1

BASIC MOS TECHNOLOGY AND MOS TRANSISTOR THEORY

1.1 INTRODUCTION

Very-Large-Scale Integration (VLSI) Design forms a modular methodology of structuring the challenges in semiconductor designs. VLSI is the process of creating integrated circuits by combining thousands of transistors into a single chip. The concept of integration provides higher speed, physically less area occupancy and relatively consuming less power.

1.2 INTEGRATED CIRCUITS ERA

A remarkable change in the history began with the invention of transistor by William Shockley, John Bardeen and Walter Brattain, in 1947 in bell labs. Figure 1.1 gives the picture of first transistor. This made a path for the development of Integrated Circuits (IC). Figure 1.2 gives the picture of the first integrated circuit developed by Jack Kilby of Texas Instruments (TI) in 1958 in which more than one transistor was fabricated in the single piece of semiconductor material. This idea of integration emerged to the beginning of commercial IC during 1960.

Figure 1.1: First Transistor
(Image Courtesy of Bell Labs)

The first integrated circuits contained only a few transistors, called "**Small-Scale Integration**" (**SSI**). The next step in the development of integrated circuits, introduced devices which contained hundreds of transistors on each chip, called "**Medium-Scale Integration**" (**MSI**) during late 1960s. Further

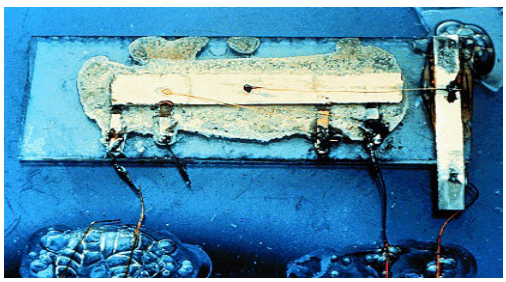

Figure 1.2: First TI-Integrated Circuit
(Image Courtesy of Texas Instruments, Inc.)

development, driven by the same economic factors, led to "**Large-Scale Integration**" (**LSI**) in the mid 1970s, with tens of thousands of transistors per chip. The future step in the development process, starting in the 1980s and continuing through the present, is **"Very Large-Scale Integration" (VLSI)**. The development started with hundreds of thousands of transistors and is still continuing at a greater scale with multiple developments.

Since then large scale improvement in the development and design is observed in VLSI. This rapid growth of technology in VLSI was forecasted by Gordon Moore, co-founder of Intel. His prediction in the year 1965, popularly known as "Moore's Law," states that, ***"the number of transistors on integrated circuits doubles about every two years"***. This is shown with reference to the microprocessor transistor counts in Figure 1.3.

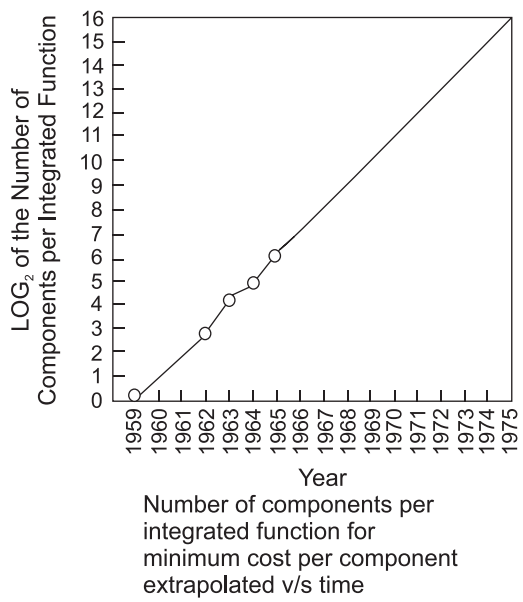

Figure 1.3: Demonstration of Moore's Law (Image Courtesy of Intel)

Today, technology continues driving Moore's Law to increase functionality and performance and decrease cost, bringing growth to industries

worldwide. A negative implication of Moore's Law is obsolescence., that is, as technologies continue to rapidly "improve", these improvements can be significant enough to rapidly render predecessor technologies obsolete.

Silicon is the most popular material which has wide range of cost performance trade-offs. As technology scales down, new design challenges in **timing (performance), power** and **area** are to be met in order to design an efficient and effective chip. Sophisticated Computer-Aided Design (CAD) tools and methodologies are developed and applied in order to manage the rapidly increasing design complexity.

VLSI engineering encompasses several distinct areas of specialization that mesh together in a unique manner. VLSI design is viewed at many different levels, from abstract to the physical implementation. Each level may have subdivisions. A generalized VLSI design flow is discussed in Figure 1.4.

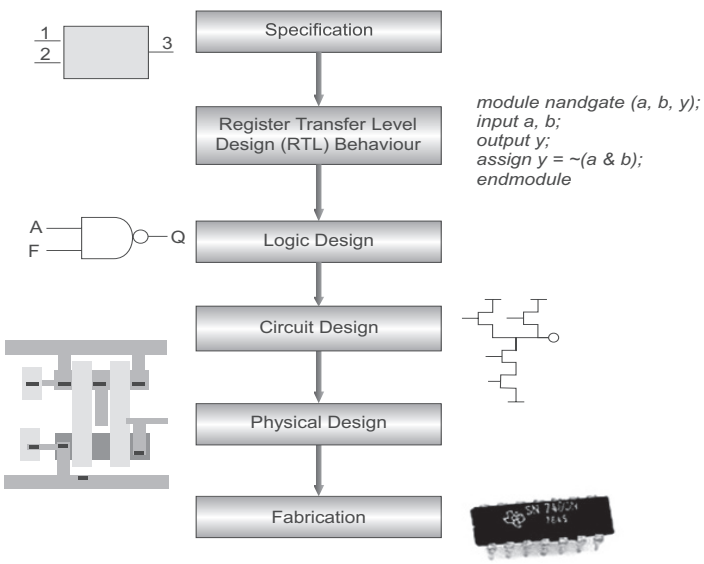

Figure 1.4: VLSI Design Flow

Based on the specification, the functional behaviour of the block is determined by applying a set of excitation vectors. A Hardware Description Language (HDL) code is written to describe the behaviour of the module. The design is tested through a simulation process; it is to check, verify, and ensure that what is wanted is what is described. Simulation is carried out through dedicated tools. With every simulation run, the simulation results are studied to identify errors in the design description. The errors are corrected and another simulation

run carried out. The resulting Logic Design is obtained with the automation tool. The next level involves the formation of schematics for the module, with the transistor gate level simulation of the logic. Functional check, Timing checks and the Power analysis checks are done. This will be the Circuit Design level in the design flow. After the schematic is simulated and verified, the corresponding mask layers of the circuit are created. The layout design includes the floor planning, placement and routing. The locations of each schematic component are decided in the floor planning and are placed accordingly. While routing, the interconnections are done between the components. This step is the Physical Design. The design is verified and tested. Failure, if any in the design is identified in each stage and corrected, to avoid the re-design at later stages. Finally the routed netlist will be sent to the foundry and the chip will be fabricated as per the technology requirement.

1.3 ENHANCEMENT AND DEPLETION MODE MOS TRANSISTORS

MOS Transistors are built on a silicon substrate. Silicon which is a group IV material is the eighth most common element in the universe by mass, but very rarely occurs as the pure free element in nature. It is most widely distributed in dusts, sands, planetoids, and planets as various forms of silicon dioxide (silica) or silicates. It forms crystal lattice with bonds to four neighbours. Silicon is a semiconductor. Pure silicon has no free carriers and conducts poorly. But adding dopants to silicon increases its conductivity. If a group V material i.e. an extra electron is added, it forms an n-type semiconductor. If a group III material i.e. missing electron pattern is formed (hole), the resulting semiconductor is called a p-type semiconductor. A junction between p-type and n-type semiconductor forms a conduction path. Source and Drain of the Metal Oxide Semiconductor (MOS) Transistor is formed by the "doped" regions on the surface of chip. Oxide layer is formed by means of deposition of the silicon dioxide (SiO_2) layer which forms as an insulator and is a very thin pattern. Gate of the MOS transistor is the thin layer of "polysilicon (poly)"; used to apply electric field to the surface of silicon between Drain and Source, to form a "channel" of electrons or holes. Control by the Gate voltage is achieved by modulating the conductivity of the semiconductor region just below the gate. This region is known as the channel.

The Metal–Oxide–Semiconductor Field Effect Transistor (MOSFET) is a transistor which is a voltage-controlled current device, in which current at two electrodes, drain and source is controlled by the action of an electric field

at another electrode gate having in-between semiconductor and a very thin metal oxide layer. It is used for amplifying or switching electronic signals.

The Enhancement and Depletion mode MOS transistors are further classified as N-type named NMOS (or N-channel MOS) and P-type named PMOS (or P-channel MOS) devices. Figure 1.5 shows the MOSFETs along with their enhancement and depletion modes.

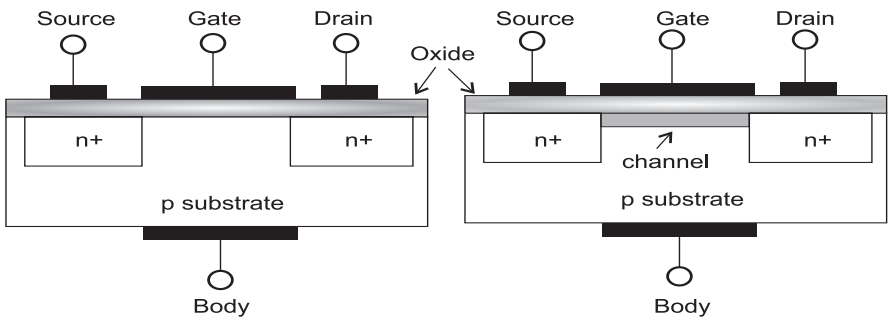

Figure 1.5: (a) Enhancement N-type MOSFET (b) Depletion N-type MOSFET

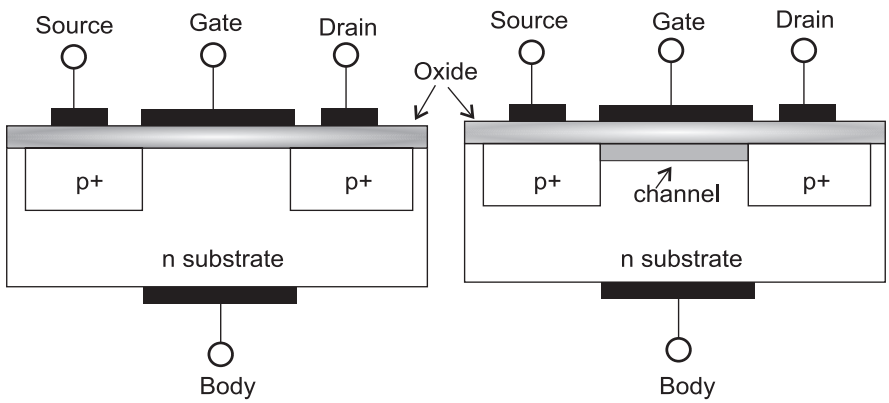

Figure 1.5: (c) Enhancement P-type MOSFET (d) Depletion P-type MOSFET

The depletion mode devices are doped so that a channel exists even with zero voltage from gate to source during manufacturing of the device. Hence the channel always appears in the device. To control the channel, a negative voltage is applied to the gate (for an N-channel device), depleting the channel, which reduces the current flow through the device. In essence, the depletion-mode device is equivalent to a closed (ON) switch, while the enhancement-mode device does not have the built in channel and is equivalent to an open (OFF) switch. Due to the difficulty of turning off the depletion mode devices, they are rarely used.

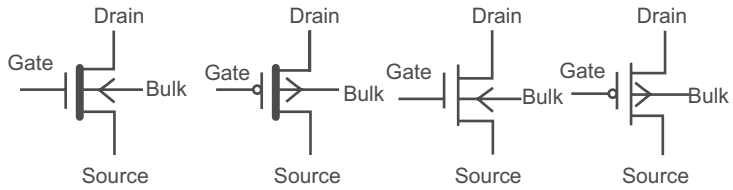

Figure 1.6: (a) N and P Depletion MOSFET's. (b) N and P Enhancement MOSFET's

Figure 1.6 shows the symbolic representation of MOS devices. It has a line indicating the channel with source and drain terminals leaving it at right angles. The gate terminal is drawn parallel to the channel. The bulk (or substrate) terminal will not act as any input or output, but it is connected to the source of the MOSFET. The arrow shown represents the direction of current flow. The thickness at the gate in Figure 1.6 shows the in-built channel for the depletion mode devices.

Working of Enhancement Mode Transistor

The enhancement mode devices do not have the in-built channel. By applying the required potentials, the channel can be formed. Also for the MOS devices, there is a threshold voltage (V_t), below which not enough charges will be attracted for the channel to be formed. This threshold voltage for a MOS transistor is a function of doping levels and thickness of the oxide layer.

Case 1: $V_{gs} = 0V$ and $V_{gs} < V_t$

The device is non-conducting, when no gate voltage is applied ($V_{gs} = 0V$) or ($V_{gs} < V_t$) and also drain to source potential $V_{ds} = 0$. With an insufficient voltage on the gate to establish the channel region as N-type, there will be no conduction between the source and drain. Since there is no conducting channel, there is no current drawn, i.e. $I_{ds} = 0$, and the device is said to be in the **cut-off region**. This is shown in the Figure 1.7 (a).

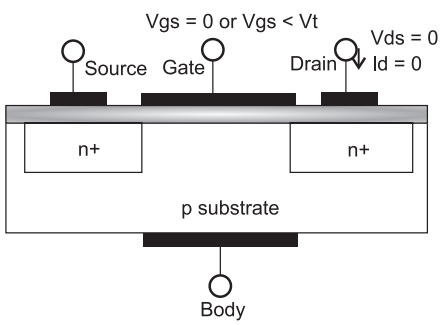

Figure 1.7: (a) Cut-off Region

Case 2: Vgs > Vt

When a minixmum voltage greater than the threshold voltage V_t (i.e. $V_{gs} > V_t$) is applied, a high concentration of negative charge carriers forms an inversion layer located by a thin layer next to the interface between the semiconductor and the oxide insulator. This forms a channel between the source and drain of the transistor. This is shown in the Figure 1.7 (b).

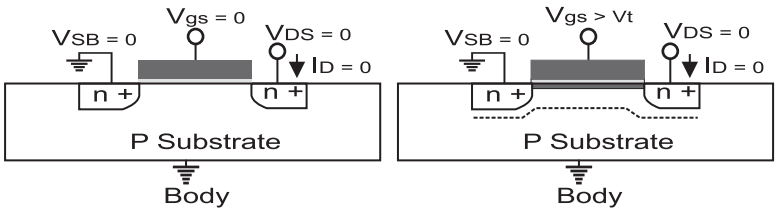

Figure 1.7: (b) Formation of a Channel

A positive V_{ds} reverse biases the drain substrate junction, hence the depletion region around the drain widens, and since the drain is adjacent to the gate edge, the depletion region widens in the channel. This is shown in Figure 1.7 (c). This results in flow of electron from source to drain resulting in current Ids. The device is said to operate in **linear region** during this phase. Further increase in $V_{ds,}$ increases the reverse bias on the drain substrate junction in contact with the inversion layer which causes inversion layer density to decrease. This is shown in Figure 1.7 (d). The point at which the inversion layer density becomes very small (nearly zero) at the drain end is termed pinch-off. The value of V_{ds} at pinch-off is denoted as $V_{ds,sat}$. This is termed as **saturation region** for the MOS device. Diffusion current completes the path from source to drain in this case, causing the channel to exhibit a high resistance and behaves as a constant current source.

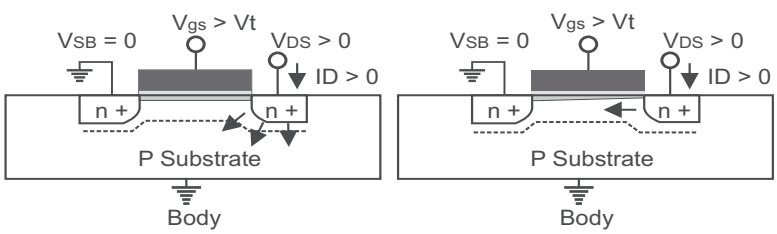

Figure 1.7: (c) Linear Region. (d) Saturation Region

The MOSFET I_D versus V_{DS} characteristics (V-I Characteristics) is shown in the Figure 1.8. For $V_{GS} < V_{t,}$ $I_D = 0$ and device is in cut-off region. As V_{DS} increases at a fixed V_{GS}, I_D increases in the linear region due to the increased lateral field, but at a decreasing rate since the inversion layer density is decreasing. Once

pinch-off is reached, further increase in V_{DS} results in increase in I_D; due to the formation of the high field region which is very small. The device starts in linear region, and moves into saturation region at higher V_{DS}.

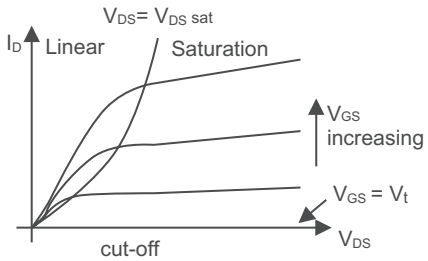

Figure 1.8: MOS V-I Characteristics

1.4 BASICS OF FABRICATION

Foundry (also called a fab for fabrication plant) is used to refer to a factory where devices like integrated circuits (IC) are manufactured. The central part of a fab is a cleanroom. A cleanroom is an environment, typically used in manufacturing or scientific research that has a low level of environmental pollutants such as dust, airborne microbes, aerosol particles and chemical vapours. More accurately, a cleanroom has a controlled level of contamination that is specified by the number of particles per cubic meter at a specified particle size.

The basic semiconductor material used in device fabrication is **Silicon.** Silicon is an abundant material, which occurs naturally in the form of sand. It can be refined using well established techniques for purification and crystal growth. It is the widely chosen material for semiconductor device manufacturing, because of its abundant availability, wide operating temperature range, good electrical characteristics, cost effectiveness, and ease for fabrication. Also silicon can be made an excellent insulator, (SiO_2 or glass) by oxidizing it. This native oxide can be used to form capacitors and MOSFETs. This oxide also serves as the barrier that can mask diffusion of unwanted impurities into nearby high-purity silicon material. This property allows the electrical property of silicon to be easily altered in predefined areas. Hence active and passive elements can be built on same piece of material (substrate). The components are then interconnected using metal layers to form the single piece of material called the integrated circuit (IC). It is also relatively easy process and reliable high volume fabrication. Other semiconductors such as gallium arsenide (GaAs) are used for special applications.

The semiconductor surface is subjected to various processing steps in which impurities and other materials are added with specific geometrical patterns. The fabrication steps are sequenced to form three dimensional regions that act as transistors and interconnect that form the network. The MOS transistors are fabricated on silicon wafer. Lithography process is used which is similar to printing process. On each step, different materials are deposited or patterned or etched on the substrate material. This method is easier to understand by viewing both top and cross-section of wafer in a simplified manufacturing process.

Silicon manufacturing process involves several steps. First, pure silicon is melted in a pot (1400º C) and a small seed containing the desired crystal orientation is inserted into molten silicon and slowly (1mm/minute) pulled out. The silicon crystal (in some cases also containing doping) is manufactured as a cylindrical ingot, with a diameter of 8-12 inches (1″ = 2.54 cm). This ingot is carefully sawed into thin (0.50 mm - 0.75 mm thick) disks called wafers, which are later polished and marked for crystal orientation. Then the process of lithography is done where the patterns are transferred on to the layer of IC. Photoresist application is then done to the surface to be patterned. It is spin-coated with a light-sensitive organic polymer called photoresist. Printing is then done on the mask pattern which is developed on the photoresist, with Ultra Violet (UV) light exposure; depending on the type of photoresist (negative or positive) the exposed or unexposed parts become resistant to certain types of solvents. The soluble photoresists are then chemically removed. The developed photoresist acts as a mask for patterning of underlying layers and then is removed. Oxide can be grown from silicon through heating in an oxidizing atmosphere. The process involved are, gate oxide, and device isolation. Oxidation consumes silicon. SiO_2 is deposited on materials other than silicon through reaction between gaseous silicon compounds and oxidizers. Insulation between different layers of metallization is done. Once the desired shape is patterned with photoresist, the etching process allows unprotected materials to be removed. Two types of etching process are followed, wet etching, which uses chemicals or dry or plasma etching which uses ionized gases. Then, doping materials are added to silicon to change its electrical characteristics. This is done by diffusion, where dopants deposited on silicon move through the lattice by thermal diffusion (high temperature process). This is used to create the well region. The other method is by ion implantation, where highly energized donor or acceptor atoms impinge on the surface and travel below it. Here, the patterned SiO_2 serves as an implantation mask. Source and Drain regions

are created by this method. The next process is the thermal annealing, a high temperature process which allows doping impurities to diffuse further into the bulk and repair lattice damage caused by the collisions with doping ions. The last process is the silicon deposition and ion implantation. Films of silicon can now be added on the surface of a wafer. This can be done by epitaxy, which is the growth of a single-crystal semiconductor film on a crystalline substrate, polysilicon formation, where polycrystalline film with a granular structure is obtained through deposition of silicon on an amorphous material. This will create the MOSFET gates. Metallization is the process for deposition of metal layers by evaporation. Interconnections are done by this process.

1.5 NMOS FABRICATION

The following description explains the basic steps used in the process of fabrication.

(*a*) The fabrication process starts with the oxidation of the silicon substrate. It is shown in the Figure 1.9 (a).

(*b*) A relatively thick silicon dioxide layer, also called field oxide, is created on the surface of the substrate. This is shown in the Figure 1.9 (b).

(*c*) Then, the field oxide is selectively etched to expose the silicon surface on which the MOS transistor will be created. This is indicated in the Figure 1.9 (c).

(*d*) This is followed by covering the surface of substrate with a thin, high-quality oxide layer, which will eventually form the gate oxide of the MOS transistor as illustrated in Figure 1.9 (d).

(*e*) On top of the thin oxide, a layer of polysilicon (polycrystalline silicon) is deposited as is shown in the Figure 1.9 (e). Polysilicon is used both as gate electrode material for MOS transistors and also as an interconnect medium in silicon integrated circuits. Undoped polysilicon has relatively high resistivity. The resistivity of polysilicon can be reduced, however, by doping it with impurity atoms.

(*f*) After deposition, the polysilicon layer is patterned and etched to form the interconnects and the MOS transistor gates. This is shown in Figure 1.9 (f).

(*g*) The thin gate oxide not covered by polysilicon is also etched along, which exposes the bare silicon surface on which the source and drain junctions are to be formed (Figure 1.9 (g)).

(*h*) The entire silicon surface is then doped with high concentration of impurities, either through diffusion or ion implantation (in this case with donor atoms to produce n-type doping). Diffusion is achieved by heating the wafer to a high temperature and passing the gas containing desired impurities over the surface. Figure 1.9 (h) shows that the doping penetrates the exposed areas on the silicon surface, ultimately creating two n-type regions (source and drain junctions) in the p-type substrate. The impurity doping also penetrates the polysilicon on the surface, reducing its resistivity.

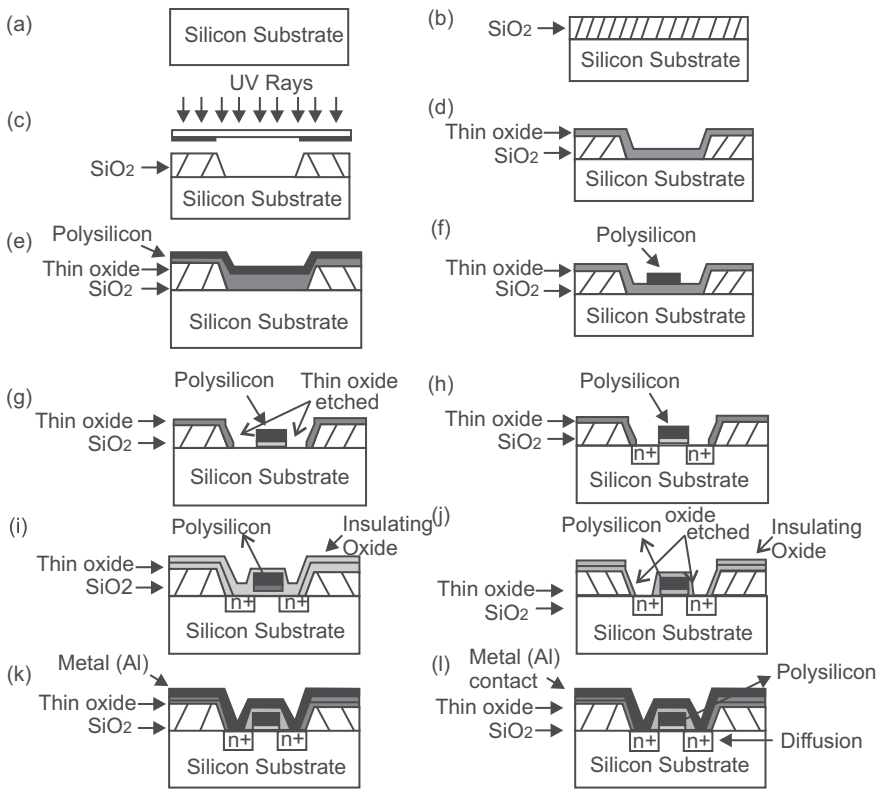

Figure 1.9: Fabrication Process of NMOS Device

(*i*) Once the source and drain regions are completed, the entire surface is again covered with an insulating layer of silicon dioxide, as shown in Figure 1.9 (i).

(*j*) The insulating oxide layer is then patterned in order to provide contact windows for the drain and source junctions, as illustrated in Figure 1.9 (j).