

## **Fabrication process: -**

Integrated Circuits are a collection of electrical circuits that are put on a tiny electronic chip. With the help of the fabrication process, a large number of devices are produced, including transistors, MOSFETs (Metal Oxide Semiconductor Field Effect Transistors), microcontrollers, computer processors, etc. Because of their stable structures, silicon and germanium are the most often used materials in the manufacturing of semiconductors.

Other materials favored in the fabrication sector include wood, thermoplastics, resins, silver, aluminum, magnesium, copper, and carbon steel. Let's first talk about fabrication and the needs it has in the semiconductor industry.

Construction of an industrial product is known as fabrication. It can also be described as a group of procedures used in the production of electronic goods. Silicon semiconductor chips, as an illustration, etc. Fabrication is a procedure used to turn raw materials into finished goods in the case of metals.

By using the smaller components created during the fabrication process, larger electrical components can be created.



The following are the fundamental techniques used in fabricating integrated circuits:

- Crystal growth
- Wafer processing
- Masking
- Oxidation
- Etching
- Deposition
- Ion implantation , diffusion and doping
- CVP
- Metallization
- Packaging

## **Manufacture**

Semiconductor devices have long been used in electronics. Semiconductor device fabrication is the process used to manufacture semiconductor devices, typically integrated circuits (ICs) such as computer processors, microcontrollers, and memory chips (such as NAND flash and DRAM) that are present in everyday electrical and electronic devices. Electronic circuits are gradually built on a wafer, normally formed of pure single-crystal semiconducting material, using a multi-step photolithographic and physio-chemical process (including phases like thermal oxidation, thin-film deposition, ion-implantation, and etching). Silicon is almost always used, but various compound semiconductors are used for specialized applications.

## **Silicon Manufacture**

In this chapter, we will consider only silicon-based (Si) technologies. Although other compound materials in groups III through V, such as gallium arsenide (GaAs) and aluminum gallium nitride (AlGaN), are also used to produce VLSI chips, silicon is still the most popular material, with excellent cost–performance trade-off. The mineral quartz consists entirely of silicon dioxide, also known as silica. Ordinary sand is chiefly composed of tiny grains of quartz and is there for also silica.

Despite the abundance of its components, elemental silicon does not exist in nature. The element can be synthesized artificially by heating silica and carbon in an electric boiler. The carbon reacts with the oxygen in the silica, resulting in nearly pure molten

silicon. As this cools, a slew of minute crystals forms and coalesces to form a fine-grained grey solid. Because it contains a large number of crystals, this type of silicon is called poly-crystalline. This metallurgical-grade polysilicon is unsuitable for semiconductor manufacturing due to impurities and a disordered crystal structure.

Metallurgical-grade silicon can be further refined to produce an extremely pure semiconductor-grade material. Purification begins with the conversion of the crude silicon into a volatile compound, usually trichlorosilane. After repeated distillation, the extremely pure trichlorosilane is reduced to elemental silicon using hydrogen gas. The final product is exceptionally pure, but still poly-crystalline. Practical integrated circuits can only be fabricated from single-crystal material, so that the next step consists of growing a suitable crystal.

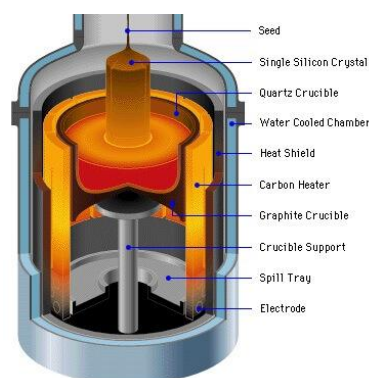
## **Crystal Growth**

The principles of crystal growth are straightforward and well-known. Assume a few sugar crystals are added to a saturated solution that then evaporates. Sugar crystals act as seeds for the deposition of more sugar molecules. The crystals eventually grow to be quite large. Even in the absence of a seed, crystal growth would occur, but the result would be a swarm of small intergrown crystals. By suppressing unwanted nucleation sites, the use of a seed allows the growth of larger, more perfect crystals.

In principle, silicon crystals can be grown in much the same manner as sugar crystals. In practice, no suitable solvent exists for silicon, and the crystals must be grown from the molten element at temperatures in excess of 1400°C. The resulting crystals are at least a meter in length and ten centimeters in diameter, and they must have a nearly perfect crystal structure to be useful to the semiconductor industry. These requirements make the process technically challenging.

The usual method for growing semiconductor-grade silicon crystal is called *Czochralski process*. This process, illustrated in figure 2.1, uses a silica crucible charged with pieces of semi-grade polycrystalline silicon. An electric furnace raises the temperature of the crucible until all the silicon melts. The temperature is then reduced slightly and a small seed crystal is lowered into the crucible. Controlled cooling of the melt causes layers of silicon atoms to deposit upon the seed crystal. The rod holding the seed slowly rises so that only the lower portion of the growing crystal remains in

contact with the molten silicon. In this manner, a large silicon crystal can be pulled centimeter-by-centimeter from the melt. The shaft holding the crystal rotates slowly to ensure uniform growth. The high surface tension of molten silicon distorts the crystal into a cylindrical rod rather than expected faceted prism.



*Figure 2.1 czochralski process for growing silicon crystal.*

The czochralski process requires careful control to provide crystals of the desired purity and dimensions. Automated systems regulate the temperature of the melt and the rate of crystal growth. A small amount of doped polysilicon added to the melt sets the doping concentration in the crystals. In addition to the deliberately introduced impurities, oxygen from the silica crucible and carbon from the heating elements dissolve in the molten silicon and become incorporated into the growing crystal. These impurities subtly influence the electrical properties of the resulting silicon. Once the crystal has reached its final dimensions, it is lifted from the melt and is allowed to slowly cool to room temperature. The resulting cylinder of monocrystalline silicon is called an *ingot*.

Since integrated circuits are formed upon the surface of a silicon crystal and penetrate this surface to no great depth, the ingot is customarily sliced into numerous thin circular sections called wafers. That are 400 $\mu$ m to 600 $\mu$ m thick. Each wafer yields hundreds or even thousands of integrated circuits. The larger the wafer, the more integrated circuits it holds and the greater the resulting economies of scale.

## **Wafer Manufacturing**

A series of mechanical processes are used to create wafers. The ingot's two tapered ends are sliced off and discarded. The remaining material is ground into a cylinder, the

diameter of which determines the size of the resulting wafers. After grinding, no visible indication of crystal orientation remains. The crystal orientation is determined experimentally, and a flat stripe is ground down one side of the ingot. Each wafer cut from it will have a facet, or flat, that clearly identifies the crystal orientation.

After grinding the flat, the manufacturer uses a diamond-tipped saw to cut the ingot into individual wafers. Approximately one-third of the precious silicon crystal is reduced to worthless dust in the process. The sawing process leaves scratches and pockmarks on the surfaces of the resulting wafers. Because the tiny dimensions of integrated circuits necessitate extremely smooth surfaces, each wafer must have one side polished. Mechanical abrasives are used first, followed by chemical milling. The resulting mirror-bright surface has a dark grey color and the characteristic silicon near-metallic luster.

The finished wafers' thickness and diameter must correspond to the mechanical strength and other physical qualities of the material the wafers will be used to make. The finished product must be sturdy enough to support its weight without shattering when handling the products of the wafers' diverse applications. As more material is added in the fabrication of the slice, the diameter of the wafer increases, as does the weight of the wafer. When enough weight has been added, the diameter cannot be increased because it compromises the slice's strength.

If the weight is not properly placed, a small amount of pressure will be enough to break down the wafer. Before they can be used in manufacturing, sliced wafers must be processed. Machines and abrasive chemicals are used to smooth out the rough surface of the wafer. The flawless surface facilitates printing circuit layouts on the wafer surface.

All silicon wafers are extremely useful components, just that there are different variations used for different purposes. It is therefore imperative to know about the different types of Silicon Wafers. There are mainly two varieties of silicon wafers commonly used today. These are undoped silicon wafer and Doped silicon wafer. Undoped Silicon Wafers, also known as Intrinsic or Float Zone (FZ) do not have any dopants in them. They are made of strictly pure crystalline silicon. This type of silicon wafer is recognized as the ideal semiconductor. Doped silicon wafers are formed by

introducing dopants (certain impurities) into the silicon crystal during the formation process. When boron is added into the mixture, a P-type doped silicon wafer is produced. P-type silicon wafers have numerous positively-charged holes. To produce an N-type doped silicon wafer, elements like phosphorus, arsenic, or antimony will be added. N-type silicon wafers have a negatively-charged electron in them. The quantity of dopant discovered in the wafer will ascertain if it is degenerate or extrinsic. To be degenerate means there's a higher concentration of dopants in it, while to be extrinsic means it has little or moderate dopants.

## **Photolithography**

photolithography is used in all structuring process steps. Its purpose is to transfer a two-dimensional image of the required structures onto the wafer surface, so that subsequent processing (e.g. implantation, etching) can be applied to a restricted area.

The wafer is first coated with a thin radiation sensitive film, called a *photoresist* or *resist*. A photomask is then exposed to light to project a black and white image of the desired structure on the photoresist. The solubility of the photoresist changes with reference to. a specific fluid, called a *developer*, at the regions exposed to light. In the *developing process*, the soluble areas of the photoresist are removed from the wafer with this fluid. The insoluble parts of the photoresist remain.

## **Photoresist**

Photoresist is deposited on the wafer by means of spin coating. Liquid coating material is applied to the center of the wafer, which is rapidly rotated. The centrifugal force causes the fluid to be distributed on the wafer. The solvent for setting the viscosity in the fluid is vaporized to produce a layer of constant thickness. Spraying the resist onto the wafer is another method that is sometimes used. Photoresists are radiation-reactive polymers. Positive and negative photoresists are available. A positive photoresist is a type of photoresist in which the portion of the photoresist that is exposed to light becomes soluble to the photoresist developer. The unexposed portion of the photoresist remains insoluble to the photoresist developer (see Fig. 2.2, top row). The effect is the opposite with negative photoresists where the portion of the photoresist that is exposed to light becomes insoluble to the photoresist developer. The unexposed

portion of the photoresist is dissolved by the photoresist developer (see Fig. 2.2, bottom row). In each case, the remaining photoresist serves as a mask for the next process step.

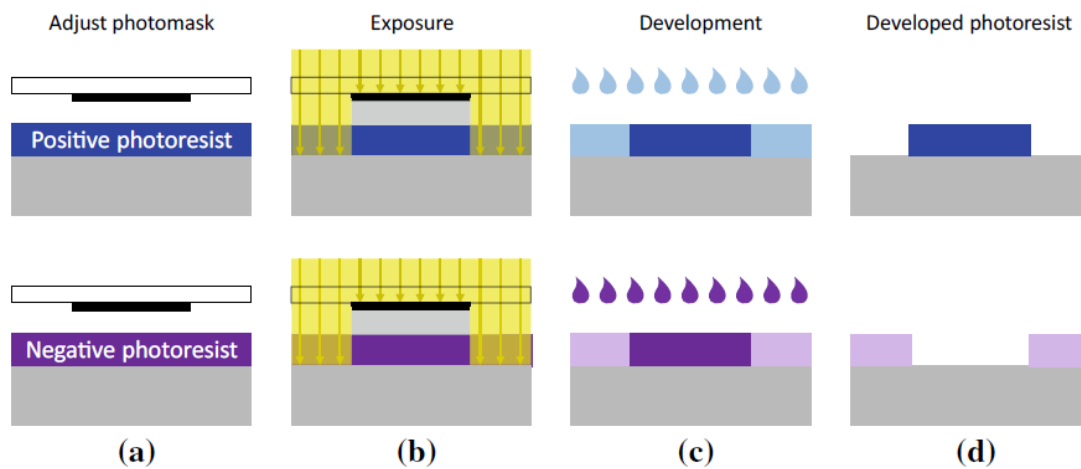


Figure 2.2 Schematic representation of the photolithography with positive resist (above), where the exposed regions are removed, and negative resist (bottom), where the exposed regions are kept.

## Photomasks and Exposure

A *photomask* is used to expose a wafer. A photomask is a sheet of glass, on which a black-and-white image of the structures to be processed is applied to an opaque layer made of chromium. When this photomask is exposed, a shadow is cast on the exposed wafer to produce the desired image. There are two types of exposure: *direct exposure* and *projection exposure*.

### Direct Exposure

With direct exposure, the photomask is placed above and in close proximity to the wafer, either in direct contact with the photoresist (*contact exposure*) or close to it (*Proximity exposure*). The result is in both cases a 1:1 (full scale) image produced by simple shadowing. Hence, the structure sizes on the photomask must correspond with the structures on the wafer (so-called “1X photomasks”).

Contact exposure is unsuitable for volume manufacture as it can cause damage to and soiling on the resist and the photomask. Proximity exposure differs in that there is no contact between photomask and resist. However, the distance between mask and resist in proximity exposure, which cannot be reduced indefinitely but must lie between 10 and 40  $\mu\text{m}$ , can cause resolution degrading.

If the dimensions of the structures being imaged are similar in size to the exposure wavelength, significant diffraction phenomena may occur. Direct exposure therefore approaches its limit for structure sizes to be imaged at this order of magnitude. The limit is even bigger (approximately 3  $\mu\text{m}$ ) for proximity exposure. Hence, the structure sizes required in state-of-the-art semiconductor technologies that are smaller than this value cannot be imaged with direct exposure; this process is therefore not deployed any more today.

## Projection Exposure

*Projection exposure* was developed to overcome the above-mentioned direct exposure issues. Here, the pattern on the photomask is projected through lenses onto a wafer coated with photoresist, as illustrated in Fig. 2.3. This system brings with it two fundamental benefits: (i) photomask and wafer are physically separated and (ii) the image can be optically reduced in size during the projection thus improving imaging accuracy. Among the most commonly used photomasks are 4X, 5X and 10X masks, in other words, photomasks with imaging ratios of 4, 5 and 10 to 1. A small part of a wafer can only be exposed in one step with this method, as the sizes of the photomasks and lenses are limited.

Photomasks therefore have structures for one or a few chips. These types of photomasks are called *reticles*. A Reticle is a special type of photomask where the data for only part of the final exposed area is present. A Reticle is loaded into a Stepper or Scanner system where multiple exposures are made to cover the full patterned area. However, the term “reticle” has become a synonym for “photomask” as state-of-the-art photomasks are seldom designed for an entire wafer.

A wafer is exposed in many single steps in the so-called “step-and-repeat technique”. The wafer is transported on a positioner table under the projection optics. This type of exposure equipment is called a *wafer stepper* or simply a *stepper*. Figure 2.3 shows the stepper principle of operation.



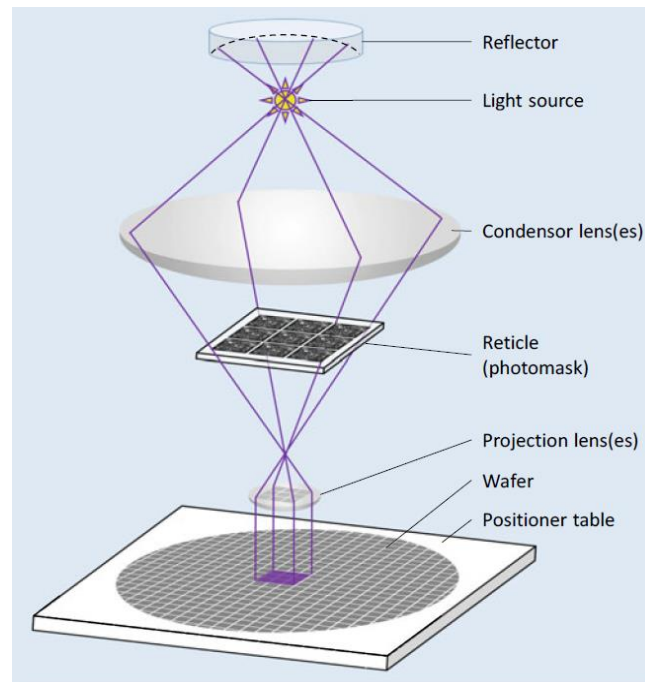


Figure 2.3 Exposing a wafer with step-and-repeat technology. Using one reticle, multiple exposures are made to cover the full patterned area

Technology is being continuously developed to improve imaging accuracy, and such Ultraviolet light is used for higher optical resolution. Exposure and photomask in addition to these developments, there are other approaches for downscaling structure sizes. Mirror systems are needed for projections using light of shorter wavelengths as the materials available for lenses are increasingly opaque.

## Alignment and Alignment Marks

Progressive structural processes that interact with one another are used to construct integrated devices. This means that a reticle's (i.e., a photomask's) position for a specific layer must always be precisely aligned with the structures that are already present on the wafer. For instance, when the devices have been installed, the contact holes need to be precisely positioned where the connecting areas are.

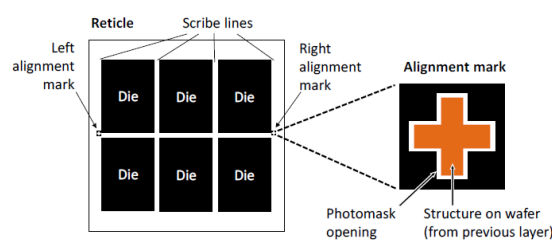
Before each exposure in the step-and-repeat procedure, the wafer's position and attitude with regard to the exposure reticle must be modified. To do this, alignment marks on the reticle are automatically identified by optical methods (Fig. 2.4). These marks, which include geometric shapes like crosses, build structures on the wafer in accordance with the actual process step, just like all other geometric elements on the reticle. If alignment marks' effects on these structures can be seen visually, the wafer position can be aligned with the help of alignment marks on subsequent photomask(s).

An opening in the photomask in the shape of a cross serves as the alignment mark in Fig. 2.4 to the right. The wafer underneath is modified so that the cross-shaped structure, which is also present, is situated in the center of the wafer.

However, some structural processes, like ion implantation doping techniques, for instance, leave no reliable optical traces. But this is not a problem. As mechanical tolerances result in variations in every adjustment, adjustments are made in as many process stages as possible using the same alignment mark on the wafer to prevent the deviations from building up. When a structure can no longer be detected, an alignment mark from a later process stage is used. The "newer" alignment mark is then used in the subsequent phases.

During the alignment procedure, the correct angular orientation of the wafer must be checked in addition to the correct location of the alignment marks. As a result, the alignment marks are placed at two distant points on the reticle to ensure accuracy. the wafer attitude is correct.

The alignment marks are placed outside the chips because they are not functional chip structures. The chips have a 50-100 $\mu$  m clearance from one another to facilitate (subsequent) chip dicing. This clearance is known as a "sawing trench," "saw street," or "saw street clearance." "scribe line". In this clearance space, the alignment marks are placed (Fig. 2.4, left).



*Figure 2.4 Alignment marks on a reticle (photomask) with six chip structures.*

## Reference to Physical Design

The extremely high effort required to create the photomasks derived from the layout design to perform exposure is the price for the degree of miniaturization currently achieved with structure sizes of the order of nanometers. Modern photomasks Semiconductor processes can be quite costly.

It's important to remember that a single design flaw can render photomasks and wafers totally worthless. In the event of a fault, this financial hit is accompanied by a significant delay in development time. Troubleshooting and a new production run can take up to six months. When one considers the delays in bringing the product to market, the financial losses become even more severe.

## Imaging Errors

The structures on the photomasks are determined by layout data. So far, we've discussed how photolithographic images and subsequent targeted process steps are used to convert these mask structures to wafer structures. When the structures created "on" and "in" the wafer are compared to the original layout structures, various types of deviations become apparent. We then go over these unavoidable deviations and how we can deal with them in the layout design.

There are three categories of imaging errors: (1) *overlay errors*, (2) *diffraction effects* and (3) *edge shifts*. Overlay errors and diffraction effects occur during exposure, while edge shifts occur in subsequent structuring process steps.

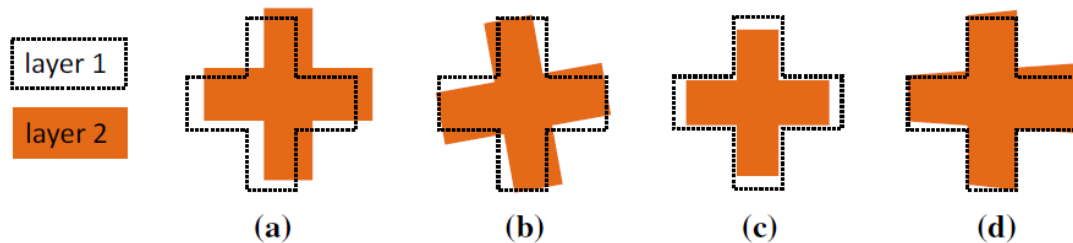
## Overlay Errors

The wafer and photomask (reticle) cannot be positioned with absolute accuracy relative to the exposure device due to mechanical tolerances and measurement inaccuracies that can occur during alignment. As a result, structures on the photomask are not precisely mapped to the wafer during exposure, as specified by the layout template.

Exaggerated representations of possible exposure faults are shown in Fig. 2.5, where we show (a) displacements and (b) rotations that can occur relative to the required positions. To adjust the depth of focus, the wafer and photomask are moved along the optical axis. The layers will be scaled relative to one another if the distances between

the masks, lenses, and wafer are changed during this focusing step (c). The perspective will be distorted if the photomask is tilted with respect to the optical axis (d).

Overlay errors are also caused by the fact that wafers and photomasks, like all materials, expand when heated. If wafer exposures take place at different temperatures, spacings between structures on the wafer will change. As a result, displacements (see Fig. 2.5a) occur, the magnitude of which is determined by where they occur on the exposed wafer. To reduce this effect, keep the exposure temperatures as constant as possible throughout the manufacturing run.



*Figure 2.5 Possible overlay errors between two layers: a displacement, b rotation, c scaling,*

This cause of overlay errors cannot be completely eliminated because it is not always possible to keep these temperatures constant. In high temperature steps of the process, the wafer can become irreversibly deformed. The severity of these deformations varies from wafer to wafer. This effect may result in displacements and scaling (see Fig. 2.5a, c).

The nature and extent of these overlay defects are unpredictable. Their effects are cumulative and can only be limited to specific areas by efforts that concentrate on specific devices and the manufacturing process. Obviously, the maximum allowable overlay error in a semiconductor process should not be greater than the minimum feature size. It should usually be much lower than this value.

Let us take a look at a typical overlay fault to see how it can affect a layout. To ensure proper electrical connection, contacts in chip fabrication must always be completely covered with metal. Because contacts and metallic interconnect layouts use different photomasks, this "design rule" for full coverage must also be met when overlay errors occur.

How overlay errors are handled in layout design is shown in Fig. 2.6. For creating the layout, an enclosure design rule is specified that requires all structures in the "Contact" layer to be covered by structures in the "Metal1" layer and to overlap on all sides by a minimum value. This is the same as the maximum allowable overlay error, which is the deviation that can occur in the worst-case scenario when all overlay errors are superimposed. This design rule is met by the layout structure in the centre of Fig. 2.6. The right-hand side of the figure depicts a possible fabrication situation. For clarity, the assumed overlay error is exaggerated here.

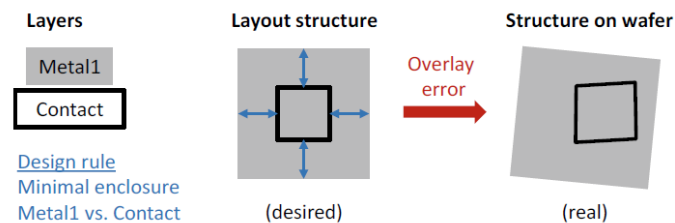


Figure 2.6 Handling of unavoidable overlay errors in physical design. Under all circumstances, contact holes must be completely covered by metal; this necessitates a design rule for a sufficiently large "minimal enclosure" between metal and contact layer that accounts for possible shifting, rotation, scaling, and perspective distortion.

## Edge Shifts

Some technology layers experience effects that cause the graphics elements on the processed wafer to be enlarged or shrunk in comparison to the associated graphics elements in the layout. (Please keep in mind that we are not discussing scaling, in which the dimensions of the elements are changed by a certain factor, as with a "Zoom.") The changes caused by these enlarging/shrinking effects are additive: the structure's boundary lines are shifted outwards by a specific value (positive shift) or inwards by a specific value (negative shift) (negative shift). We call this category of imaging errors edge shifts because the individual structures in the layout are typically modelled as polygons (i.e., as geometrical elements bounded by edge strips).

Figure 2.7 depicts a simple example of a process step with a negative edge shift. The structure on the wafer contracts in relation to the element on the photomask (shown with red arrows in the figure on the right). These edge shifts have layer-specific sizes that are defined for each semiconductor process. As a result, the effect can be mitigated

by pre-sizing the photomask geometry. In particular, the prepared layout data are automatically modified in a layout to mask preparation process that is part of the layout post process. If a value  $k$  edge shift occurs during the process, the edges of the layout geometries are shifted by a value  $k$  before the new data is transferred to the photomask. This operation is shown in the bottom row (b) in Fig. 2.7.

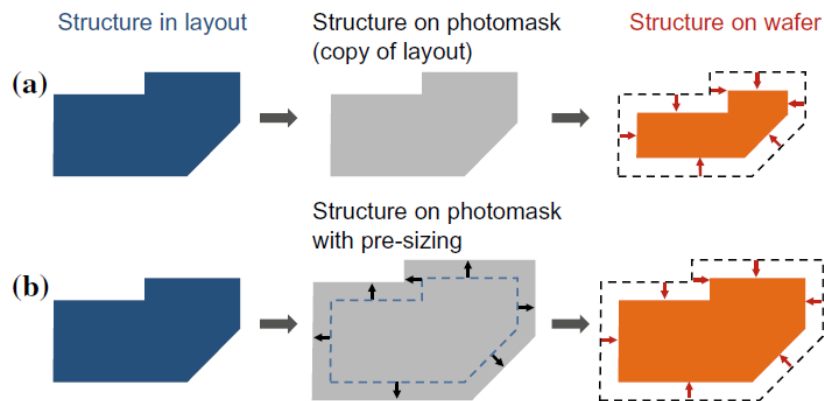


Figure 2.7 Edge shift inwards in the wafer process; a without pre-sizing, b with pre-sizing

## Diffraction Effects

Diffraction phenomena occur at the structural edges of the chromium layer on the photomasks due to the wave properties of light, limiting the optical resolution of photolithography. The smaller the feature size (for a constant exposure wavelength), the greater the impact of diffraction effects on imaging accuracy. In Fig. 2.8, we show these effects with an L-shaped layout structure.

The layout element is transferred to the photomask (in grey) unchanged in the top row. As the feature size decreases, the shape of the area exposed in the photoresist (blue) deviates more from the shape of the photomask opening and thus from the desired layout structure (moving from left-to-right in the figure).

The small fillets at the corners are insignificant as long as the wavelength of light is greater than the feature size (see Fig. 2.8, top left). However, when the feature size to wavelength ratio falls below one (so-called sub-wavelength lithography), significant line-end shortening occurs. Corner rounding has also increased significantly (see Fig. 2.8, top center and right).

These diffraction effects can be corrected by slightly enlarging the photomask opening where it is underexposed and slightly shrinking it where it is overexposed. This procedure, known as optical proximity correction, is another example of a preventive measure (OPC).

As long as the diffraction effects are of the same order of magnitude as those shown in the middle of the figure, these corrective measures can be defined using simple rules based on structural shapes. If the line width is roughly the same as the feature size, "hammer heads" are attached to the ends of "thin" lines. Square elements, so-called "serifs", are added (this means more light) to the outer corners, or "punched" (means less light, hence also called "jogs") from the inner corners. These measures, known as *rule-based OPC*, are illustrated in Fig. 2.8, bottom center. In addition, the change in line widths, caused by interference from many parallel lines, can be corrected by rule-based OPC (not shown in Fig. 2.8).

The extent of imaging errors increases as feature sizes shrink in relation to light wavelength (see Fig. 2.8, top right). In this case, rule-based OPC is ineffective because the exposure result is increasingly influenced by surrounding features. Corrections to the photomasks must then be calculated individually for each structure. The algorithms used in this case are based on models that describe wave-optical effects. A model-based OPC result is shown in Fig. 2.8, bottom right.

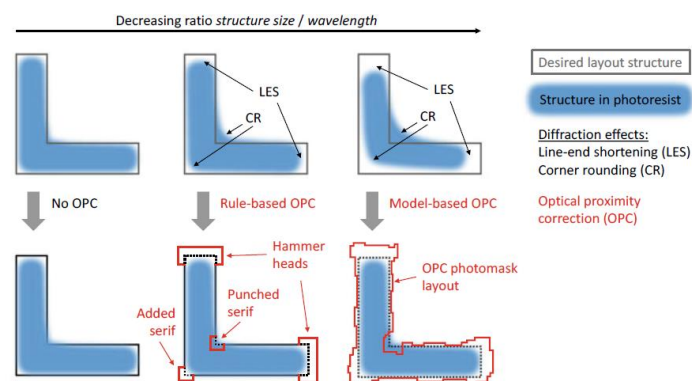


Figure 2.8 Diffraction effects in photolithography (top row) and possible corrective measures using optical proximity correction (OPC, bottom row). The imaging errors increase as the ratio of feature size to optical wavelength decreases (from left to right).

## **Applying and Structuring Oxide Layers**

One of silicon's major advantages over other semiconductors is that it forms a very stable intrinsic oxide: silicon dioxide ( $\text{SiO}_2$ ). Silicon dioxide, which we will refer to as "oxide" in the following for simplicity, has many beneficial properties.

Oxide is a good electrical insulator and a dielectric in capacitive applications. It is mechanically stable and thus appropriate for a strong layer structure. It is simple to create from a process standpoint, and it serves as a masking layer for many process steps, as we will see. It is also visible. Because alignment features can be detected beneath the oxide, this is a useful factor in fabrication. This also enables a wide range of applications, including LEDs (silicon can emit light), solar cells, and photodiodes (light can penetrate silicon from outside). Different processes are available for producing and structuring oxide layers. We will examine these now.

### **Thermal Oxidation**

Thermal oxidation is used to create oxide on the wafer surface. Once formed, an oxide layer can only grow if oxygen atoms diffuse through it until they reach the silicon beneath. As a result, as the thickness of the oxide layer increases, the rate of oxide growth slows. There are two different thermal oxidation processes: *dry oxidation* and *wet oxidation*.

#### **Dry oxidation**

Wafers are heated in an oxidation furnace and exposed to pure oxygen ( $\text{O}_2$ ) at temperatures ranging from 1000 to 1200 °C (approximately 2000 °F). The oxide grows slowly and produces good quality oxide with few vacancy defects.

#### **Wet oxidation**

In this process, the oxygen first flows through boiling water. The wafer is thus exposed to steam, as well. It is much faster than dry oxidation. It is more difficult to control however and produces a lower quality oxide. This is the first thick oxide layer that is created directly on the silicon.



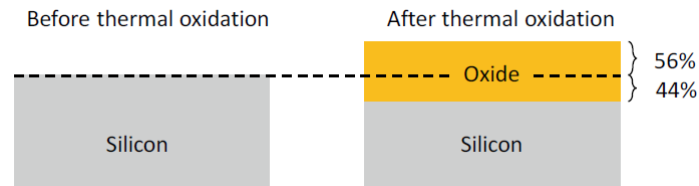


Figure 2.9 Thermal oxidation of silicon (Si) to silicon dioxide (SiO<sub>2</sub>)

## Oxide Structuring by Etching

Etching is a chemical removal process that is frequently repeated for different materials in chip fabrication. It is critical that etchants that remove material be used selectively with respect to the substance, i.e., not etching away other substances (or at least only very slightly).

Figure 2.10 shows how etching can be used to structure an existing oxide layer. A photomask is used to expose and develop an etch-resistant photoresist. There are two types of etching that can be used: wet etching and dry etching.

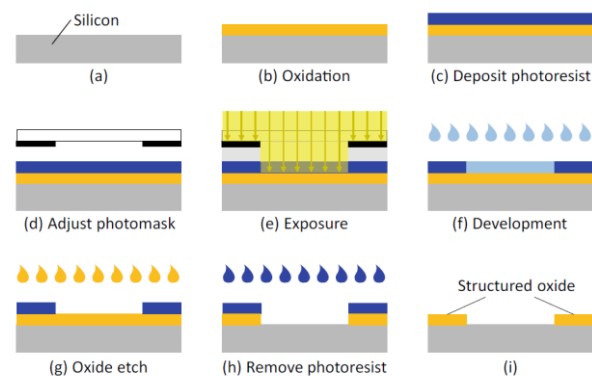


Figure 2.10 Structuring an oxide layer

## Wet Etching

A fluid chemical etching agent dissolves and removes oxide during wet etching. This is a straightforward and widely used method, and the fast etching rate can be adjusted. The disadvantage of wet etching is that it is isotropic, meaning that it acts in all directions. This causes undesirable lateral etching beneath the photoresist. Because of these so-called undercuts, the oxide openings are always larger than the photoresist openings. This results in an edge shift (Fig. 2.11, left). The undercut etching rate is slightly lower than the vertical etching rate because the etching agent cannot circulate as easily under the photoresist and is thus more highly saturated. The lateral undercut is typically 80% of the etching depth.

Due to this undercutting effect, wet etching is no longer suitable for imaging typical feature sizes in advanced processes. As a result, in these processes, wet etching is only used to dissolve and remove entire layers.

## Dry Etching

Reactive ion etching (RIE) is a popular dry-etching method. In general, the etching agent is ionized and used as a gas plasma. An electrically alternating field causes the ions to oscillate. The field is parallel to the wafer surface. Chemically active ions oscillate in this direction and etch material only vertically. The main advantage of RIE is that there is no edge shift in this process (Fig. 2.11, right).

In this process, the etching effect is a combination of a physical effect (where the material to be etched is bombarded with particles in a specific direction) and a chemical effect, i.e., etching. The RIE process can produce extremely fine structures. Furthermore, the trenches formed can be much deeper than they are wide.

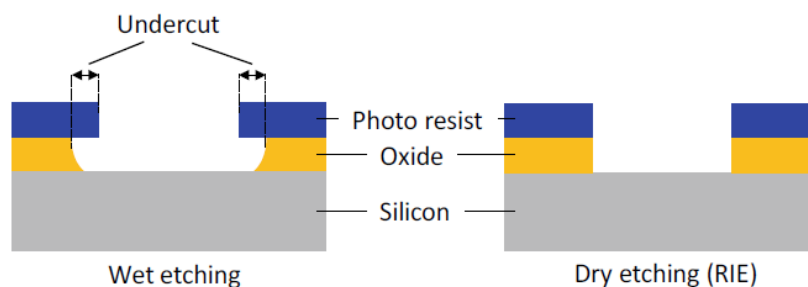
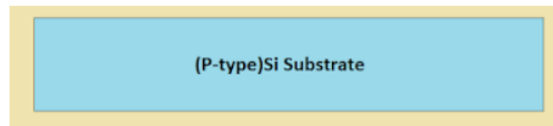


Figure 2.11 A comparison of wet and dry etching (RIE, reactive ion etching).

## Example:

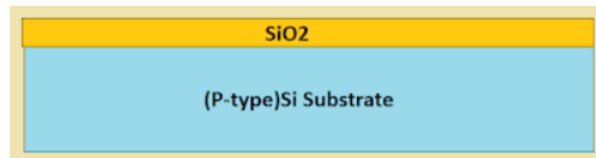
### CMOS fabrication process

- 1- **Substrate:** First we choose a substrate as a base for fabrication. For N- well, a P-type silicon substrate is selected.

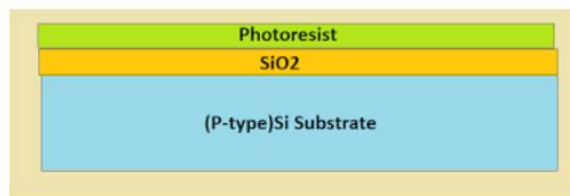


Si substrate p-type

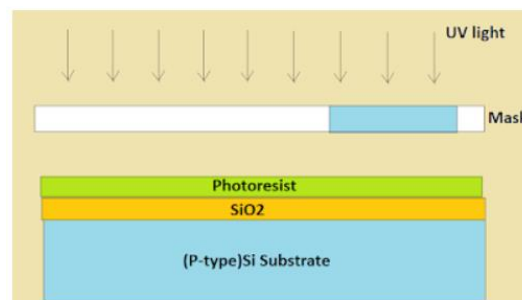
- 2- **Oxidation:** Creating a protective  $\text{SiO}_2$  layer on the wafer surface. It protects the junction from moisture and other atmospheric contaminants. It is used to isolate one device from another.



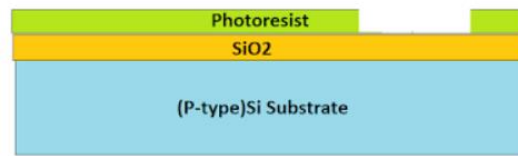
- 3- **Photoresist coating:** A photoresist is a light-sensitive material used in several processes, such as photolithography, to form a specific patterned coating on a surface



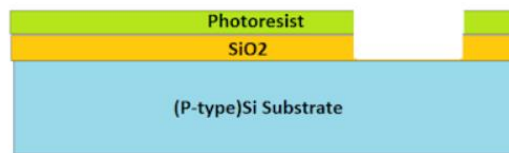
- 4- **Masking:** The process of photolithography includes masking with a photographic mask and photo etching. a mask with the desired pattern is used as a medium to expose UV (Ultra-Violet) lights. The UV light through the mask reaches the photoresist material. The exposed resist remains on the surface and the unexposed part is removed from the surface.



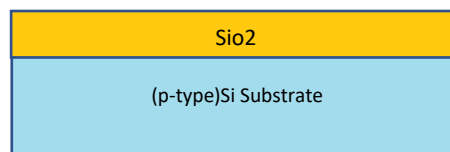
- 5- **Removal of Unexposed Photoresist:** The mask is removed and the unexposed region of photoresist is dissolved by developing wafer using a chemical such as Trichloroethylene.



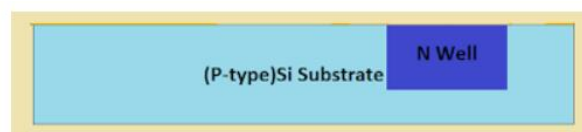
- 6- **Etching:** The unexposed window is removed from the surface and the regions are etched together to form a clean wafer surface. It removes the oxide from the areas through which dopants are to be diffused.



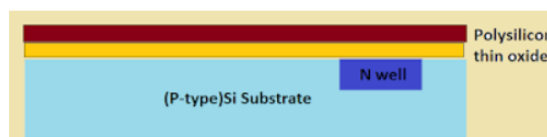
- 7- **Removal of Whole Photoresist Layer:** During the **etching process**, those portions of SiO<sub>2</sub> which are protected by the photoresist layer are not affected. The photoresist mask is now stripped off with a chemical solvent.



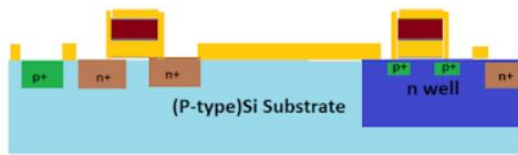
- 8- **N-WELL formation and SiO<sub>2</sub> removing:** The n-type impurities are diffused into the p-type substrate through the exposed region thus forming an N- well.



- 9- **Deposition of Polysilicon:** Chemical Vapor Deposition (CVD) process is used to deposit a very thin layer of gate oxide.



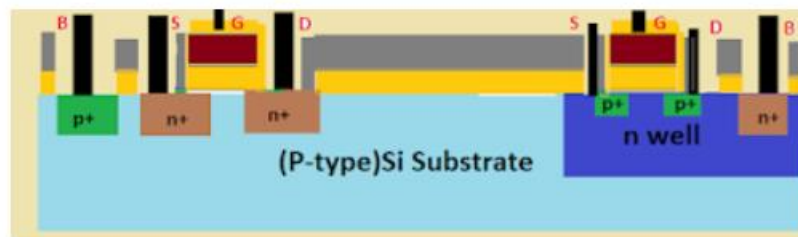
10- **Ion implementation:** it is the process of adding impurities to a silicon wafer.



11- **Metallization:** This step is used for the formation of metal terminals which can provide interconnections. Aluminum is spread on the whole wafer.



12- **Removal of Excess Metal and Formation of Terminals:** In the gaps formed after removal of excess metal terminals are formed for the interconnections.



Following photolithography, one of two procedures is usually carried out. One method involves etching thin films, which is used to transfer photoresist patterns to the thin film(s) beneath. The second involves employing ion implantation and photoresist patterns to prevent certain dopants from entering certain areas.

a wafer's surface in certain places. We go through wet chemical etching and dry etching methods for thin-film etching in this part. We also cover chemical mechanical polishing (CMP), a technique used to remove unpattern thin coatings.

## Thin Film Etching

Wet etching and dry etching are two commonly used methods for transferring patterns into underlying layers after a photoresist pattern has been created. The most important variables for both wet and dry etching are etching rate (amount of thickness reduced per unit time), selectivity, and degree of anisotropy. Temperature and solution concentration are frequently significant determinants of etch rates. The etch rate ratio of one material to another, as determined by the selectivity equation, is the definition of selectivity ( $S$ ).

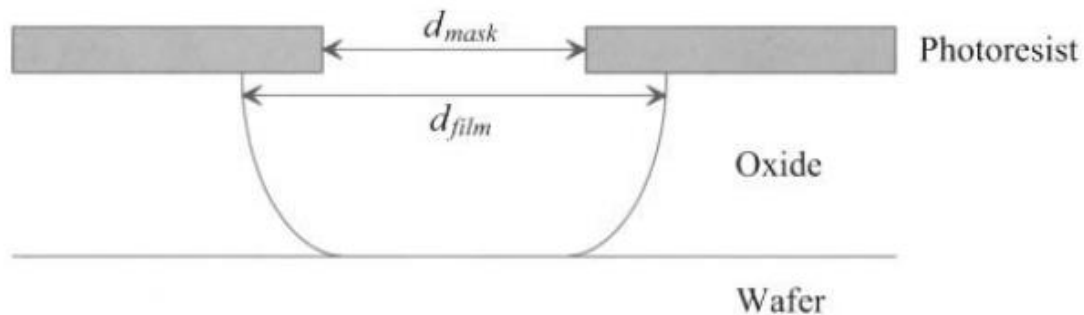
$$S = \frac{R_2}{R_1}$$

where  $R_2$  is the desired material's etch rate and  $R_1$  (7.11) is the intended material's planned removal-resisting etch rate for the material beneath, around, or next to it. The rate at which an etchant dissolves material in various directions is expressed mathematically as the degree of anisotropy,  $A_f$ .

$$A_f = 1 - \frac{R_l}{R_v}$$

where  $R_l$ , is the rate of lateral etch, and  $R_v$ , is the rate of vertical etch. Keep in mind that the etchant is entirely anisotropic if  $A_f = 1$ . The etchant is entirely isotropic, though, if  $A_f = 0$ . The level of anisotropy in conjunction with photolithography plays a significant role in the resolution that can be achieved. The effects of etch bias (i.e.,  $d_{\text{film}} - d_{\text{mask}}$ ) on the final feature size are shown in Figure 7.13. Dry etch methods are referred to wet etch methods for the submicron features needed in CMOS. This is

because dry etching processes often have a higher level of anisotropy. Metal, semiconductor, and insulator removal are accomplished using wet and dry etching.



*Figure: Picture of a post-etch profile. Observe that the mask opening does not match the manufactured opening in the underlying oxide film due to isotropy in the etch process. Etch-bias refers to the variation between these dimensions.*

## Deposition

a method of making semiconductors in which a thin coating is used to add electrical properties to a wafer surface. A thin film of a certain chemical is placed to a wafer's surface during the deposition process to give it electrical properties. The procedure entails depositing several molecule or atomic-thick layers of a substance onto a wafer. The degree of homogeneity and thinness of the film define the quality of the semiconductor, making it an essential step in semiconductor production. Because the covering is so thin, sophisticated technology is required.

Physical vapor deposition (PVD) and chemical vapor deposition are the two forms of deposition (CVD). Physical vapor deposition (PVD), which does not involve a chemical reaction, is typically used to deposit metal layers.

Chemical vapor deposition (CVD), on the other hand, entails supplying outside energy to the vapor of particles produced by a chemical reaction of gas. On semiconductors, insulators, and conductors alike, it can be utilized to deposit thin films.

**Types of deposition: -**

- Thin film deposition
- Chemical-vapor deposition (CVD)
- Low-pressure chemical-vapor deposition
- Plasma-assisted chemical-vapor deposition
- Sputter deposition
- Materials deposited
  - Silicon nitride ( $\text{Si}_3\text{N}_4$ )
  - Silicon dioxide ( $\text{SiO}_2$ )
  - Aluminum
  - Polysilicon

**Chemical Vapor Deposition**

Chemical vapor deposition (CVD) is a process by which gases or vapors are chemically reacted, leading to the formation of solids on a substrate. CVD can be used to deposit various materials on a silicon substrate including  $\text{SiO}_2$ ,  $\text{Si}_3\text{N}_4$ , polysilicon, and so on. For instance, if silane gas and oxygen are allowed to react above a silicon substrate, the end product, silicon dioxide, will be deposited as a solid film on the silicon wafer surface. The properties of the CVD oxide layer are not as good as those of a thermally grown oxide, but they are sufficient to allow the layer to act as an electrical insulator.

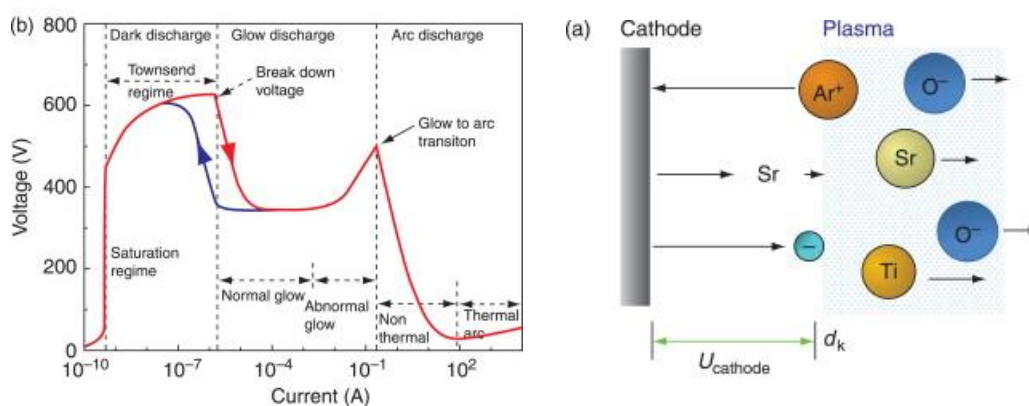
The advantage of a CVD layer is that the oxide deposits at a faster rate and a lower temperature (below  $500^\circ\text{C}$ ).

If silane gas alone is used, then a silicon layer will be deposited on the wafer. If the reaction temperature is high enough (above  $1000^\circ\text{C}$ ), the layer deposited will be a crystalline layer (assuming that there is an exposed crystalline silicon substrate). Such a layer is called an epitaxial layer, and the deposition process is referred to as epitaxy instead of CVD. At lower temperatures, or if the substrate surface is not single-crystal silicon, the atoms will not be able to align along the same crystalline direction. Such a layer is called polycrystalline.



## Sputter Techniques

Sputter deposition is well known for being a method for creating thin films of alloys and complicated materials for use in industry. Its foundation is a discharge of free ions and electrons in a gaseous environment (see Figure 11). Arc discharge, glow discharge, and dark discharge are the three types of discharges that can be recognized. In addition to their luminescence, they can be identified by their breakdown voltage, current density, and current-voltage characteristic (Figure 11(b)). These key features are influenced by the electrode material, process gas, and geometry of the electrodes and deposition chamber.



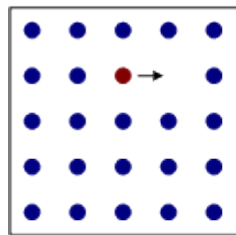
## Doping

Doping is the intentional addition of impurities to a semiconductor crystal in order to change the conductivity of the crystal due to a lack or surplus of electrons. This article explains the partial doping of silicon, which differs from wafer manufacture doping, which involves doping the complete wafer. Diffusion and ion implantation are two methods that can be used to introduce foreign materials (or alloy).

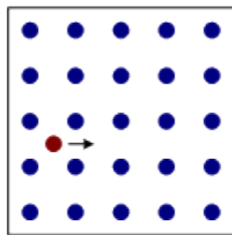
## Diffusion

Molecular diffusion, also known as just diffusion, is the net movement of molecules by random molecular motion from one area of higher concentration to another. Diffusion causes materials to gradually mix together. To give an example, after a specific amount of time, an ink drop in a glass of water spreads out equally. A solid lattice of atoms is present in a silicon crystal, through which the dopant must pass. There are several ways to accomplish this:

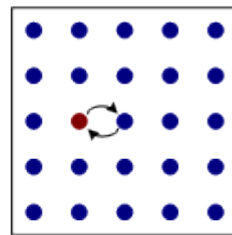
- **empty space diffusion:** Even in flawless single crystals, there are always gaps in the crystal lattice that impurity atoms can fill.
- **inter lattice diffusion:** The impurity atoms in the crystal lattice migrate between the silicon atoms.
- **changing of places:** The silicon atoms trade places with the impurity atoms in the crystal lattice.



Dopants ● fill empty places in the lattice



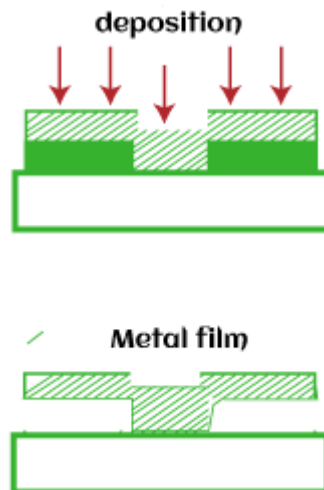
Dopants move in-between atoms



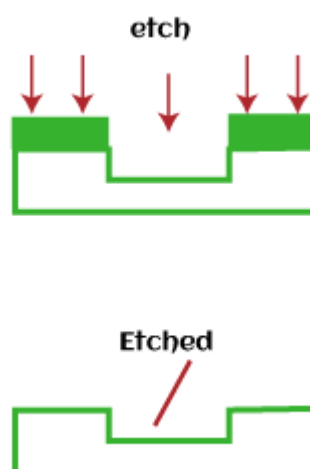
Dopants exchange places with silicon atoms

## Metallization

The process of depositing a metal layer on a metallic or non-metallic surface is known as "metallization." The coating may be made of aluminum, zinc, or silver. Aluminum is the metal used in CMOS production to protect the surface from external environmental elements like dust, air, water, etc.



Metallization is also used to interconnect many components that form an Integrated circuit (IC). The many components that make up an integrated circuit are also connected using metallization. Components include things like relays, transistors, capacitors, and resistors. According to the previous discussion, the metal layer is initially placed on the silicon wafer's surface. After that, as illustrated below, the necessary pattern or region for connecting is carved.



The metallization process uses chamber to apply metal layer. The wafer is placed inside the chamber, which coats the entire surface inside it. The thickness of the metal layer can vary depending on the requirements.

Conductive films enable electrical connection between devices and the environment. And here are Desired properties of the metallization for integrated circuits:-

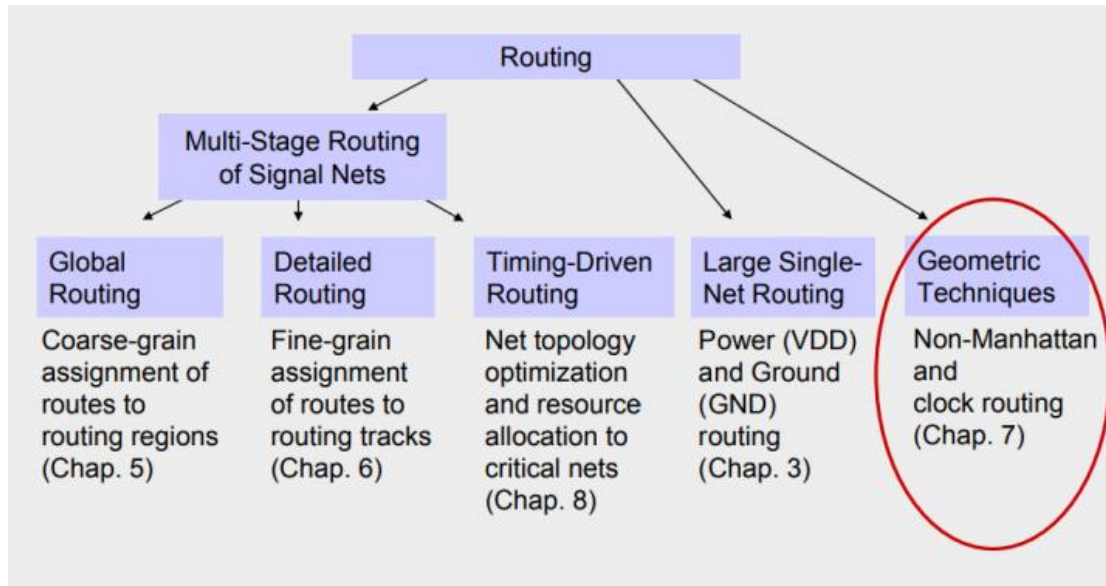
1. Low resistivity
2. Easy to form
3. Easy to etch for pattern generation
4. Should be stable in oxidizing ambients and oxidizable
5. Mechanical stability, good adherence, and low stress
6. Surface smoothness
7. Stability throughout processing, including high temperature sinter, dry or wet oxidation, gettering, phosphorus glass (or any other materials) passivation, and metallization
8. No reaction with final metals
9. Should not contaminate devices, wafers, or working apparatus
10. Good device characteristics and lifetimes
11. For window contacts - low contact resistance, minimal junction penetration, and low electromigration

In addition to offering contacts, gates, and connections, metallization serves two crucial functions. The resistance and capacitance of the connecting runners serve to regulate the circuit speed. The flatband voltage  $V_{FB}$  is also determined by it:

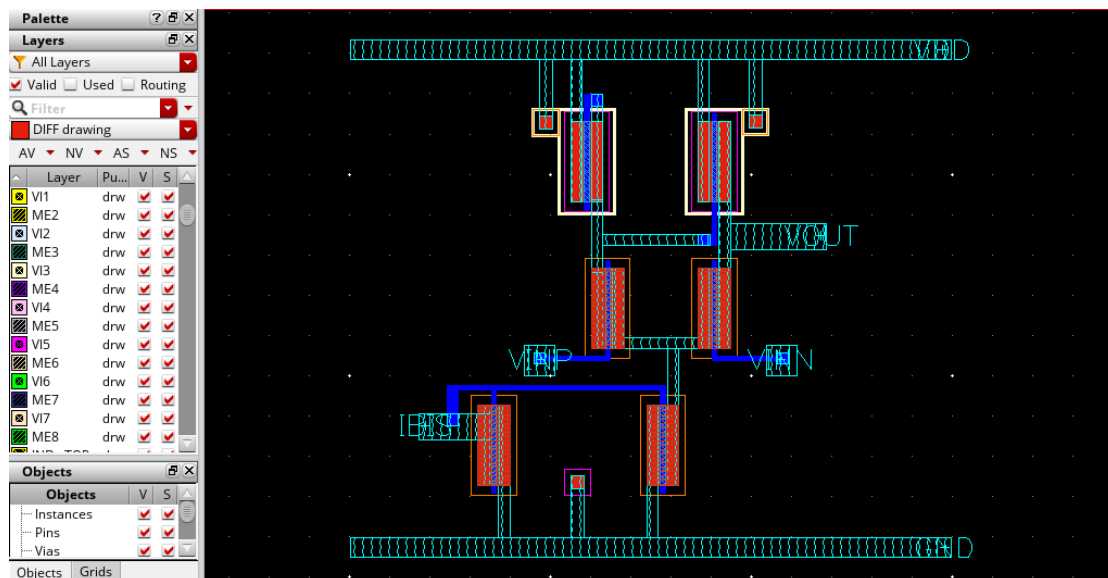
$$V_{FB} = \phi_m - \phi_s \equiv \phi_{ms}$$

where  $\phi_m$  represents the work function of the gate metallization and  $\phi_s$  represents the work function of the semiconductor. The voltage known as  $V_{FB}$  is necessary to maintain a flatband state in the semiconductor by balancing the work function difference between the metal and semiconductor. Thus,  $V_{FB}$  adds to the threshold voltage  $V_T$ , which is the voltage needed at the gate metal to achieve conduction between the source and drain regions.

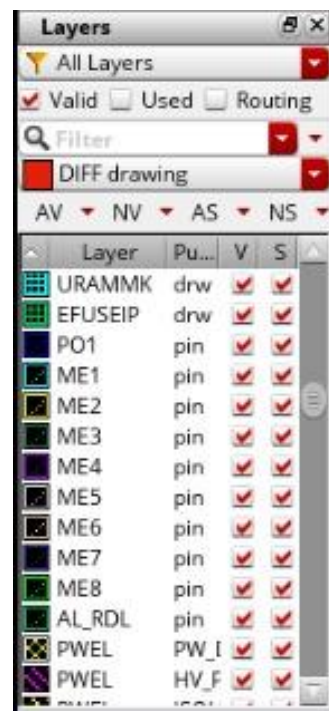
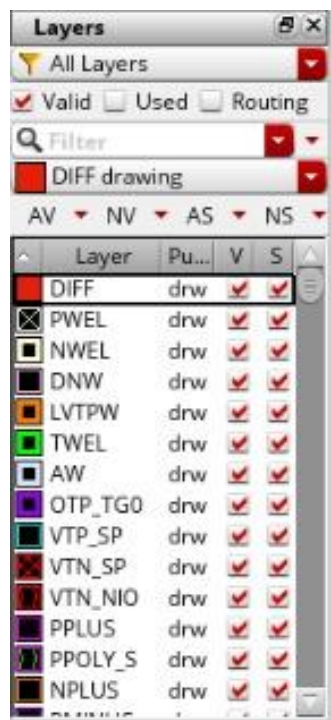
## Signal Routing Matching:



## Routing examples



Metals used in the example:



## **Physical Design Routing Process:**

### **Metal routing:**

- a) It involves generating metal wires to connect the pins of same signal while obeying manufacturing design rules.
- b) Before routing is performed on the design, cell placement has to be carried out wherein the cells used in the design are placed.
- c) The connections between the pins of the cells pertaining to same signal need to be made. At the time of placement, there are only logical connections between these pins.
- d) The physical connections are made by routing. More generally speaking, routing is to locate a set of wires in routing space so as to connect all the nets in the netlist taking into consideration routing channels' capacities, wire widths and crossings etc.
- e) The objective of routing is to minimize total wire length and number of vias and that each net meets its timing budget. The tools that perform routing are termed as routers.
- f) You typically provide them with a placed netlist along with list of timing critical nets. These tools, in turn, provide you with the geometry of all the nets in the design.

### **Routing stages:**

1. Global Routing
2. Track Assignment
3. Detail Routing
4. Search and repair

## Recourses:

- [\[Semiconductor Glossary\] Deposition | Samsung Semiconductor USA](#)
- [Low Pressure Chemical Vapor Deposition - an overview | ScienceDirect Topics](#)
- **EE501Lecture1Process**
- [Plasma-Enhanced Chemical Vapor Deposition - an overview | ScienceDirect Topics](#)
- [Doping techniques - Waferfabrication - Semiconductor Technology from A to Z - Halbleiter.org](#)
- **CMOS Circuit Design Layout and Simulation 3<sup>rd</sup> Edition**
-