical functions. In this chapter, we describe the basic processes used in the fabrication of integrated circuits. Both bipolar and MOS fabrication are treated. These circuits naturally offer a number of distinct advantages over those made by interconnecting discrete components. These may be listed as follows:

1. Miniaturization and hence increased equipment density
2. Cost reduction due to batch processing
3. Increased system reliability due to elimination of soldered joints
4. Improved functional performance (as it is possible to fabricate even complex circuits for better characteristics)
5. Matched devices
6. Increased operating speeds (due to the absence of parasitic capacitance effect)
7. Reduction in power consumption.

CLASSIFICATION

Integrated circuits offer a wide range of applications and could be broadly classified as:

Digital ICs

Linear ICs

Based upon the above requirements, two distinctly different IC technology namely, Monolithic technology and Hybrid technology have been developed.

In monolithic integrated circuits, all circuit components, both active and passive elements and their interconnections are manufactured into or on top of a single chip of silicon. The monolithic circuit is ideal for applications where identical circuits are required in very large quantities and hence provides lowest per-unit cost and highest order of reliability. In hybrid circuits, separate component parts are attached to a ceramic substrate and interconnected by means of either metallization pattern or wire bonds. This technology is more adaptable to small quantity custom circuits.

BASIC PLANAR PROCESSES

The basic processes used to fabricate ICs using silicon planar technology can be categorised as follows:

1. Silicon wafer (substrate) preparation
2. Epitaxial growth
3. Oxidation
4. Photolithography
5. Diffusion
6. Ion implantation
7. Isolation technique
8. Metallization
9. Assembly processing and packaging

Silicon Wafer Preparation

The following steps are used in the preparation of Si-wafers

1. Crystal growth and doping
2. Ingot trimming and grinding
3. Ingot slicing
4. Wafer polishing and etching
5. Wafer cleaning

The starting material for crystal growth is highly purified (99.99999) polycrystalline silicon. The Czochralski crystal growth process is the most often used for producing single crystal silicon ingots. The polycrystalline silicon together with an appropriate amount of dopant is put in a quartz crucible and is then placed in a furnace. The material is then heated to a temperature in excess of the silicon melting point of 1420°C. A small single crystal rod of silicon called a seed crystal is then dipped into the silicon-melt and slowly pulled out as shown in Fig. 1.5. As the seed crystal is pulled out of the melt, it brings with it a solidified mass of silicon with the same crystalline structure as that of seed crystal. During the crystal pulling process, the seed crystal and the crucible are rotated in opposite directions in order to produce ingots of circular cross-section. The diameter of the ingot is controlled by the pulling rate and the melt temperature. Ingot diameter of about 10 to 15 cm is common and ingot length is generally of the order of 100 cm.

Next the top and bottom portions of the ingot are cut off and the ingot surface is ground to produce an exact diameter (D = 10, 12.5, 15 cm). The ingot is also ground flat slightly along the length to get a reference plane. The ingot is then sliced using a stainless

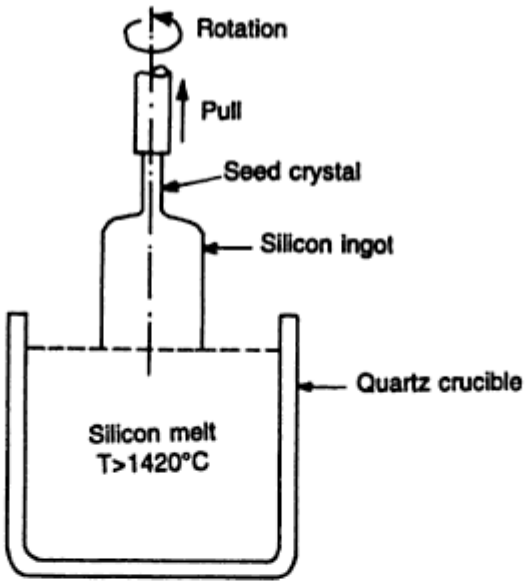


Fig. 1.5 Czochralski crystal growth

steel saw blade with industrial diamonds embedded into the inner diameter cutting edge. This produces circular wafers or slices as shown in Fig. 1.6. The orientation flat portion serves as a useful reference plane for the various processes described later. The silicon wafers so obtained have very rough surface due to slicing operation. These wafers undergo a number of polishing steps to produce a flat surface. Then one side of the wafer is given a final mirror-smooth highly polished finish, whereas the other side is simply lapped on an abrasive lapping machine to obtain an acceptable degree of flatness. Finally, the wafers are thoroughly rinsed and dried. A raw cut slice of thickness 23–40 mils produces wafers of 16–32 mils thickness after all the polishing steps.

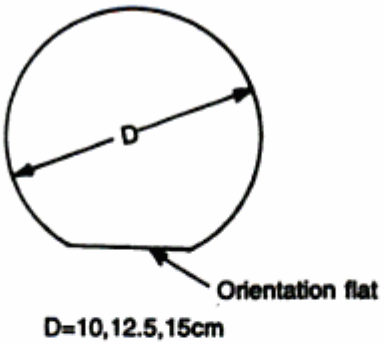


Fig. 1.6 Silicon wafer, D = 10, 12.5, 15 cm showing flat orientation

These silicon wafers will contain several hundred rectangular chips, each one containing a complete integrated circuit. After all the IC fabrication processes are complete, these wafers are sawed into 100 to 8000 rectangular chips having side of 10 to 1 mm. Each chip is a single IC and may contain hundreds of components. The wafer thickness therefore is so chosen that it is possible to separate chips without breaking and at the same time, it gives sufficient mechanical strength to the IC chip.

Epitaxial Growth

The word epitaxy is derived from Greek word *epi* meaning 'upon' and the past tense of the word *teinon* meaning 'arranged'. So, one could describe epitaxy as, arranging atoms in single crystal fashion upon a single crystal substrate, so that the resulting layer is an extension of the substrate crystal structure.

The basic chemical reaction used for the epitaxial growth of pure silicon is the hydrogen reduction of silicon tetrachloride.

$$SiC_4 + 2H_2 \underset{}{\overset{1200°C}{\rightleftharpoons}} Si + 4HCl$$

Mostly, epitaxial films with specific impurity concentration are required. This is accomplished by introducing phosphine (PH_3) for the *n*-type and bi-borane (B_2H_6) for *p*-type doping into the silicon-tetrachloride hydrogen gas stream.

The process is carried out in a reaction chamber consisting of a long cylindrical quartz tube encircled by an RF induction coil. Figure 1.7 shows the diagrammatic representation of the system used. The silicon wafers are placed on a rectangular graphite rod called a boat. This boat is then placed in the reaction chamber where the graphite is heated inductively to a temperature 1200°C. The various gases required for the growth of desired epitaxial layers are introduced into the system through a control console.

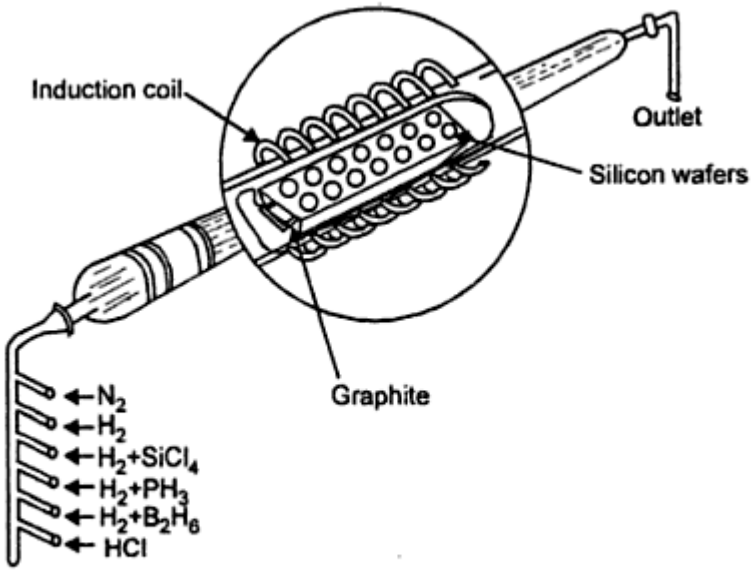


Fig. 1.7 A diagrammatic representation of a system for growing silicon epitaxial films

Oxidation

SiO_2 has the property of preventing the diffusion of almost all impurities through it. It serves two very important purposes.

1. SiO_2 is an extremely hard protective coating and is unaffected by almost all reagents except hydrofluoric acid. Thus it stands against any contamination.
2. By selective etching of SiO_2, diffusion of impurities through carefully defined windows in the SiO_2 can be accomplished to fabricate various components.

The silicon wafers are stacked up in a quartz boat and then inserted into quartz furnace tube. The Si-wafers are raised to a high temperature in the range of 950 to 1115°C and at the same time, exposed to a gas containing O_2 or H_2O or both. The chemical reaction is

$$Si + 2H_2O \longrightarrow SiO_2 + 2H_2$$

This oxidation process is called thermal oxidation because high temperature is used to grow the oxide layer. The thickness of the film is governed by time, temperature and the moisture content. The thickness of oxide layer is usually in the order of 0.02 to 2 μm.

Photolithography

With the help of photolithography, it has become possible to produce microscopically small circuit and device patterns on Si-wafers. As many as 10,000 transistors can be fabricated on a 1 cm × 1 cm chip. The

Photolithography

With the help of photolithography, it has become possible to produce microscopically small circuit and device patterns on Si-wafers. As many as 10,000 transistors can be fabricated on a 1 cm × 1 cm chip. The conventional photolithographic process uses ultraviolet light exposure and device dimension or line width as small as 2 μm can be obtained. However, with the advent of latest technology using X-ray or electron beam lithographic techniques, it has become possible to produce device dimension down to submicron range (< 1 μm).

Photolithography involves two processes, namely:
Making of a photographic mask
Photo etching

The making of a photographic mask involves the following sequence of operations—first the preparation of initial artwork and secondly, its reduction. The initial layout or artwork of an IC is normally done at a scale several hundred times larger than the final dimensions of the finished monolithic circuit. This is because, for a tiny chip, larger the artwork, more accurate is the final mask. For example, it is often required to make an opening of width about 1 mil (25 μm). Obviously, this cannot be managed by any draftsman even with his thinnest of sketch pens. So the drawings are made magnified and often by a factor of 500. With this magnification, it is easy to see that a width of one mil is magnified to a width of 500 mils, that is, about 1.2 cm. Therefore, for a finished monolithic chip of area 50 mil × 50 mil, the artwork will be made on an area of about 60 cm × 60 cm.

This initial layout is then decomposed into several mask layers, each corresponding to a process step in the fabrication schedule, e.g., a mask for base diffusion, another for collector diffusion, another for metallization and so on.

For photographic purpose, artwork should not contain any line drawings but must be of alternate clear and opaque regions. This is accomplished by the use of clear Mylar coated with a sheet of red photographically opaque mylar (trade name-Rubylith). The red layer can be easily peeled off thus exposing clear areas with a knife edge from the regions where impurities have to be diffused. The artwork is usually produced on a precision drafting machine, known as coordinatograph. The coordinatograph has a cutting head that can be positioned accurately and moved along two perpendicular axes. The coordinatograph outlines the pattern cutting through the red mylar without damaging the clear layer underneath.

This rubylith pattern of individual mask is photographed and then reduced in steps by a factor of 5 or 10 several times to finally obtain the exact image size. The final image also must be repeated many times in a matrix array, so that many ICs will be produced in one process. The photo repeating is done with a step and repeat camera. This is an imaging device with a photographic plate on a movable platform. Between exposure, the plate is moved in equal steps so that successive images form in an array. When the exposed plate is developed, it becomes a master mask. The masks, actually used in IC processing are made by contact printing from the master. These working masks wear out with use and are replaced as required.

Photo-etching is used for the removal of SiO_2 from desired regions so that the desired impurities can be diffused. The wafer is coated with a film of photosensitive emulsion (Kodak Photoresist KPR). The thickness of the film is in the range of 5000–10000 Å as shown in Fig. 1.8 (a). The mask negative of the desired pattern) as prepared by steps described earlier is placed over the photoresist coated wafer as shown in Fig. 1.8 (b). This is now exposed to ultraviolet light, so that KPR becomes polymerized beneath the transparent regions of the mask. The mask is then removed and the wafer is developed using a chemical (trichloroethylene) which dissolves the unexposed/ unpolymerized regions on the photoresist and leaves the pattern as shown in Fig. 1.8 (c). The polymerised photoresist is next fixed or cured, so that it becomes immune to certain chemicals called etchants used in subsequent processing steps. The chip is immersed in the etching solution of hydrofluoric acid, which removes the SiO_2 from the areas which are not protected by KPR as shown in Fig. 1.8 (d) After diffusion of impurities, the photoresist is removed with a chemical solvent (hot H_2SO_4) and mechanical abrasion.

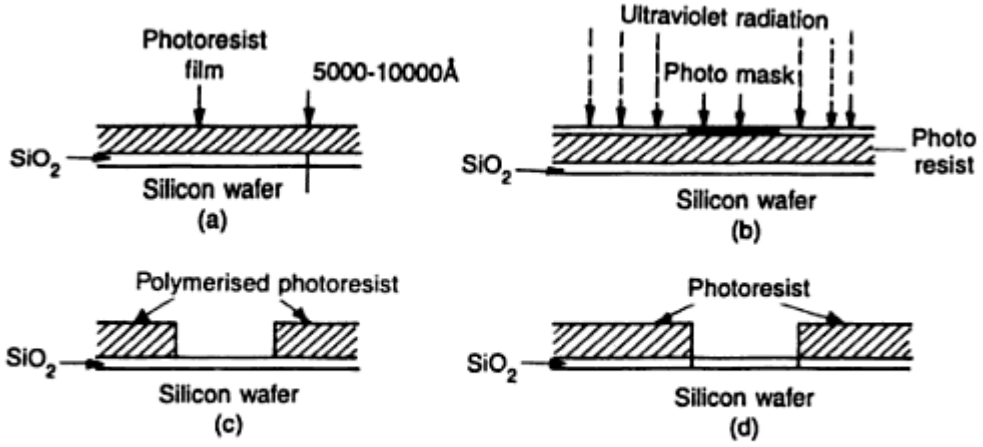


Fig. 1.8 Various steps for photo-etching

The etching process described is a wet etching process and the chemical reagents used are in liquid form. A new process used these days is a dry etching process called plasma etching. A major advantage of the dry etching process is that it is possible to achieve smaller line openings (≤ 1 μm) compared to wet process.

X-Ray and Electron Beam Lithography

With conventional ultraviolet (UV) photolithography process in which the UV wavelengths used are in the range 0.3 to 0.4 μm, the minimum device dimensions or line widths are limited by diffraction effects to around five wavelengths or about 2 μm. This is what puts an upper limit on the IC device density using UV photolithography.

With the advent of X-ray and electron beam lithography techniques, it has become possible to produce device dimensions down to submicron range (< 1 μm). This is due to much shorter wavelengths involved. With these techniques, MOSFET with gate length as small as 0.25 μm have been made. The cost of X-ray or electron beam equipment is very high and the exposure times very much longer than with UV photolithography. So this becomes a very expensive process and is used only when very small device dimensions (≤ 1 μm) are needed.

Diffusion

Another important process in the fabrication of monolithic ICs is the diffusion of impurities in the Silicon chip. This uses a high temperature furnace having a flat temperature profile over a useful length (about 20″ length). A quartz boat containing about 20 cleaned wafers is pushed into the hot zone with temperature maintained at about a 1000°C. Impurities to be diffused are rarely used in their elemental forms. Normally, compounds such as B_2O_3 (Boron oxide), BCl_3 (Boron chloride) are used for Boron and P_2O_5 (Phosphorous pentaoxide) and $POCl_3$ (Phosphorous oxychloride) are used as sources of Phosphorous. A carrier gas, such as dry oxygen or nitrogen is normally used for sweeping the impurity to the high temperature zone. The depth of diffusion depends upon the time of diffusion which normally extends to 2 hours.

The diffusion of impurities normally takes place both laterally as well as vertically. Therefore, the actual junction profiles will be curved as shown in Fig. 1.9.

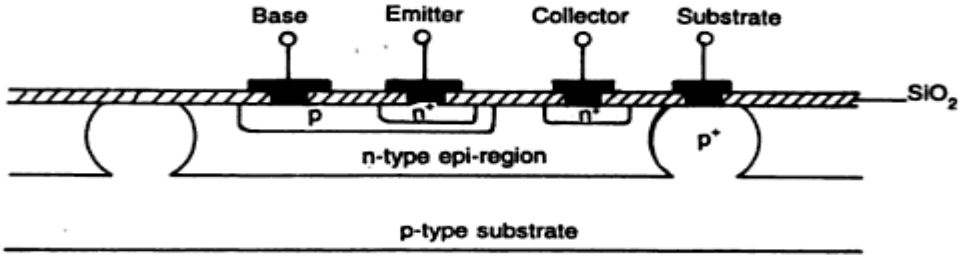


Fig. 1.9 The cross-section of an *npn* transistor showing curved junction profiles as a result of lateral diffusion

Ion Implatation

Ion implantation is the other technique used to introduce impurities into a silicon wafer. In this process, silicon wafers are placed in a vacuum chamber and are scanned by a beam of high-energy dopant ions (borons for *p*-type and phosphorus for *n*-type) as shown in Fig. 1.10. These ions are accelerated by energies between 20 kV to 250 kV.

As the ions strike the silicon wafers, they penetrate some small distance into the wafer. The depth of penetration of any particular type of ion increases with increasing accelerating voltage. Ion implantation technique has two important advantages.

1. It is performed at low temperatures. Therefore, previously diffused regions have a lesser tendency for lateral spreading.
2. In diffusion process, temperature has to be controlled over a large area inside the oven, whereas in ion implantation technique, accelerating potential and the beam current are electrically controlled from outside.

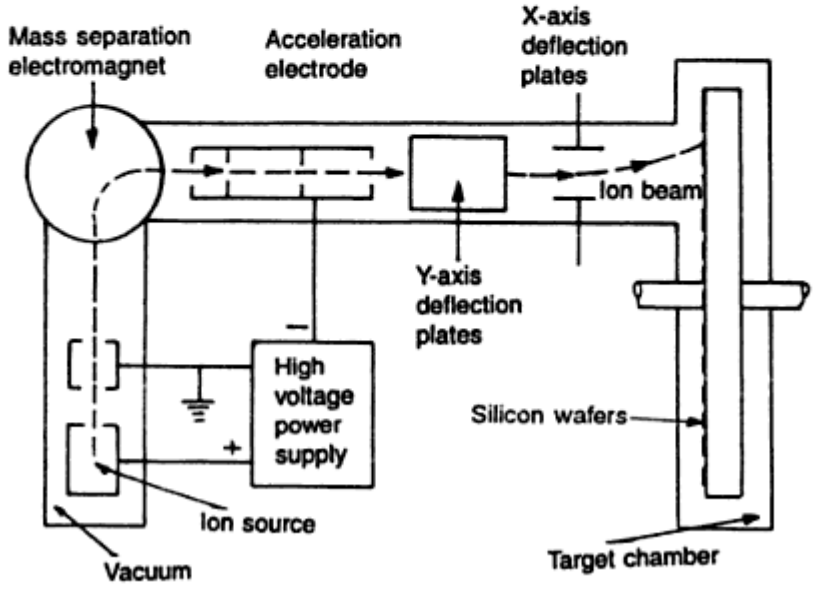


Fig. 1.10 Ion implantation system

Isolation Techniques

Since a number of components are fabricated on the same IC chip, it becomes necessary to provide electrical isolation between different components and interconnections. Various types of isolation techniques have been developed. However, we shall discuss here only two commonly used techniques namely:

pn junction isolation
Dielectric isolation

p-n Junction Isolation

In this isolation technique, p^+ type impurities are selectively diffused into the *n*-type epitaxial layer so as to reach *p*-type substrate as shown in Fig. 1.11 (a). This produces islands surrounded by *p*-type moats. It can be seen that these regions are separated by two back-to-back *p-n* junction diodes. If the *p*-type substrate material is held at the most negative potential in the circuit, the diodes will be reverse

biased providing electric isolation between these islands. The different components are fabricated in these isolation islands. The concentration of the acceptor atoms in the region between isolation islands is usually kept much higher (p^+) than the *p*-type substrate. This prevents the depletion region of the reverse biased diode from penetrating more into p^+ region and possibly connecting the isolation islands.

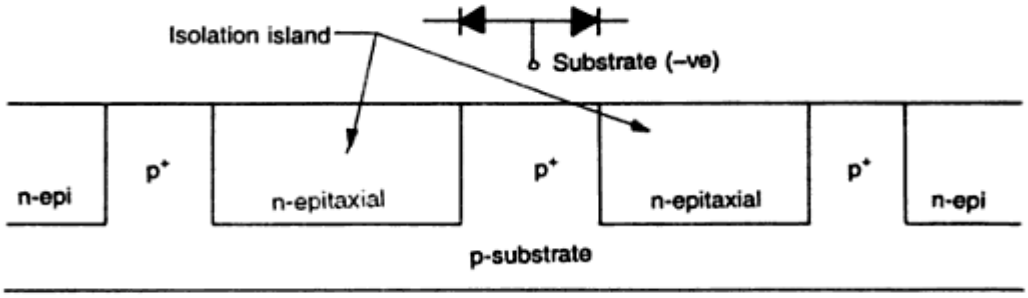


Fig. 1.11 (a) *p-n* junction isolation

There is, however, one undesirable by-product of this isolation process. It is the presence of a transition capacitance at the isolating *pn* junctions, resulting in an inevitable capacitor coupling between the components and the substrate. These parasitic capacitances limit the performance of the circuit at high frequencies. But being economical, this technique is commonly used for general purpose ICs.

Dielectric Isolation

Here a layer of solid dielectric such as silicon dioxide or ruby completely surrounds each component, thereby producing isolation, both electrical and physical. This isolating dielectric layer is thick enough so that its associated capacitance is negligible. Also, it is possible to fabricate both *pnp* and *npn* transistor within the same silicon substrate. Since this method requires additional fabrication steps, it becomes more expensive. The technique is mostly used for fabricating professional grade ICs required for specialised applications viz, aerospace and military, where higher cost is justified by superior performance. Figure 1.11 (b) shows such a structure.

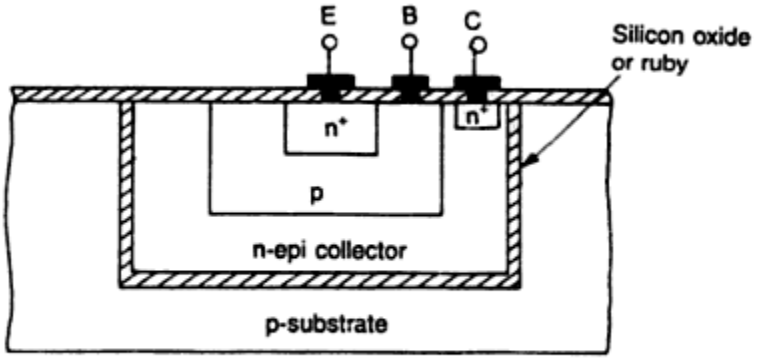


Fig. 1.11 (b) Dielectric isolation

Metallization

The purpose of this process is to produce a thin metal film layer that will serve to make interconnections of the various components on the chip. Aluminium is usually used for the metallization of most ICs as it offers several advantages.

1. It is relatively a good conductor.
2. It is easy to deposit aluminium films using vacuum deposition.
3. Aluminium makes good mechanical bonds with silicon.
4. Aluminium forms low resistance, non-rectifying (i.e. ohmic) contact with *p*-type silicon and the heavily doped *n*-type silicon.

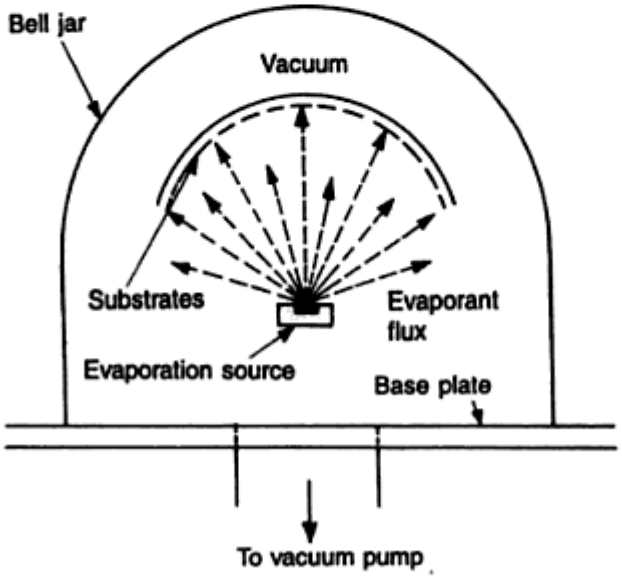


Fig. 1.12 Vacuum evaporation for metallization

The film thickness of about 1 μm and conduction width of about 2 to 25 μm are commonly used. The process takes place in a vacuum evaporation chamber as shown in Fig. 1.12. The pressure in the chamber is reduced to the range of about 10^{-6} to 10^{-7} torr (1 atmosphere = 760 torr = 760 mm Hg). The material to be evaporated is placed in a resistance heated tungsten coil or basket. A very high power density electron beam is focussed at the surface of the material to be evaporated. This heats up the material to very high temperature and it starts vaporizing. These vapors travel in straight line paths. The evaporated molecules hit the substrate and condense there to form a thin film coating.

After the thin film metallization is done, the film is patterned to produce the required interconnections and bonding pad configuration. This is done by photolithographic process and aluminium is etched away from unwanted places by using etchants like phosphoric acid (H_3PO_4).

Assembly Processing and Packaging

Each of the wafer processed contains several hundred chips, each being a complete circuit. So these chips must be separated and individually packaged. A common method called scribing and cleaving used for separation makes use of a diamond tipped tool to cut lines into the surface of the wafer along the rectangular grid separating the individual chips. Then the wafer is fractured along the scribe lines and the individual chips are physically separated. Each chip is then mounted on a ceramic wafer and attached to a suitable package.

There are three different package configurations available.

1. TO-5 glass metal package
2. Ceramic flat package
3. Dual-in-line (ceramic or plastic type)

TO-5 packages are available in 8, 10 or 12 leads, whereas the flat or dual-in-line package is commonly available in 8, 14 or 16 leads, but

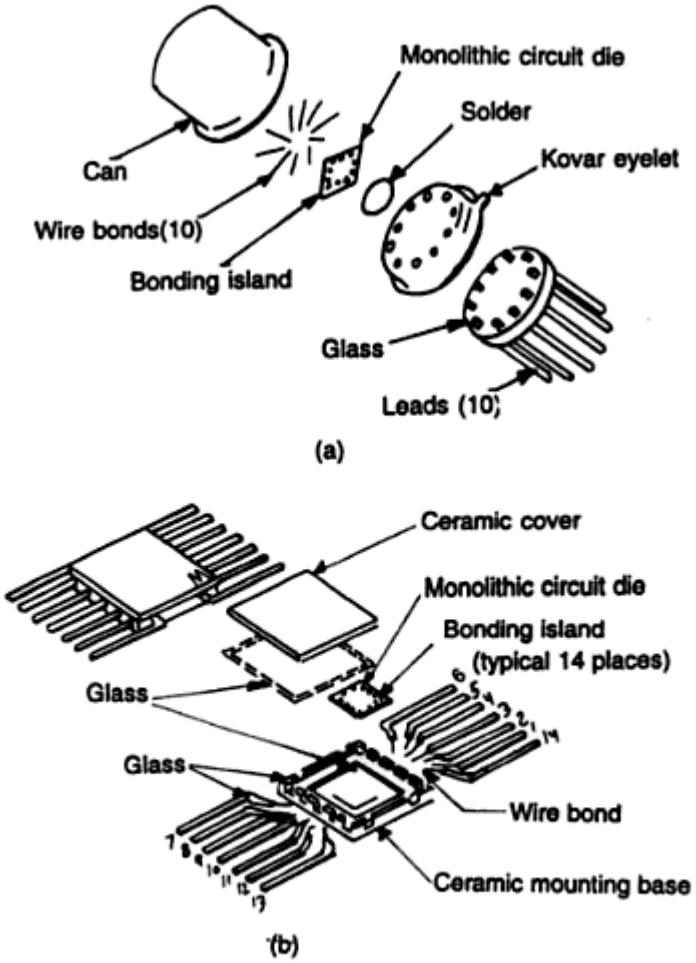


Fig. 1.13 Exploded view of (a) lead TO-5 package (b) 14-lead version of the flat package, showing the various components as well as the completed flat package

even 24 or 36 or 42 leads are also available for special circuits. Ceramic packages, whether of flat type or dual-in-line are costly due to fabrication process, but have the advantage of best hermetic sealing. Most of the general purpose ICs are dual-in-line plastic packages due to economy. Figure 1.13 (a, b) shows the exploded view of TO-5 and flat package.

Integrated Resistors

The basic technique for obtaining a resistor in integrated circuit is by utilising the bulk resistance of a defined volume of semicondector region. Four different methods are available for fabricating integrated resistors, namely:

Diffused Resistor
Epitaxial Resistor
Pinched Resistor
Thin Film Resistor

Diffused Resistor

In this method, resistor is formed in one of the isolated regions of epitaxial layer during base or emitter diffusion common to bipolar transistor fabrication. As no extra fabricating steps are required, this type of resistor is very economical. However, a severe limitation is the small range of resistances possible.

The value of the resistance depends upon the surface geometry. That is, length, width and upon the diffused impurity profile. In this context, a very useful quantity sheet resistance is defined as diffused layers are very thin.

Sheet Resistance R_S

Consider the square $L \times L$ of a material of resistivity ρ, thickness t, and cross-sectional area $A = L \times t$ shown in Fig. 1.22 (a). The resistance of this sheet of meterial can be written as

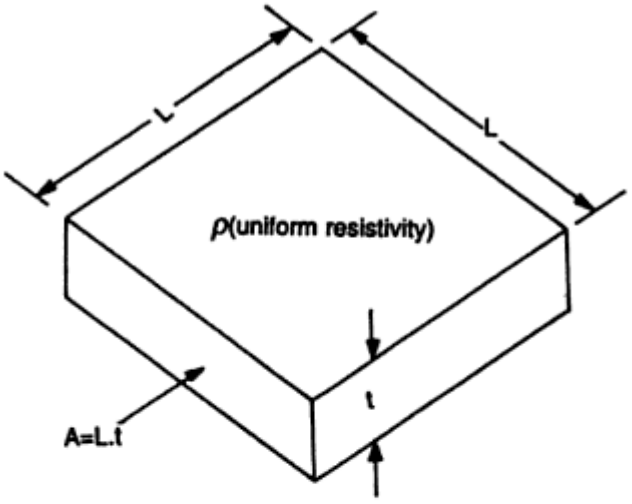
$$R_s = \frac{\rho L}{L \times t} = \frac{\rho}{t} \text{ (ohms per square)}$$

This quantity R_s is independent of the size of the square and mainly depends upon the diffusion characteristic of the material. Now consider Fig. 1.22 (b, c) which shows a base resistor and the emitter resistor. The resistance for these resistors can be expressed in terms of the sheet resistance R_s and the surface dimensions L and W.

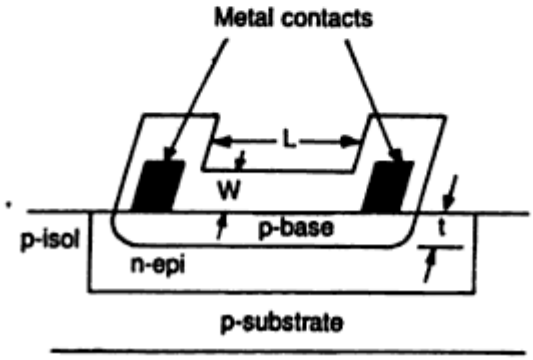
Now $$R = \rho \frac{L}{W \times t}$$

or, $$R = R_s \frac{L}{W}$$

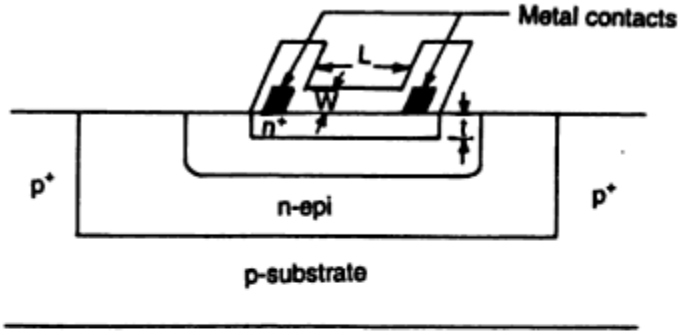
where the ratio L/W is called the aspect ratio of the surface geometry and is, therefore, the effective number of square contained in the resistor. The base resistor in the range of 20 Ω to 300 kΩ can be easily fabricated due to medium resistivity (200 Ω/square) p-type base region. However, the sheet resistance of the emitter diffusion is of the order of 5 Ω/sq. only. So emitter resistors are usually in the range of 10 to 1 kΩ.



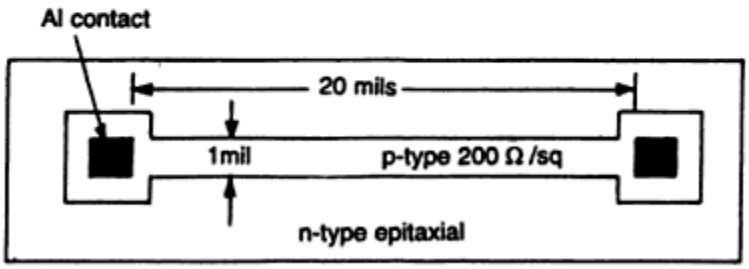
(a)



(b)



(c)



(d)

Fig. 1.22 (a) Sheet resistance (b) Base resistor (c) Emitter resistor (d) Top view showing dimensions for a 4000 Ω diffused resistor

Example

Design a 4 kΩ diffused resistor.

The sheet resistance of p-type diffusion is 200 ohm/sq.

Then $$\frac{L}{W} = \frac{R}{R_s} = \frac{4 \times 10^3}{200} = \frac{20}{1}$$

So a 4 kΩ resistor can be fabricated by using a pattern of 20 mils long by 1 mil wide as shown by the top view in Fig. 1.22 (d).

Epitaxial Resistor

Large value of resistance than possible by base or emitter resistor can be achieved by using n-epitaxial collector region as shown in Fig. 1.23 (a). Sheet resistance of epitaxial layer in the order of 1 to 10 kΩ/ square can be obtained.

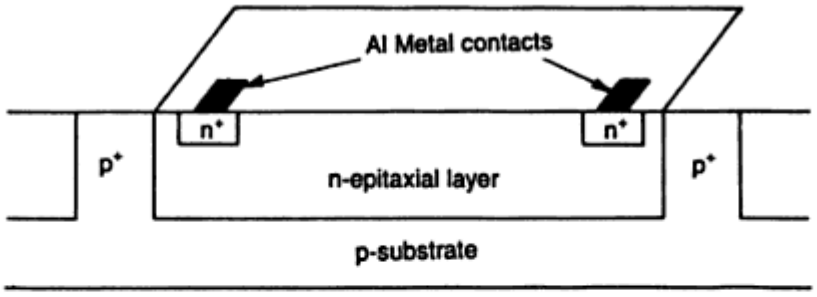


Fig. 1.23 (a) Epitaxial resistor

Pinched Resistor

The sheet resistivity of a semiconductor region can be increased by reducing its effective cross-sectional area. In a pinched resistor, this technique is used to achieve a high value of sheet resistance from the ordinary base diffused resistor.

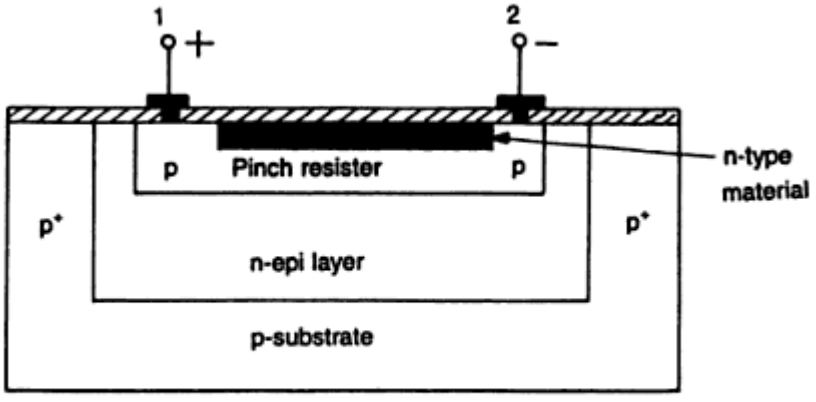


Fig. 1.23 (b) Cross-sectional view of a pinch resistor

A pinched base diffused resistor is shown in Fig. 1.23 (b). It can give resistances of the order of mega-ohm in a reasonably small area. In this structure, no current can flow through the *n*-type material (dark region) due to the diode at contact 2 in the reverse direction. Only a small reverse saturation current can flow through *n*-material. So, by creating this *n*-type region the effective cross-sectional area for the conduction path has been reduced and thus resistance between 1 and 2 increases.

Thin Film Resistor

Vapour thin film deposition techniques as discussed in Sec. 1.5.8 can also be used for the fabrication of IC resistors. In this, a very thin metallic film usually of Nichrome (NiCr) of thickness less than 1 μm is vapour deposited on the SiO_2 layer. Using masked etching, desired geometry of this thin film is achieved to obtain suitable values of resistors. The ohmic contacts are made using Al metallization and usual masked etching techniques. Nichrome resistors are available with typical sheet resistance values of 40 to 400 Ω/square depending upon film thickness, so that resistance in the range of 20 to 50 kΩ can be easily obtained. Figure 1.23 (c) shows the cross-sectional view of such a resistor. These thin film resistors have three distinct advantages over the diffused resistors.

1. Thin film resistors have lesser and smaller parasitic components and hence their high frequency behaviour is better.
2. The values of thin-film resistors can be easily adjusted even after fabrication by cutting a part of the resistor with a laser beam (Laser trimming).
3. Thin film resistors have low temperature coefficient, thereby making them more stable.

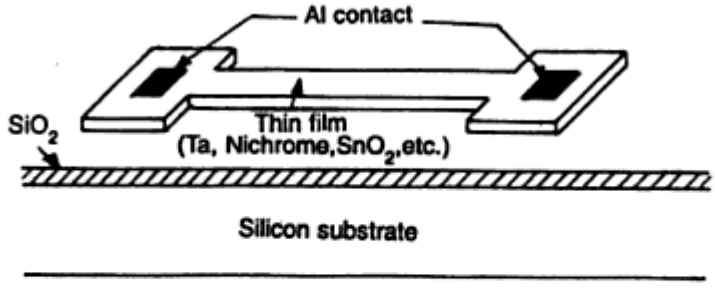


Fig. 1.23 (c) Cross-section of a thin film resistor

Higher values of thin film resistors have been obtained by depositing Tantalum over SiO_2 layer. The main disadvantage of thin film resistors is the additional process steps required in their fabrication.

Integrated Capacitors

Two commonly used methods for obtaining integrated capacitors are as follows:

(1) Junction capacitor
(2) MOS and thin film capacitor

Junction Capacitor

It is possible to use the junction capacitance of a reverse biased diode as an element in monolithic ICs. Figure 1.24 shows the cross-sectional view of a junction monolithic capacitor and its equivalent circuit. It can be seen that there are two junctions in this type of diffused capacitor. Junction J_2, if reverse biased, will produce the desired capacitance. However, a parasitic capacitance C_1 is inevitable due to the junction J_1 between n-type epitaxial layer and the substrate. Also a series resistance results due to the bulk resistance of the n-region. In the equivalent circuit, two diodes are the idealized diodes of the two junctions. The substrate must be held at the most negative point in the circuit, to minimize C_1. The value of the capacitance C_2 will depend upon the area of the junction, impurity concentration of the n-type epitaxial layer and the voltage across the junction. It can be seen that the capacitor C_2 is polarised and is obtained only when the junction J_2 is reverse biased.

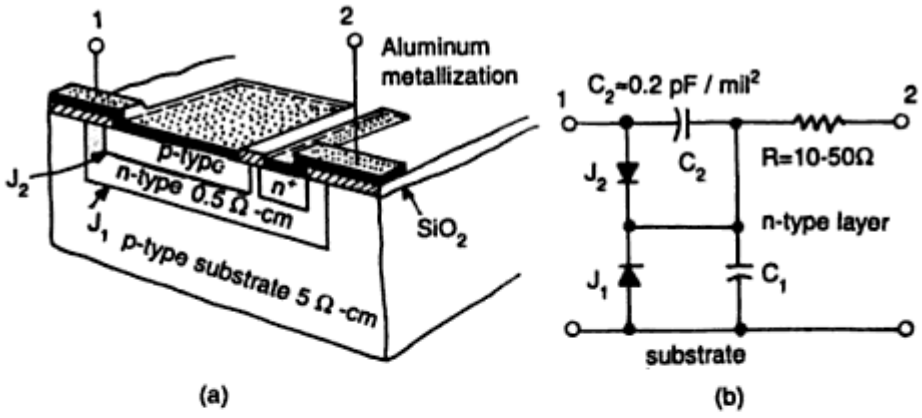


Fig. 1.24 (a) Junction-type IC capacitor, (b) Equivalent circuit

MOS and Thin Film Capacitor

A commonly used capacitor is the metal-oxide semiconductor capacitor for which the cross-sectional view and equivalent circuit is shown in Fig. 1.25. It is basically a parallel plate capacitor with SiO_2 as the dielectric. The heavily doped n^+ region formed during the emitter diffusion forms the lower plate and the thin film of aluminium metallization forms the upper plate of the capacitor with SiO_2 as the

dielectric. As shown in the equivalent circuit, the parasitic effects consist of a small series resistance R due to n^+ region, a collector substrate junction J_1, and its associated capacitance C_1.

This type of capacitor has the advantage of being nonpolar, that is, it does not matter which of the plates is positive or negative. Nor it is a function of the voltage applied, thus giving more circuit flexibility Silicon Nitride (Si_3N_4) can also be used as dielectric layer. Silicon Nitride offers a higher value of capacitance because of higher dielectric constant value, but requires extra processing steps.

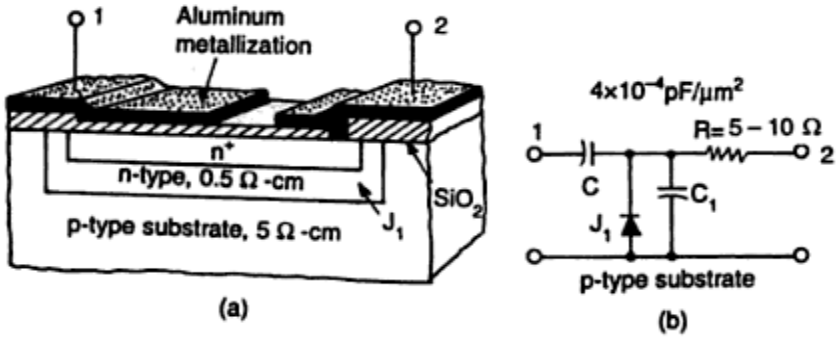


Fig. 1.25 (a) Structure and (b) equivalent circuit of a MOS capacitor

Thin film capacitor structures which use thin dielectric film layer between two metal layers are also in use. Although such a capacitor structure is almost free from substrate parasitics, it requires a number of additional masking and deposition steps beyond the basic MOS structure. In such a structure, either aluminium or tantalum is used as capacitor plates and aluminium oxide (Al_2O_3) or tantalum oxide (Ta_2O_5) as the dielectric material. Ta_2O_5 is particularly preferred for large value capacitors. One basic serious disadvantage of thin film capacitor is that they fail when the voltage rating exceeds due to breakdown of the dielectric. This is a destructive and irreversible failure mechanism and may require over-voltage protection.

Integrated Inductors

Inductors, transformers and chokes are the missing link in the chain of IC components. IC devices are essentially two dimensional as the depth dimension is usually very small (~ 1 to 10 μm) compared to the lateral dimensions. IC inductors can be made in the form of a flat metallic thin film spirals by successive deposition of conduction patterns. Very small values of inductance of the order of nano-henry with low quality factor can be only obtained. For any reasonable inductance value, a three dimensional coil structure is needed to obtain a large number of turns.

Most circuit designers go to great lengths to avoid the use of inductors or otherwise simulate them by using RC active networks. In applications such as RF and IF circuits, where inductors cannot be avoided, inductors external to the IC-package are used. However, in thin-film hybrid microwave integrated circuit (MIC), thin film inductor spirals are used giving upto 250 nH.