UNIVERSIDAD CARLOS III DE MADRID

Manufacturing and packaging of integrated circuits

Integrated circuits and microelectronics

Luis Entrena

Enrique San Millán

Mario García

Celia López

Almudena Lindoso

Marta Portela

1. Integrated Circuits Manufacturing and packaging

1.1 Main processes for integrated circuits manufacturing

1.1.1 Wafer preparation

The usual procedure for obtaining the silicon wafer is the Czochralski process. High-purity silicon is melted in a crucible. A small Si crystal (seed) is used to contact the melted silicon, so that the Si atoms solidify around the seed creating a crystalline structure. The crystal is slowly extracted with a gentle rotation movement to reduce the temperature gradients. Usually a small amount of dopant atoms (such as borom or phosphorus) is added to the melted silicon to obtain a determined initial doping (obtaining thus n-type or p-type silicon).

The dimensions of the solidified silicon ingot have evolved with time and typically have a diameter of up to 150 mm and a length of up to 2 m. The wafers are obtained by slicing the ingot using a wire saw. Then they are smoothed and cleaned by a combination of physical and chemical polishing methods until a high quality and planarity surface is obtained. The typical thickness of a wafer is less than 1 mm.

Wafers usually have one or two side cuts. The longest or main one shows the direction of the crystal, which is obtained using X-ray techniques. The secondary is generally shorter and shows the orientation (100 or 111) and the dope type (p or n).

1.1.2 Oxidation

Oxidation of Si produces SiO_2. Silicon dioxide plays an important role in the manufacturing of integrated circuits, as it has the following main properties:

a) It is an excellent insulator that allows to separate the different conductor materials

b) It works as a barrier for diffusion or dopant implantation

c) It is a component of the MOS transistor itself, as it forms the dielectric that separates the transistor gate from the channel.

Oxidation may be dry or wet. Dry oxidation is performed in the presence of oxygen, while wet oxidation in the presence of water vapor. In both cases it is done at high temperatures (950-1250ºC). Dry oxidation is slower, but it produces a higher purity oxide.

1.1.3 Diffusion

The goal of diffusion is creating zones with a determined doping (n-type or p-type). To obtain it the wafer is exposed to high temperatures in presence of a dopant (B for p-type; P, As, Sb for n-type) which is usually in liquid or gas form.

Diffusion is produced not only vertically, but also laterally, being both diffusions of comparable magnitudes. The effect of lateral diffusion is that the effective dimension of the doped zone is bigger than the one of the window opened for the dopant to penetrate. As technologies reduce sizes of transistors, lateral diffusion becomes a more important undesired effect.

Another feature of diffusion is that it produces a non-uniform concentration of impurities.

1.1.4 Ion implantation

Ion implantation is an alternative process to diffusion for creating doped zones. The zone that is going to be doped is bombarded with ions that come from a source, and that are accelerated by an electromagnetic field. The crystalline structure of the bombarded zone gets deteriorated, such that the body in that zone generally gets transformed into amorphous silicon. After bombarding the wafer, it is heated so that the implanted zone gets back its crystalline structure with the introduced dopants replacing the silicon atoms in the structure.

Ion implantation has the following advantages over diffusion:

a) It provides a much more homogeneous impurities concentration

b) Lateral diffusion is very small

c) Ions may cross oxide barriers, if they are thin enough, and reach the body.

1.1.5 Epitaxial growth and deposition

Epitaxial growth consists in re-growing some material film upon the surface of the wafer ("epi": over/upon). The most common method for epitaxial growth is the "Chemical Vapor Deposition" (CVD). This method is based on the thermal decomposition of silicon compounds, such as silane (SiH_4) or hidrochlorides of Si. As a result from this decomposition, silicon is deposited on the waver surface creating a film whose thickness depends on the exposition time.

If the deposition is produced upon a crystalline silicon body, then the deposited Si grows "copying" the crystalline structure of the body, which acts as a seed. If the deposition is produced upon some other material (for instance, SiO_2), then Si gets a polycrystalline structure (it does not have the structure of a monocrystal, but the one of multiple aggregated crystals), known as polysilicon, or just "*poly*". Polysilicon is an essential material in the current manufacturing of MOS transistors, as it is the building material for the gate terminal. This is due to two very important properties:

1. As it is a semiconductor, polysilicon provides a high resistivity. However, if properly doped, polysilicon is an acceptable conductor (although not as good as a metal).
2. Polysilicon acts as a barrier for diffusion o ion implantation, which allows manufacturing enhancement MOS transistors with high precision by a process denominated "*self-alignment*". In this process both drain and source are doped, while the channel is masked by polysilicon.

Epitaxial growth can also be used for depositing SiO_2 when there is no way of obtaining it by silicon oxidation of the wafer (for example, to isolate intermediate layers or for the final passivation).

1.1.6 Metalization

The metal that forms the interconnections, contacts and vias are deposited using the "Physical Vapour deposition" (PVD) or CVD. PVD consists in evaporating the metal or

bombarding it with ions so that the metal particles get separated and drop upon the wafer surface. The metal patterns are obtained by a photolithographic process as the one used for other materials.

The most used material is aluminum (Al) because of its low resistivity. Other materials as copper (Cu), different silicides or alloys of Al-Si-Cu can be also used for enhancing the interconnections characteristics.

1.1.7 Photolithography

Photolithography is a mechanism that allows applying the different studied processes selectively upon the wafer surface. It consists of the following steps:

1. The wafer is covered by a thin layer of a sensible to light material, denominated photoresist. This material changes its physical and chemical properties (polymerization) when it is exposed to light.
2. The waver is illuminated with UV light, through a mask that let pass the light selectively. This way, exposed zones change their physical-chemical propertis as mentioned before..
3. Using a selective chemical agent (developer) the polymerized material can be removed. Then, opening a window to expose the wafer surface, the mask patter can be reproduced on the wafer surface.
4. Then the exposed material through the window (SiO_2 o metal) can be removed by a chemical attack (etching). The photoresist protects the non-exposed zones as they are not sensible to this chemical attack.

Photoresists can be positive or negative. In the case of a positive photoresist, the polymerized material that has been exposed to light is removed, so the obtained pattern corresponds to the mask. For a negative photoresist, it is the non-exposed part the one that gets removed, corresponding to the “negative” of the mask.

Photolithography has a paramount importance in the manufacturing of integrated circuits, as the size of the transistors depends on the resolution that this process is able to achieve. In order to get a high resolution in this process it is necessary to use a very low light wavelength, and very high quality chemical agents and photoresists. In practice, illumination is performed with UV light or electronic beam.

1.2 Manufacturing process of a CMOS circuit

1.2.1 Manufacturing process of a CMOS transistor[1]

1.2.2 Manufacturing process of a CMOS inverter. Single and twin tub processes ("twin tub")[2]

1.3 Manufacturing of integrated passive components

1.3.1 Integrated resistors

Resistor types:

- poly
- diffusion or ion implantation
- thin film

1.3.2 Integrated capacitors

Capacitor types:

- poly-poly
- poly-diffusion

1.4 Packaging and mounting technologies[3]

1.4.1 Introduction

Once the *die* is manufactured and passivized it is necessary to include it into a package. The package has several important functions:

- Protect the circuit from the outside environment, mainly against wet that humidity that can produce the metallization corrosion.
- Dissipate effectively the heat generated inside the chip. For this reason it is needed a material with a low thermal resistivity.
- Interconnect the chip pads and the external pins. The former should have an easy to handle size for their interconnection on the board, and at the same time a capacity that does not limit the operating speed of the circuit.

These functions are apparently feasible, but are not easy to obtain in practice. The cost of the package may be an important percentage of the final cost of the circuit (more than 50% in some cases), especially if the power consumption of the circuit is high.

---

[1] Slides

[2] Slides

[3] Herbst 2.6

1.4.2 Packaging types

Depending on the utilized material, packages may be plastic or ceramic. Plastic packages are more economical, but they have higher thermal resistivity and provide less protection against humidity. Ceramic packages are more expensive, but on the other hand they dissipate heat more effectively and provide hermetic sealing.

The main characteristics that distinguish some packages from others are:

- dimensions
- material of the package (plastic or ceramic)
- maximum number of pins
- spacing between pins (“pitch”)
- Mounting scheme: “through-hole” (TH) or “surface mounted” (SM)
- Thermal resistivity

Some of the more common types of packages and their characteristics are shown in the following table:

Package type	Number of pins	Mounting scheme
Dual-In-line (DIP)	8-64	TH
Single-In-line (SIP)	5-40	TH
Zig-zag-In-line (ZIP)	14-28	TH
Quad-In-line (QUIP)	14-64	TH
Small Outline (SO, SOIC) Shrunk Small Outline (SSOP)	8-32	SM
Leaded Chip Carrier (LCC) Plastic Leaded Chip Carrier (PLCC)	16-200	SM (TH with socket)
Flat Pack (FP) Quad Flat Pack (QFP) Ceramic Quad Flat Pack (CQFP) Plastic Quad Flat Pack (PQFP) Thin Quad Flat Pack (TQFP) Fine Pitch Quad Flat Pack (FQFP)	10-300	SM
Pin Grid Array (PGA) Plastic Pin Grid Array (PPGA) Ceramic Pin Grid Array (CPGA)	68-500+	TH
Ball Grid Array (BGA)	500+	SM

1.4.3 Multi-chip modules (MCM). Types of multi-chip modules

A multi-chip module or MCM ("Multi-Chip Module") is a packaging type where several integrated circuits ("bare chip") are mounted upon a common substrate and are packaged together.

The main advantages that a MCM provides are:

- Reduction of the interconnection distances between the different chips, which enhances speed and miniaturization.
- Reliability improvement
- Weight reduction, which increases the mechanical reliability (shock, vibrations, etc.)

There are several types of MCM technologies, that can be differentiated by the substrate type used for mounting the chips:

- MCM-L. Substrate based on multilayer PCB laminated technology. Basically it is as a printed circuit board of thin line.
- MCM-C. Ceramic type substrate.
- MCM-D: Dielectric or conductors deposited upon a substrate that may be silicon, ceramic or metal. The deposition is performed with similar techniques to the ones used in integrated circuits manufacturing, and for this reason they are also known as thin layer MCMs. The materials that are deposited are silicon oxide or poly (dielectrics) and metals (conductors)

MCM-D is the technology that provides better performance and routing density, but it is also the most expensive one. On the other hand, MCM-L is the most economical, but it provides the worst performance from these three technologies. MCM-C is a middle term in performance and cost..

Characteristics	MCM-L	MCM-C	MCM-D
Line density (cm/cm^2)	30	20-40	200-400
Line width (μm)	750	125	10
Separation (μm)	2250	125-375	10-30

1.5 Application-Specific Integrated Circuits. Types and characteristics

Application-Specific Integrated Circuits (ASICs) can be classified into two groups, depending on how they have been designed:

- *"full custom"* or just *"custom"*. Design and manufacturing is completely customized. The cost of design and manufacturing of *full custom* prototypes is very high, so it only makes sense for big quantities, as in standard products (products that are not focused on a specific application).
- *"semicustom"*. Partially pre-designed and/or partially pre-manufactured. This way the cost is significantly reduced by re-use of parts of the design (partially pre-designed) or manufacturing masks (partially pre-manufactured). However, the

characteristics of the circuit (area, delays, etc.) are worse than in a full custom circuit, as pre-designed or pre-manufactured components impose some design constrains.

Semicustom circuits can be classified into:

Standard Cell

They are partially pre-designed circuits whose elemental components (as logic gates or flip-flops) are pre-designed. These components are organized in libraries, so that the designer only has to select which components he wants to use. On the other hand, they allow a systematic application, so it is a design method very suitable for being used with automatic synthesis tools.

The advantage of standard cells is that the designer does not have to design each logic gate, while de disadvantage is that if he needs a logic gate with some special characteristics (for example a very high speed logic gate) it may happen that it is not available in the library.

Gate Array and Sea of Gates

They are circuits partially pre-manufactured. Concretely, all the transistors are pre-manufactured but some or all the metallization layers. Wafers with these transistors are manufactured in big quantities, with very economical prices, and can be used for any circuit. In order to manufacture a specific circuit, it is only necessary to create the metallization masks and apply them on the wafer that contains the pre-manufactured transistors

The difference between the *Gate Array* and *Sea of Gates* technologies are shown in the attached slides. A Gate Array has transistor rows interleaved by spaces, denominated *routing channels,* which are free zones used to create interconnections.

In the periphery of the circuit are located input/output blocks that contain the *pads* and their associated circuitry. In the Sea of Gates technology, more recent, all the surface of the circuit is occupied by pre-manufactured transistors. The interconnections can be created over the transistors, which may become unusable for this reason. Nevertheless, the Sea of Gates technology provides a higher routing flexibility.

In summary, the advantage of these technologies is the costs reduction that is obtained by starting with wafers with pre-manufactured transistors. Once again, the main disadvantage is the design limitation, as all the transistors have the same characteristics.

Circuits based on macrocells

In modern integrated circuits of high complexity, and especially in the denominated System-on-Chip (SoC), it is common to use different combinations of techniques. These circuits may contain:

- Some parts designed with standard cells
- Some pre-manufactured parts, that can be configured at a later stage
- Macrocells. Complex pre-designed components, such as memories, microprocessors, etc.

An example of circuit based on macrocells is a microcontroller family. The manufacturer may pre-manufacture a basic circuit that contains the microprocessor

(macrocell), a pre-manufactured region as a Gate Array, and another free region for standard cells. In any of the former two, components as I/O ports, peripherals, etc. can be used, so that from the same wafer all the microcontrollers in the family can be obtained, having to implement only those components that are specific for each circuit.

Macrocells can be bought for a third party, and can be delivered at different abstraction levels. These macrocells are known as IP ("Intellectual Property"). In general, it can be distinguished between *soft macrocells*, that are behavioral or logic designs (for instance, a schematic), or *hard macrocells*, that are designs with a defined layout and it is only needed to place them on the chip surface. Macrocells can be re-used in lots of designs, which significantly reduce the cost of the circuits that use them. The main disadvantage is the difficulty of protecting them against misuse, as they are products distributed on digital media that can be easily copied and their utilization is difficult to control.

Programmable Circuits

They are completely pre-manufactured circuits that the user can configure to implement a determined function[4]. The advantages and disadvantages are similar to those of the semicustom technologies, described previously. Programmable circuits present great ease of use, low programming cost (in comparison with the cost of manufacturing a circuit), and low time to market. In contrast, the cost per unit is higher than in a custom circuit as they require more area for the same functionality. For this reason, programmable circuits are very suitable for prototype manufacturing or small production volumes (up to 30,000 units). The maximum operating frequency is typically one third of the cost of a custom circuit.

Analog and mixed signal circuits

Currently is common to find circuits that combine analog and digital parts, which are known as *mixed-signal* circuits. These circuits pose an additional manufacturing difficulty, as the manufacturing processes for analog and digital circuits are noticeably different. In general, the analog part is separated from the digital part through wells and guard rings to avoid problems with electric noise.

---

[4] See summary about FPGA types

2. Analysis and design of integrated circuits at the physical level

2.1 Necessary masks for manufacturing a CMOS circuit

The number of masks in a current process can be very high. The most typical masks for NMOS and CMOS processes are shown next:

Typical masks of a NMOS process:

1. Diffusion or active mask: it defines the areas where the transistors will be placed, it is used to create the thin or gate oxide pattern.
2. Depletion transistors mask: it defines the channel for depletion transistors that, in contrast to enhancement transistors, must be doped.
3. Poly mask: it defines the polysilicon patterns; the doping of n+ does not require a new mask, as it is created by *self-alignment*.
4. Vias and contact masks: as many as metal layers there are.
5. Metal masks: as many as metal layers there are.
6. Pads mask: it defines the areas where the pads are to be placed; it is applied after the final passivation.

Typical masks of a CMOS process:

1. Well masks: they define the areas where the Wells are going to be placed; it is required at least one if the process is single tub, or two if it is twin-tub.
2. Active mask: it defines the areas where the transistors will be placed, it is used to create the thin or gate oxide pattern.
3. Poly mask.
4. p+ doping mask: it allows the diffusion of p-type impurities to obtain PMOS transistors, while the NMOS transistors are protected.
5. n+ doping mask: it allows the diffusion of n-type impurities to obtain NMOS transistors, while the PMOS transistors are protected.
6. Vias and contact masks: as many as metal layers there are.
7. Metal masks: as many as metal layers there are.
8. Pads mask: it defines the areas where the pads are to be placed; it is applied after the final passivation.

2.2 Design rules

Design rules are geometrical minimum specifications for the layout elements to guarantee that the design patterns may be accurately reproduced on the wafer. These rules have effect on the dimensions, spacing, overlapping, etc. of the layout pattern. The ultimate goal is to obtain circuits with an optimal performance or *yield* (percentage of

valid circuits over the total manufactured), with the minimum possible area but without risking reliability.

Design rules are derived from the manufacturing process itself, and therefore they are provided by the integrated circuit manufacturer. In summary, they are the essential information that the manufacturer provides to the designer, so that the former can realize the designs without the need of knowing in detail all the integrated circuits manufacturing problematic issues.

2.2.1 Motivation for design rules: "latch-up", manufacturing tolerances, etc...

The origin of design rules are manufacturing tolerances and removal of parasitic effects that may affect the circuit operation. Some of the most important rules that affect each element are shown next:

- Wells and diffusions: The separation distance between wells must be high enough to give room for lateral diffusion.
- Poly: Poly must protrude outside the active zone. In other case, an active mask alignment error could make poly not to cover completely the active zone, producing a short circuit between drain and source (r307). There must be a minimum diffusion size at both sides of poly to guarantee that the transistor is formed under the alignment tolerances constrains (r306).[5]
- Contacts and vias: The rules for contacts are not minimum distances but exact distances. Although it would be desirable to use wide contacts to reduce their resistance, such contacts cannot be manufactured easily. Instead, it is preferable to utilize multiple contacts connected in parallel. This is due to the fact that the chemical attack that removes the oxide to open the contact requires higher depth as we get further from the edges; therefore, the depth of the attack at a contact center could become excessive. The upper zone of a contact (metal) is always bigger than the bottom zone, so that the contact can be created correctly in case of an small alignment error (r403). The rules that correspond to vias are similar to the rules for contacts as they have the same characteristics.
- Metals: Metal lines have a minimum width and separation sizes (r501 y r502). In many cases, these distances increase as we reach higher metal layers. The reason is that the higher the layer is, the more irregular the wafer surface becomes, and therefore it is necessary to consider higher tolerances.
- Pads: Pad dimensions have not decreased with the general reduction in technologies, due to the fact that its function is to allow soldering to a metal thread. All the pad distances are very high in comparison for this reason.

2.2.2 Lambda rules (λ)

Design rules can be specified in μm or in "mils" (thousandths of inches) for big sizes. However, usually it is considered a characteristic λ parameter, associated to the minimum possible resolution of the manufacturing process. The design rules specifications that make use of this parameter have the advantage of ease of scaling to a

---

[5] See figures p. 151 Weste

technology with higher resolution: it is just necessary to change the value of λ, and all the rules scale linearly. Design rules based on λ are portable and scalable. However, it is important to note that for current sub-micron technologies, scaling is not linear in many cases, so λ is reminiscence from old technologies (with channel lengths over a micron).

It is usual referring to a technology or process by the polysilicon width, which determines the channel length of transistors. This parameter is indicative of the accuracy of the manufacturing process, as for obtaining the best performance transistors have to be fabricated with the minimum possible channel length. Typical polisilicon width is approximately 2λ. For instance, in a 1µm technology, the minimum width of a polysilicon line and thus the transistor channel length is 1µm, and λ = 0,5 µm. All the other dimensions are integer multiples of λ or at least in some cases of 0,5 λ.

2.2.3 Design rules based on λ

See examples shown in the slides.

2.2.4 Design Rules Checking (DRC)

When the design of the layout for a circuit is finished, it always should be checked if the design rules are met. For this purpose layout design tools usually provide a design rule checker tool for Automatic Rules Checking (DRC).