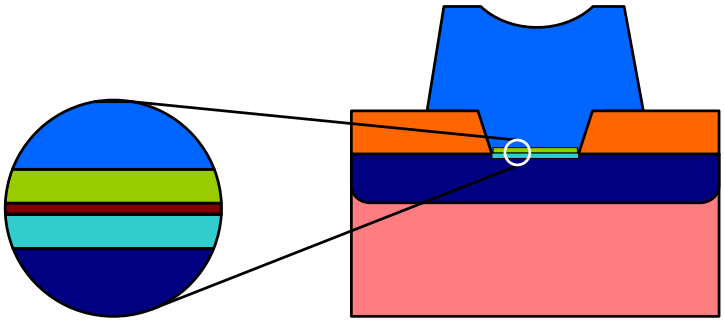
Semiconductor Technology

von A bis Z



Metallization

www.halbleiter.org

Contents

List of Figures

List of Tables

1 Metallization

1.1 Requirements on metallization

The metallization realizes contacts to the doped regions in semiconductor manufacturing with conducting paths. From here the connectors are lead to the edge of the microchip to finally connect it to the package or for testing reasons.

Following requirements are essential for metallizations in integrated micro electronic devices:

- excellent adhesion on silicon oxides (insulators between metallization layers)
- high ampacity, low resistance
- low contact resistance between metal and semiconductor
- simple processes to deposit the metallization layers
- low susceptibility to corrosion for long life times
- excellent contacting with wire bonds
- possibility to integrate stacked layers
- high purity of the metal

Aluminum meets many of this requirements and therefore has been the material of choice for many years. However, since structures are getting smaller and smaller, aluminum can't fulfill the requirements any longer. For this reason, copper will replace aluminum in the future.

1.2 Aluminum technology

1.2.1 Aluminum and aluminum alloy

Because of its properties aluminum and its alloys are widely used for wiring in microchips:

- excellent adhesion on SiO2 and interlayers as BPSG or PSG
- excellent contacting with wire bonds (ie gold and aluminum wires)
- low electrical resistance (3 $\mu\Omega \cdot cm$)
- simple to structure in dry etch processes

Aluminum fulfills the requirements in electrical toughness and resistance against corrosion only partial. Metals like silver or copper have better properties, however, these metals are more expensive and cannot be etched in dry etching this easily.

1.2.2 Diffusion in silicon

The use of pure aluminum leads to a diffusion of silicon into the metal. The semiconductor reacts with the metallization at only 200-250 °C. This diffusion of silicon causes cavities at the interface of both materials which are then filled by aluminum. Thus leads to spikes which can cause short circuits if they reach through the doped regions into the silicon crystal beneath.

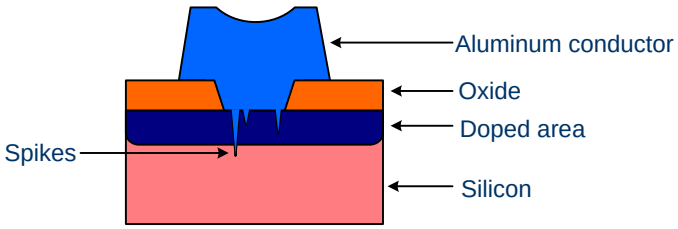


Fig. 1.1: Spikes

The size of these spikes depends on the temperature at which the aluminum was deposited onto the wafer. To avoid spikes there are several possibilities. A deep ion implantation - contact implantation - can be introduced at the location of the vias. Thus the spikes do not reach into the substrate.

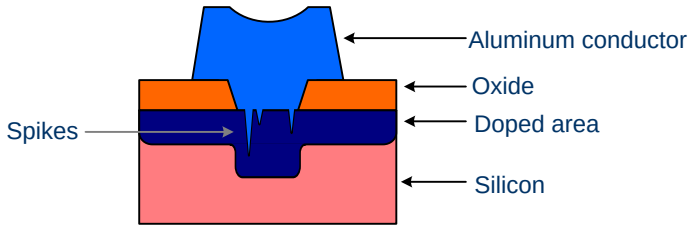


Fig. 1.2: Contact implantation

The disadvantage is that there has to be an additional process step, furthermore the electrical properties change because the doped regions are enlarged.

Instead of pure aluminum an alloy of aluminum and silicon can be used (silicon 1-2 %). Because the aluminum now already contains silicon there will be no diffusion out of the substrate. However, if the vias are very small, the silicon can drop out at the contact area and result in an increased resistance.

For contacts with a high quality a separation of aluminum and silicon is essential. A barrier of different materials (e.g. titan, titan nitride or tungsten) is deposited. To avoid an increased contact resistance at the interface of titan in silicon a thin layer of titan silicide is used.

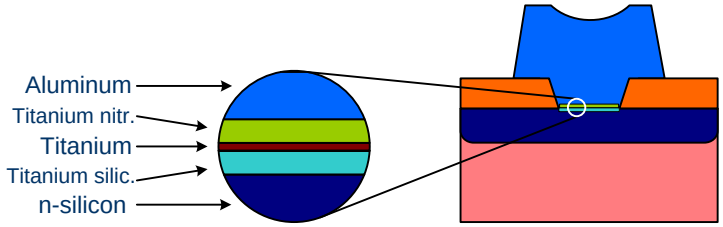


Fig. 1.3: Barrier layer between aluminum and silicon

1.2.3 Electromigration

A high current density results in a friction of electrons and fixed metal ions. The ions are moved due to collisions with the electrons. At locations with small cross sections the current density increases, thus more ions are moved and the cross section decreases which leads to a higher current density. In extreme case the aluminum wires can break.

1.2.4 Hillocks

Electromigration leads to moved material which is accumulated on locations with a lower current density. This hillocks can break through adjacent layers and cause short circuits. In addition moisture can penetrate into the material and cause corrosion. Another reason for hillocks are different coefficients of thermal expansion. Layers expand in different ways due to heating which causes stress. To minimize strain additional layers with an adjusted coefficient of expansion can be deposited between the other layers (e.g. titan, titan nitride).

Further problems which can occure during metallization:

- **Diffracted expose:** the metal layer can reflect light in that way that adjacent regions are exposed. To avoid reflections an anti reflective coating can be deposited
- **Bad edge coverage:** on edges there can be increased aluminum growth, while in corners there is a decreased growth. Therefore edges have to be rounded:

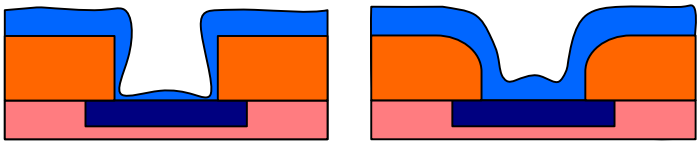


Fig. 1.4: Edge rounding

The layout of the wires has to be planned exactly to avoid these issues. A small additive of copper in the aluminum can increase the life time by far. However, the structuring of the aluminum-copper conductors is much more difficult. To avoid corrosion, the surface is sealed with layers of silicon oxide, silicon tetranitride or silicon nitride. The material of the packages for microchips is some kind of ceramics because synthetic materials are not as resistant.

1.3 Copper technology

1.3.1 Copper technology

Copper has significant advantages compared to aluminum and thus is a good alternative for ever smaller structures. The metal has a much lower resistance than aluminum and is much more efficient in view of power consumption; beacuse of the needs of

smaller and smaller feature sizes, aluminum does not fulfill the electrical requirements any longer. Also the electromigration of copper is much less than of aluminum. A change-over can not be averted.

Copper, however, has the disadvantage that it contamintes alomst everything which gets in contact to it. Therefore areas and equipment on which copper is processed, have to be separated from others. In addition, copper is susceptible for corrosion as well as aluminum and has to be covered with a passivation layer. Another advantage of copper is, that there is no need of tungsten to connect the individual copper layers with each other, and therefore additional process steps are omitted as well as thermal issues at the interface of different materials. One of the biggest disadvantages in semiconductor device fabrication is, that copper can not be structured as easy as aluminum in dry etch processes.

The traditional subtractive process for structuring - as it is used for aluminum and other materials - is done as follows:

- deposition of the layer which has to be structured
- resist deposition, exposure and development
- transfer of the resist mask into the layer beneath via dry etching
- resist removal
- passivation

In copper technology one has to use not a subtractive but an additive method: the so called damascene process.

1.3.2 Damascene process

The damascene process makes use of existent interlayer dielectrics in which the vias and trenches for conduction paths are etched. Subsequent, copper is deposited by CVD, PVD + reflow, or in electrochemical/galvanically processes. Finally, the copper is planished by chemical mechanical polishing (CMP).

The damascene process can be seperated into the single and dual damascene process and the latter can be seperated further into the VFTL (VIA First Trench Last) and the TFVL approach (Trench First VIA Last). In the following, both the TFVL as well as the VFTL process are described.

Dual damascene: Trench First VIA Last:

On top of the wafer (in this example on an existing copper layer) different layers are deposited which act as protection, isolation or passivation layers. As an etch stop and protection against gaseous molecules, silicon nitride (SiN) or silicon carbide (SiC) can be used. As interlayer dielectric (ILD), materials with a low relative static permittivity are used, like silicon dioxide SiO_2. Upon a resist mask is patterned.

1. The wafer is coated with a resist layer which is structured in photolithography (Fig. 1.5).

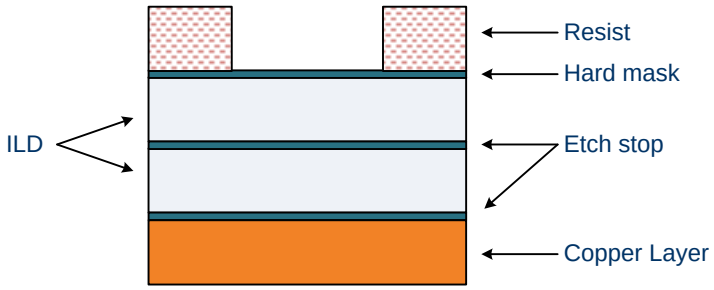


Fig. 1.5: Basic film stack of a damascene approach

2. The hardmask (SiN) and the ILD are etched in a anisotropic dry etch process until the first etch stop layer (SiN) is reached. The resist is removed and the trench for the conduction path is finished. (Fig. 1.6(a)).

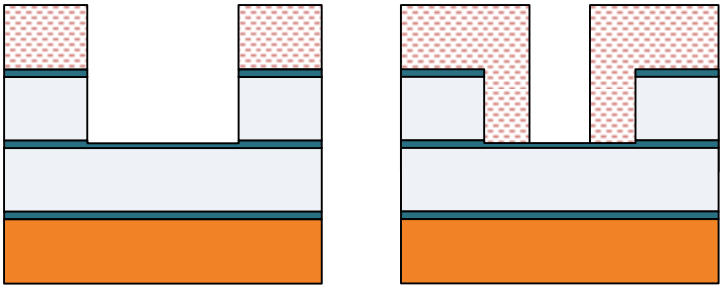


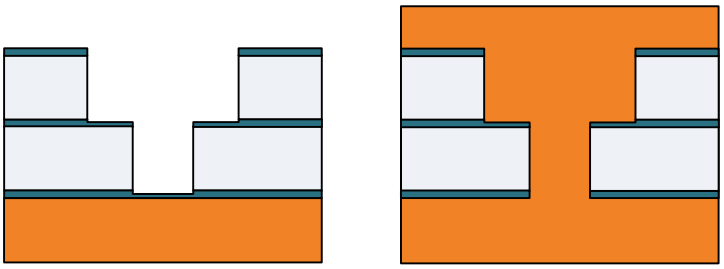
Fig. 1.6: Trench etch and via resist mask

The hardmask on top protects the ILD during the resist ash. This is necessary since the ILD has a similar composition as the resist and therefore is affected by the same process gases. In addition the hardmask acts as a barrier layer during terminal CMP.

3. Next a new resist layer is deposited and structured (Fig. 1.6(b)).

4. Finally the vias are opened in an anisotropic etch process.

With a low energetic etch process, the bottom etch stop is opened to avoid sputtering of the copper beneath which could deposit on the sidewalls and diffuse into the ILD. The resist is removed and a thin layer of tantalum is deposited as a barrier which prevents later deposited copper from diffusing into the ILD (fig. 1.7(a)).



(a) (b)

Fig. 1.7: Final structure after etch and copper deposition

5. A thin copper layer acts as a seed layer, so that the vias and trenches can be filled in a galvanically process (Fig. 1.7(b)).

6. The deposited copper is planished in a CMP process (Fig. 1.8).

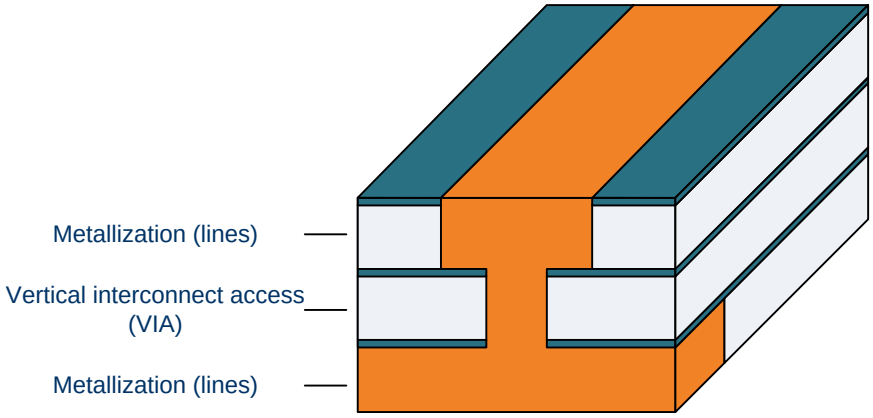


Fig. 1.8: Damascene stack after metallization and CMP

The big disadvantage of this process is the thick resist layer which is deposited after the trenches have been etched (3). To etch the tiny vias in such a thick resist layer is very difficult. For this reason the TFVL approach is done at larger structures only.

Dual damascene: VIA First Trench Last:
The VFTL approach is alike the TFVL process but now the vias are created first.

1. A resist layer to form the vias is structured and the vias transferred into the ILD by an anisotropic etch process till the bottom etch stop layer is reached. To prevent copper from being sputtered out of the metallization beneath, the etch stop must not be opened (Fig. 1.9).

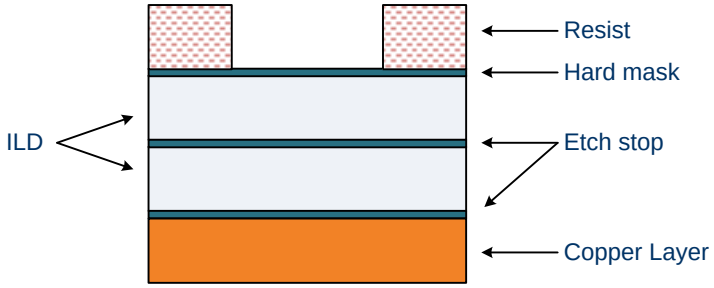


Fig. 1.9: Damascene stack after VIA etch

2. Subsequent the resist is removed, a new resist layer is patterend which represents the trenches; also the previously opened vias are filled with resist (Fig. 1.10(a)).

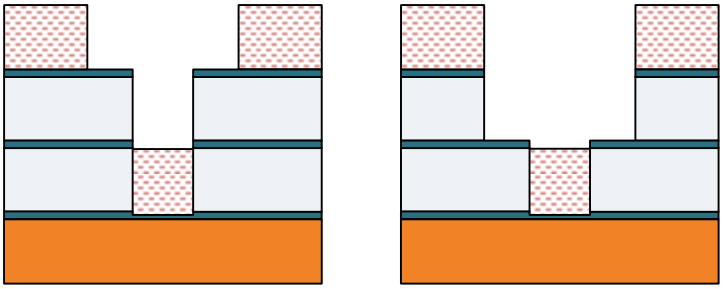


Fig. 1.10: Photolithography and trench etch

3. During the trench etch, the bottom etch stop is covered by resist. Next the bottom etch stop is opened in a low energetic process and a tantalum barrier and a copper seed layer are deposited (Fig. 1.10(b)).

4. After the copper deposition, the metal is planished in a CMP process (Fig. 1.11).

In the single damascene process, the via layer and the trench layer are deposited and structured one after each other, so that there are more process steps needed (ILD deposition ⇒ VIA structuring ⇒ copper deposition ⇒ planarization ⇒ ILD deposition ⇒ trench structuring ⇒ copper deposition ⇒ planarization).

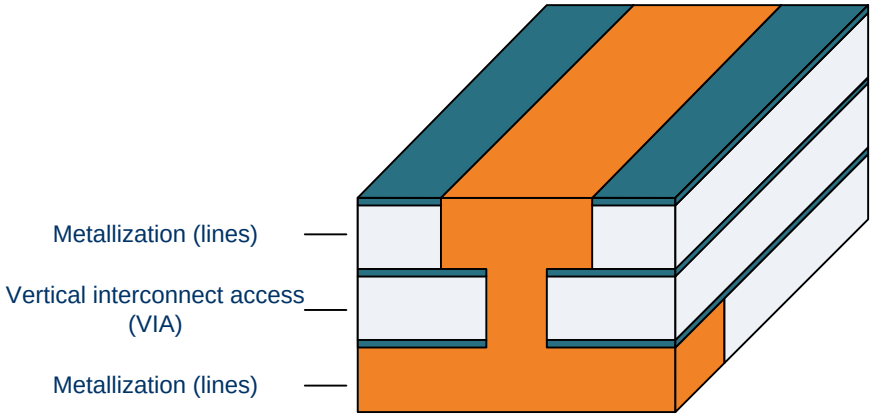


Fig. 1.11: Damascene stack after metallization and CMP

1.3.3 Low-k technology

Since there is a proceeding miniaturization of the structures on microchips to increase packing density, reduce power consumption, and increase switching speeds, the conductors for wiring are moving closer and closer together in vertical and horizontal direction. To isolate the conductors from each other additional films like silicon dioxide SiO_2 have to be deposited as an interlayer dielectric (ILD).

If conductors run parallel or cross each other on different layers upon another, parasitic capacities are created. The conductors represent the electrodes while the SiO_2 in-between is the dielectric.

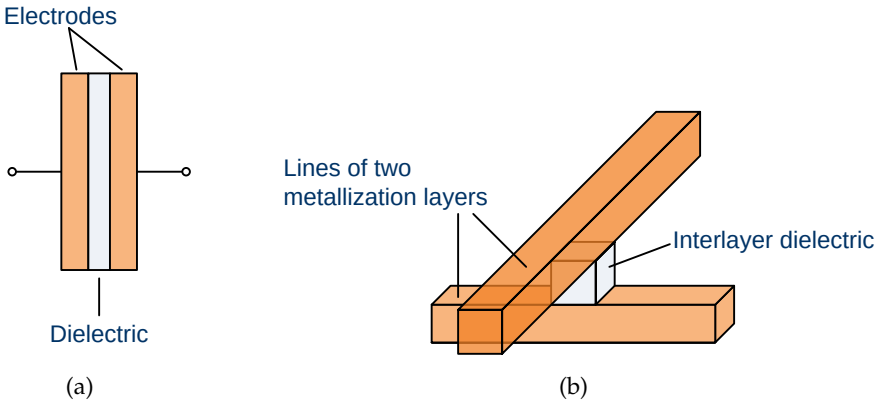


Fig. 1.12: (a) Traditional capacitor and (b) crossing of two conductors

The capacity C is given by:

$$C = \frac{e_0 e_r A}{d}$$

Where d stands for the distance of the electrodes, A is the area of the electrodes, ϵ_0 the vacuum permittivity and ϵ_r (often κ (Kappa) or simplified k) the relative static permittivity of the ILD.

The value of the parasitic capacity influences the electric properties such as the switching speed or the power consumption of a chip, and therefore one tries to decrease C. Theoretical this can be done if ϵ_0, ϵ_r and A are decreased or if d is increased. However, as mentioned above d is getting smaller and smaller, A is preset by electrical requirements and ϵ_0 is a physical constant. Thus the capacity can only be reduced by decreasing ϵ_r.

To sum it up, one needs dielectrics with a $low\epsilon_r$: *low-k*.

The traditional dielectric, SiO_2, has a relative permittivity of about 4. Low-k referes to materials whose ϵ_r is less than that of silicon dioxide. Beyond that there will be Ultra-Low-k materials with an ϵ_r of less than 2.4. The permittivity referes to the polarization (dislocation of charge carriers in the insulator) in the dielectric and is the factor by which the charge of a capacitor is increased relative to vacuum or by which the electric field inside the capacitor is weakened.

To reduce the permittivity there are basically two possibilities:

- the polarizability of bonds inside the dielectric has to be decreased
- the quantity of bonds has to be reduced by introduction of porosity in the dielectric

The polarizability can be decreased by materials with less polar groups. Candidates are fluorined (FSG, ϵ_r 3.6) or organic (OSG) silicon oxides. However, for smaller and smaller structures this approach isn't sufficient, thus porous films have been introduced. By introduction of porosity there is "empty space" inside the ILD which has - in case of air - a permittivity of about 1, and therefore ϵ_r is reduced for the entire layer. The pores can be introduced by adding polymers which are expelled later by thermal annealing. In case of silicon dioxide one needs about 50 % of pores in the material to achieve a permittivity of 2. If a dielectric is used whose permittivity is 2.5 without porosity, only 22 % pores have to be introduced to achieve a permittivity of 2.

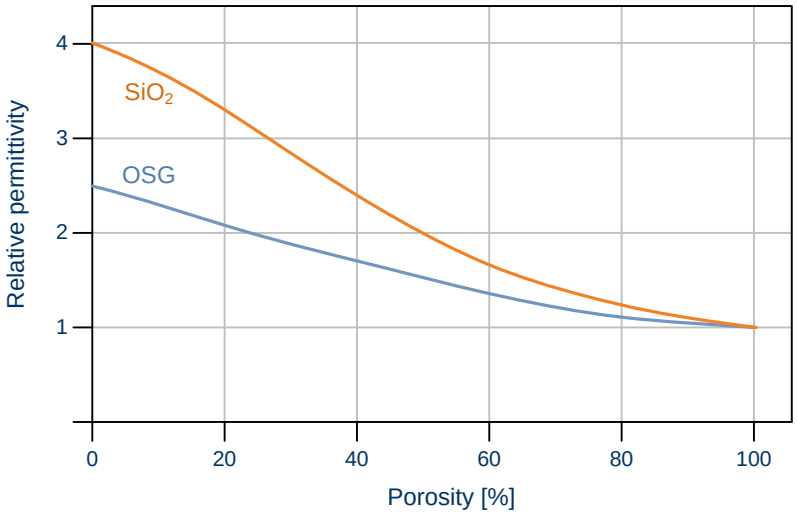


Fig. 1.13: Relative static permittivity against porosity

However, there are several issues to overcome, if one wants to bring in such new materials into semiconductor fabrication.

Due to porosity the density is reduced which results in a lower mechanical resistance. In addition process gases or copper can diffuse into the ILD and cause damage and thus increasing the permittivity or leakage. To counteract this issues, the pores have to be distributed evenly and must not be in contact with each other. To avoid a diffusion of copper, a thin barrier layer has to be deposited in a seperate process or the pores at the surface have to be closed by ion bombardment.

Like the photoresist used for manufacturing, the organic ILD is composed of hydrocarbon. If the resist is stripped in an ash process, the ILD is affected as well. To avoid this issue, additional layers (like silicon nitride as hardmask) have to be introduced in the film stack.

Chemical formula	**k value**
SiO_2	4,0
$SiO_{1,5}CH_3$	3,0
$SiO(CH_3)_2$	2,7
$SiO_{0,5}(CH_3)_3$	2,55

Tab. 1.1: Overview of organic silicon oxides

1.4 Metal semiconductor junction

1.4.1 Metal semiconductor junction

Subsequent to the manufacturing of transistors in the silicon substrate, the devices have to be connected to each other to realize an integrated circuit. The gate has to be contacted to control the current through the transistor, while the doped source and drain electrodes have to be contacted as well. This results in problems because of the doping. Both source and drain are impurified with dopants and thus contain additional charge carriers; electrons in n-doped crystals, holes in p-doped crystals.

Here the Fermi level is of interest. The Fermi level represents the highest energy level at which electrons can occur at absolute zero (-273,15 °C). In conductors there are electrons in the valence band and in the energetic higher conduction band, thus the Fermi level is at the level of the conduction band. For illustration: a see has a surface - Fermi energy - with water molecules - electrons - beneath it.

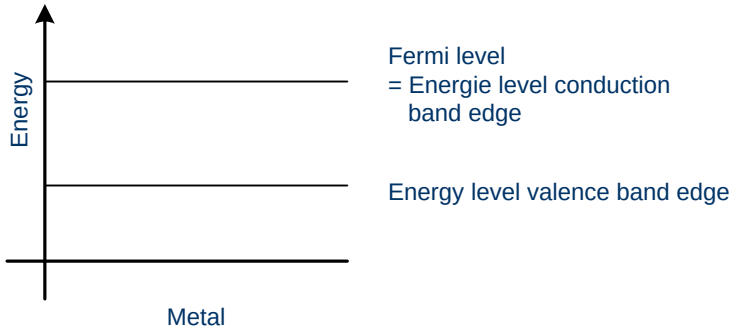


Fig. 1.14: Fermi level in metals

In doped semiconductors there are impurities as donors or acceptors in the lattice. In a p-doped semiconductor the Fermi level is next to the valence band, since electrons from the valence band can easily be lifted into the energy level of the dopant. According to this, the Fermi level in n-doped material lies next to the conduction band since electrons of donors can easily be lifted into the conduction band of the silicon crystal.

If one connects a metal to a semiconductor the Fermi levels of both materials have to equal, next to the interface the Fermi level is constant.

Because the conduction band in the semiconductor is energetic higher than the Fermi level, electrons flow into the metal since they always want to achieve the lowest energy

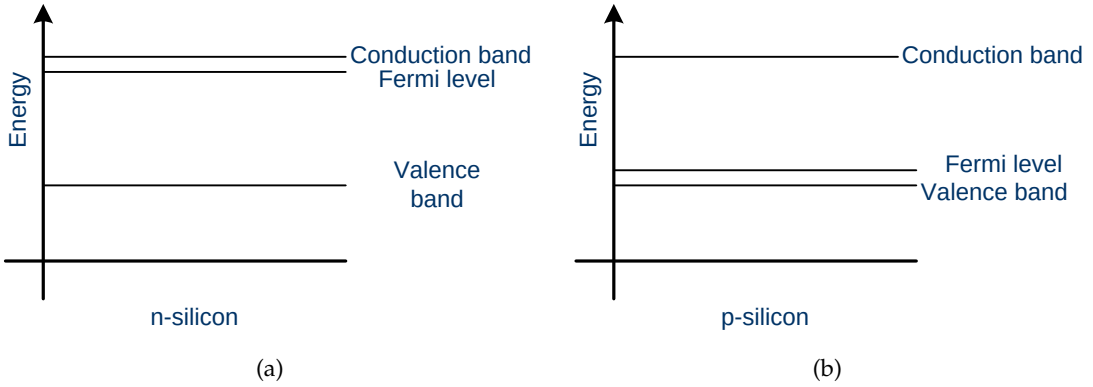


Fig. 1.15: Fermi level in doped semiconductors

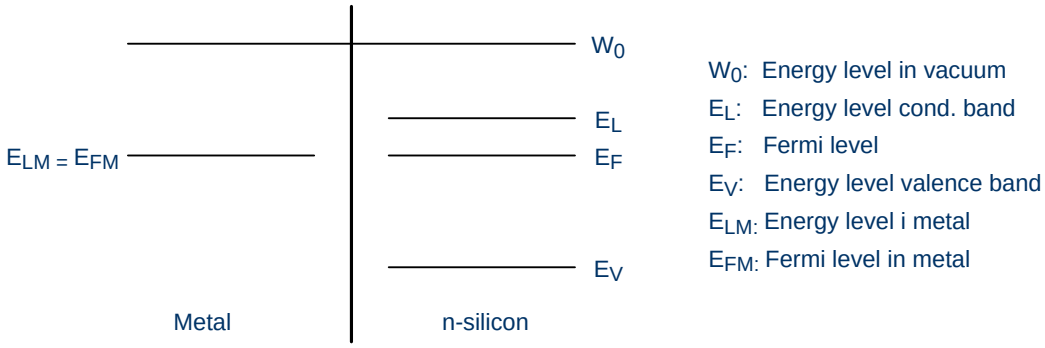


Fig. 1.16: Band model before the contact

state. Thus the probability density of electrons in the conduction band of the semiconductor decreases and therefore the distance between the conduction band and the Fermi level increases (the Fermi level represents the highest energy state at which electrons occur, and those flowed off). The electrons leave positively charged ions behind, and therefore a depletion zone remains. The bending of the energy band illustrates the potential barrier (Schottky barrier) which remaining electrons have to surpass to flow into the metal.

The width w of the depletion zone depends on the intensity of the doping. The migrated electrons lead to a negatively charged region in the metal which is limited to the surface.

This metal semiconductor junction results in a nonlinear current-voltage characteristic, a so-called Schottky diode. This barrier can be surpassed by electrons due to temperature or by tunneling due to an electric field.

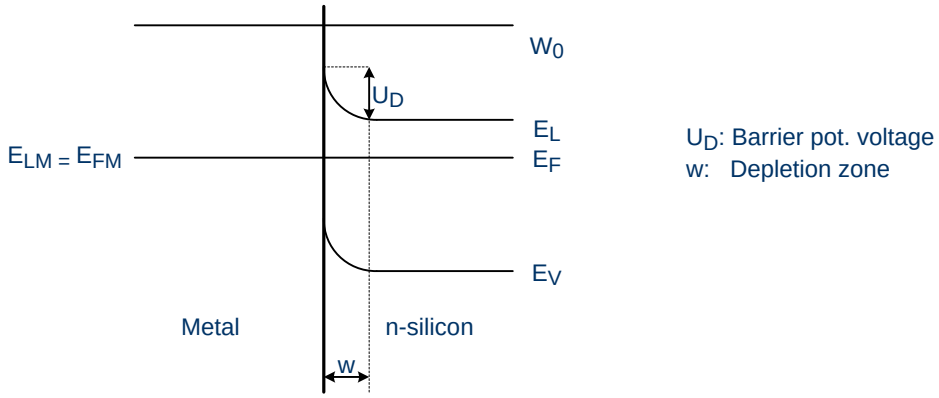


Fig. 1.17: Band model after the contact

Depending on the application this diode effect either is wanted or not. To achieve an ohmic contact (which means a contact without this potential barrier) the contact can be doped with high intensity, thus the width of the depletion zone decreases and the contact has a linear current-voltage characteristic because of tunneling.

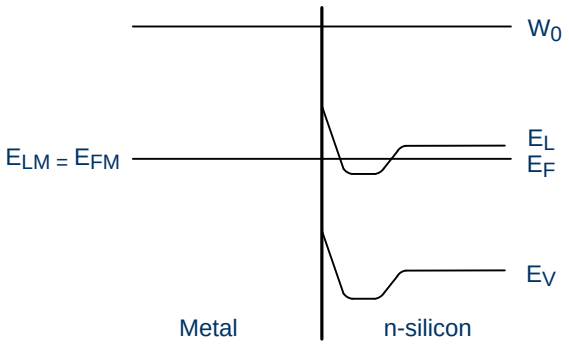


Fig. 1.18: Band model after n^+-doping

Because aluminum is integrated as an acceptor (picks up electrons) a p-doped interface occurs, leading to an ohmic contact in case of a p-doped semiconductor. In case of a n-doped semiconductor, however, aluminum leads to an inverted doping which results in a p-n junction: a diode. To avoid this there are two possibilities:

- the n-doped region is doped with such a high intensity that aluminum can only decrease the doping but not invert it
- an interlayer made of titan, chrome, or palladium prevents the inversion of the n-doped region

To enhance the contact, silicides (silicon on combination with metals) can be deposited

at the contact area.

In metal p-semiconductor junctions there is a band bending downwards, due to the exchance of charge carriers of the metal and the semiconductor. The problems mentioned above do not appear in this case because electrons from the energetic higher conduction band can flow into the metal continuously.

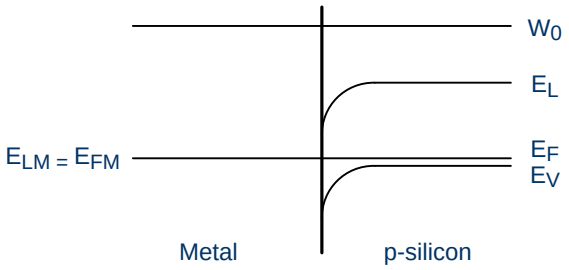


Fig. 1.19: Band model after the contact of metal and p-semiconductors

In contrast to the diode in the p-n junction, whose switching speed depends on the diffusion of electrons and holes, Schottky diodes have a very high switching speed. Thus they are suitable as protective diodes to inhibit voltage peaks.

1.4.2 Band model of p-n junctions

Because the Fermi level has to be constant, there is a band bending in p-n junctions as well. This bending illustrates the depletion zone which occurs because of the migrated charge carriers which is the potential barrier that prevents a further diffusion of electrons and holes in equilibrium (without an applied voltage). In silicon this potential barrier is about 0.7 V.

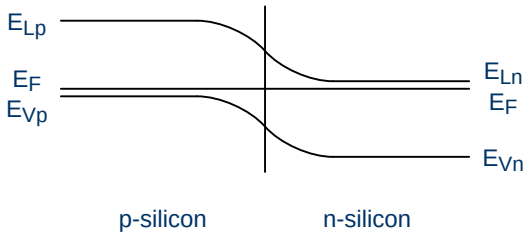


Fig. 1.20: Band model at the interface of n- and p-doped semiconductors

1.5 Wiring

1.5.1 Wiring

The wiring of an integrated circuit can take up to 80 % of the chip's surface, that's why techniques habe been developed to stack the wiring on top of the wafer in multiple layers. The amount of wires with only one additional layer can be reduced about 30 %.

Between the wires, isolation layers (oxide) are deposited, the metal layers are connected through vertical interconnect accesses (via). In today's microchips there are seven or more layers integrated. Edges and steps have to be rounded since the conformity of the metallization layers is not very good. This leads to bottlenecks in which current densities are increased so that electromigration occurs. To remove edges and steps there are several possibilities for planarization.

1.5.2 BPSG reflow

The reflow technique uses doped glasses like phosphorus silicate glass (PSG) or boron phosphorus silicate glass (BPSG). In a high temperature process the glasses melt and result in an uniform surface. Due to high temperature this technique can't be used for planarization of a metallization layer.

1.5.3 Reflow back etching

On top of the wafer a layer of silicon dioxide is deposited which is at least as thick as the highest step on the wafer. Next the oxide is coated with a resist or polyimide layer which is thermal treated for stabilization. In dry etching, the resist/polyimide and the silicon dioxide are removed with identical etch rates (selectivity of 1), thus resulting in a planished surface.

Besides the resist/polyimide, a so-called spin on glass (SOG) can be deposited on the wafer. Thus a planished layer can be produced which is stabilized during a post anneal step. An additional oxide layer is not necessary. However, all of these techniques can planish local steps only and are not sufficient for total leveling.

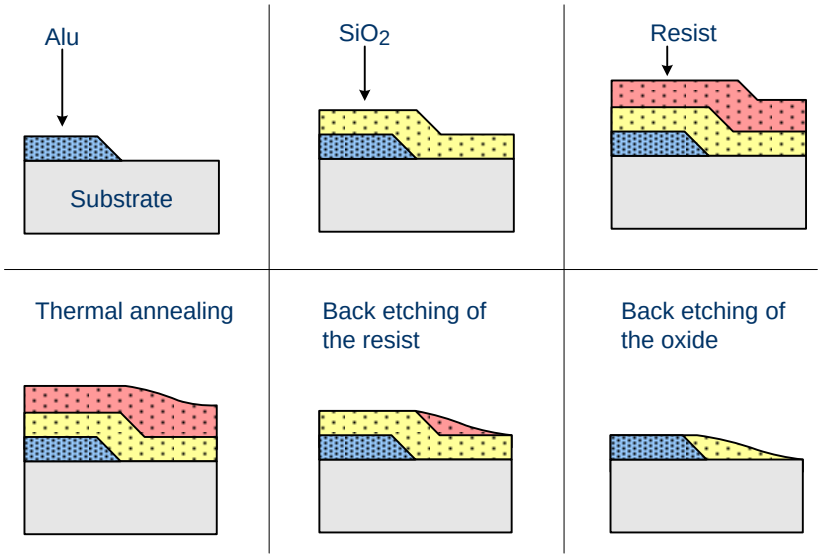


Fig. 1.21: Reflow back etching process

1.5.4 Chemical mechanical polishing

The chemical mechanical polishing/planarization (CMP) provides an uniform surface of the entire wafer. For this, an oxide is deposited on the wafer which is as thick as the highest step. The wafer is held upside down and pressed onto a polish plate. The wafer as well as the plate rotate in opposite directions and also move in horizontal directions. To support the process a slurry is used which contains abrasives and chemicals.

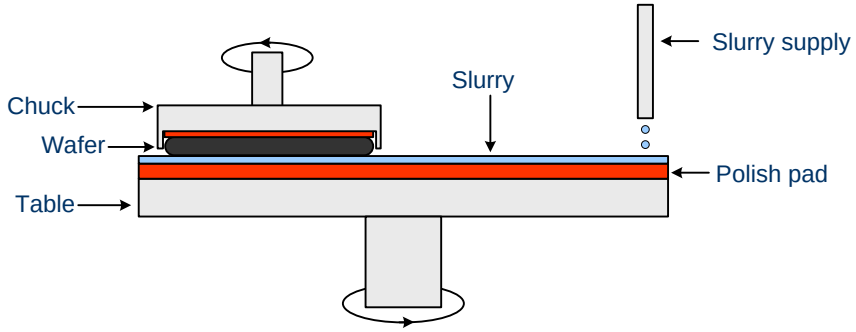


Fig. 1.22: Illustration of a polisher

Even if this process seems to be very rough it allows a surface which has an irregularity of only a few nanometers and thus is the optimal process for planarization.

1.5.5 Contacting

To contact the metallization layers, vias are etched into the isolation layers with high anisotropy. The vias have to be filled in this way, that an optimal contact is realized and the surface is not affected in a bad way.

For filling the vias, tungsten is the material of choice. With silane as additive a thin layer of tungsten is deposited as a seed layer in a CVD process with tungsten hexafluoride; byproducts as silicon tetrafluoride and hydrogen fluoride are exhausted:

$$4\,\mathrm{WF_6} + 3\,\mathrm{SiH_4} \longrightarrow 4\,\mathrm{W} + 3\,\mathrm{SiF_4} + 12\,\mathrm{HF}$$

With hydrogen as an additive to the tungsten hexafluoride the vias are filled thereafter:

$$\mathrm{WF_6} + 3\,\mathrm{H_2} \longrightarrow \mathrm{W} + 6\,\mathrm{HF}$$

On top if it the next metallization layer can be deposited, structured and planarized. If copper is used for wiring, tungsten will only be needed for the contact to the silicon substrate. The connection of the individual copper layers is done with copper itself.