Semiconductor Technology from A to Z

Deposition

Plasma, the fourth aggregation state of a material

Plasma state

In many semiconductor manufacturing processes a plasma is used, be it in sputtering, deposition processes, or dry etching. An important point here is that the wafer remains cold in the plasma. This may sound contradictory at first, since the electrons in the plasma are extraordinarily hot, at several tens of thousands of degrees. However, they are so light that upon collision they hardly transfer any energy to the heavy gas particles – the gas itself, and with it the wafer, therefore remain close to room temperature. Such a plasma, in which the electrons and the gas do not have the same temperature, is called a non-equilibrium plasma. Only for this reason can wafers that already carry a metallization be processed in a plasma at all.

Plasma is also referred to as the fourth state of matter. A state of matter is a qualitative condition of substances that depends on temperature and pressure. The three states solid, liquid, and gaseous are also encountered in everyday life. If the temperature is low, every atom in a substance is fixed rigidly at one point. Attractive forces prevent them from moving. At absolute zero (-273.15 °C) substances also do not undergo any reaction. As the temperature increases, the particles begin to oscillate, and the bonds between the atoms become more unstable. Once the melting point is reached, a substance transitions from the first to the second state of matter: ice (solid) turns into water (liquid).

The attractive forces in liquid substances are still present, but the particles can shift position and no longer have fixed places as in the solid state; the particles adapt, for example, to a given shape. If the temperature is increased further, the bonds break apart completely, and the particles move independently of one another. At the boiling point, a substance transitions from the second to the third state of matter: water (liquid) turns into water vapor (gaseous).

While the volume of solids and liquids is constant, gaseous substances completely fill the available space; the particles distribute themselves evenly throughout the entire volume.

Every substance has a very specific melting point and boiling point. Silicon melts at 1414 °C and transitions into the gaseous state at 2900 °C. If even more energy is supplied to a substance, collisions between the particles knock electrons out of the outermost electron shell. Free electrons and positively charged ions are now present in the space: the plasma state has been reached.

Plasma generation

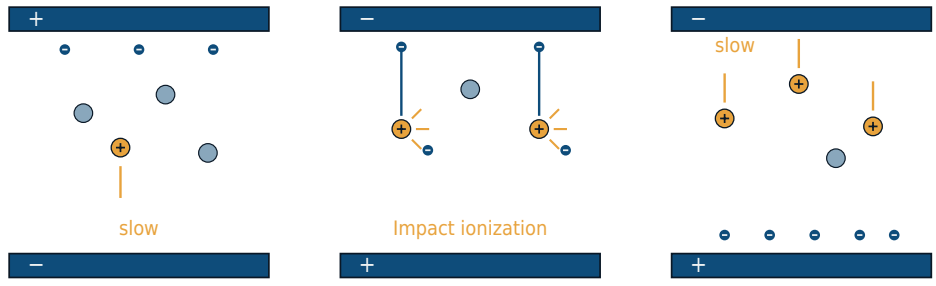
A plasma in semiconductor technology is usually generated by a high-frequency voltage; argon, for example, serves as the gas. The gas is located in a high-frequency field between two oppositely charged plates (electrodes) and is ionized there. This requires electrons that knock further electrons out of the argon atoms. These initial electrons can be generated in different ways:

- electrons are emitted from a thermionic cathode
- electrons are pulled out of the negative electrode by a very high voltage
- in every gas, brief free electrons occur through collisions of the particles

Since the electrons are much lighter than the ions, they are immediately attracted to the positively charged electrode; the heavy ions move slowly toward the negative electrode. However, before they reach it, the polarity of the electrodes is reversed; the electrons are accelerated toward the other electrode and, on their trajectory, knock further electrons out of atoms through collisions. Typical frequencies for plasma generation are 13.56 megahertz and 2.45 gigahertz, meaning the voltage at the electrodes is reversed 13.56 million or 2.45 billion times per second, respectively.

How the Plasma Ignites

The electrodes are reversed millions of times per second



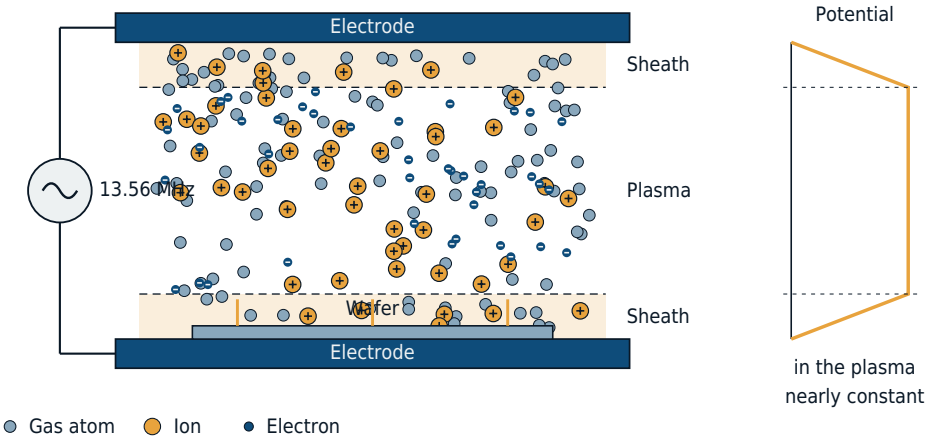
The light electrons follow the alternating field, the heavy ions barely do.
Each polarity reversal creates new charge carriers, until the plasma ignites.

Inside the plasma, electrons and ions are roughly balanced, so it is electrically neutral overall. The situation is different immediately in front of the electrodes: the light, fast electrons reach there much more often than the sluggish ions, which cannot follow the rapid voltage changes. As a result, the electrode charges negatively, and an electron-depleted boundary layer (sheath) forms in

front of it. Almost the entire voltage drops across this sheath – it is this voltage that accelerates the ions toward the wafer and thus accounts for the directional component of dry etching.

Plasma Generation Between Two Electrodes

Charges balance out in the bulk, but not in front of the electrodes



The electrons follow the alternating field, the heavy ions do not.
This depletes electrons in front of the electrodes; almost the entire voltage drops across this sheath and accelerates the ions onto the wafer.

Plasma generation takes place under vacuum; the plasma produced is not heated, which is important for many processes. Depending on the gas and the equipment in which the plasma is generated, it can be used in deposition, sputtering, etching, or also in ion implantation. Due to the rapid oscillations of the positive ions in the high-frequency field, they are very energetic. The plasma does not contain only positive ions and free electrons, since other particles are also created through collisions; the state of the plasma changes continuously. Electrons are partially recaptured by the ions and knocked out again; however, these additional particles play no role in the further use of the plasma. Depending on the particle density in the plasma (10^8–10^{12} particles per cm^3), the degree of ionization is 0.001–10 %, meaning the majority of the particles are uncharged.

The simple arrangement with two plates has one drawback: the same voltage determines both how many ions are generated and how hard they strike the wafer. The two cannot be set independently of one another. Equipment for fine structures therefore generates the plasma differently – via a coil that couples in an alternating field from outside, or via microwaves – and applies the voltage at the wafer electrode independently of this. Ion density and ion energy are then two separate, independently adjustable parameters.

Chemical vapor deposition

Silicon vapor phase epitaxy

Epitaxy means "on top" or "arranged upon," and represents a process in which a layer is created on top of another layer and inherits its crystal structure. If the deposited layer is of the same material as the substrate, one speaks of homoepitaxy; with two different materials, of heteroepitaxy. The most significant case of homoepitaxy is the deposition of silicon on silicon; in heteroepitaxy, a different material grows on top, such as silicon-germanium on silicon or gallium nitride on silicon and sapphire. A crystalline base is required: nothing single-crystalline grows on an amorphous layer such as silicon dioxide. Silicon-on-insulator wafers (SOI) are therefore not created by epitaxy, but rather by bonding two wafers together and removing one of them down to a thin layer.

Homoepitaxy

Depending on the process, the wafers can already be delivered by the manufacturer with an epitaxial layer (e.g. in CMOS technology), or the chip manufacturer has to carry this out themselves (e.g. in bipolar technology).

As gases for generating the layer, pure hydrogen is used in combination with silane (SiH_4), dichlorosilane (SiH_2Cl_2), or silicon tetrachloride ($SiCl_4$). At about 1000 °C, the gases cleave off silicon, which deposits on the wafer surface. The silicon adopts the structure of the substrate and, for energetic reasons, grows plane by plane in succession. To prevent the silicon from growing polycrystalline, there must always be a shortage of silicon atoms present, i.e. slightly less silicon must always be available than could actually grow. With silicon tetrachloride, the reaction proceeds in two steps:

$SiCl_4 + H_2 \rightarrow SiCl_2 + 2HCl$
$2\ SiCl_2 \rightarrow Si + SiCl_4$

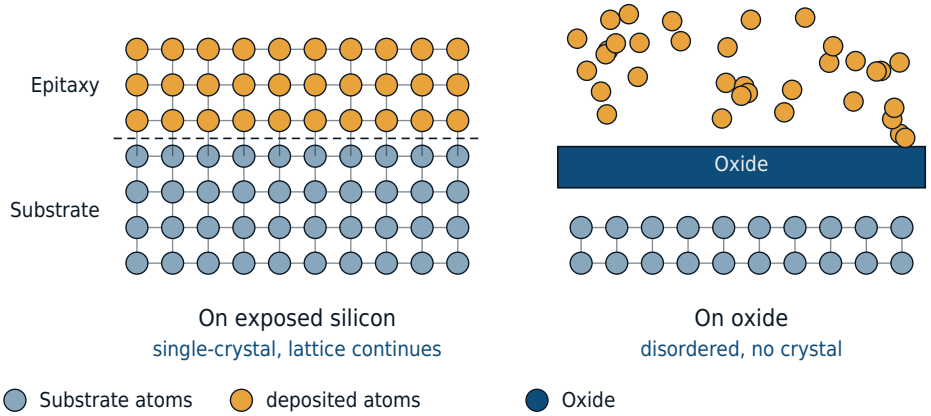
In order for the silicon on the substrate to adopt the crystal structure, the surface must be absolutely clean; this makes use of the equilibrium reaction. Both reactions can also proceed in the opposite direction, depending on the ratio of the gases used. If there is only little hydrogen present in the atmosphere, silicon is removed from the wafer surface due to the high chlorine concentration – as in the trichlorosilane process used to purify silicon. Only with an increasing concentration of hydrogen is growth achieved.

With $SiCl_4$, the growth rate is about 1–2 µm per minute. Since single-crystalline silicon only grows on the cleaned surface, certain areas can be masked with oxide, on which polycrystalline silicon then grows. Compared to single-

crystalline silicon, however, this is etched away very easily by the backward-running reaction. If diborane (B_2H_6) or phosphine (PH_3) are added to the process gases, doped layers can be produced, since the doping gases decompose at the high temperatures and the dopants are incorporated into the crystal lattice.

The process for producing homoepitaxial layers is carried out under vacuum. The process chamber is first heated to 1200 °C so that the native oxide, which always forms on the silicon surface, is volatilized. As mentioned above, too low a hydrogen concentration leads to a back-etching of the wafers. This is exploited prior to the actual process to clean the wafer surface in this way. By changing the gas concentrations, deposition then takes place in an epitaxy reactor.

Epitaxy: the layer adopts the crystal structure
It grows plane by plane, continuing the lattice of the surface below



This is exactly what selective epitaxy relies on: at the right chlorine level,
the disordered material is continuously etched away again, while the single-crystal material remains.

Selective epitaxy and strained layers

The circumstance described above – that single-crystalline silicon only grows on exposed silicon, while polycrystalline material is removed again by the backward-running reaction – is no longer merely a side effect today, but the basis of a process in its own right. With a suitably chosen chlorine content, the layer grows exclusively in the opened windows and not at all on oxide or nitride – this is referred to as selective epitaxy.

Its most important application lies not in adding to the silicon, but in straining it. If a silicon-germanium alloy is grown into the source and drain regions of a transistor instead of silicon, this alloy requires more space than the lattice

allows, due to the larger germanium atoms. As a result, it presses on the channel located between them. In a lattice compressed in this way, holes move noticeably more easily, and the transistor switches faster. For the opposite case – tensile strain on the channel, favorable for electrons – silicon-carbon is used instead. Since the mid-2000s, these strained layers have been part of the standard structure of every high-performance transistor.

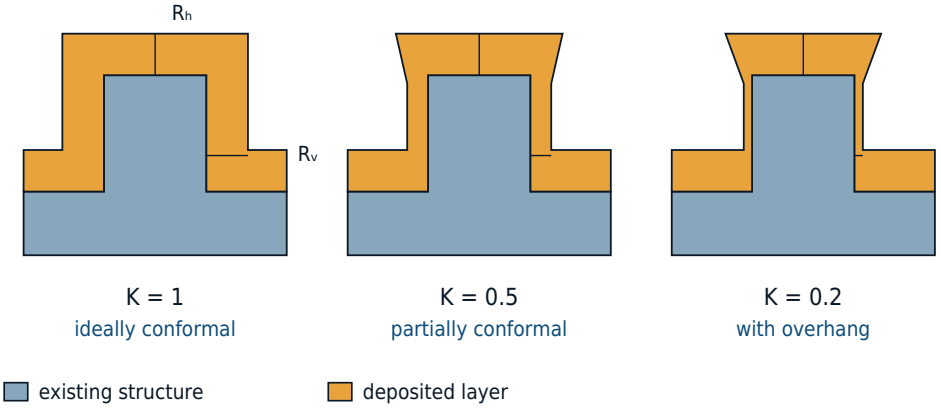
CVD process: Chemical Vapor Deposition

Semiconductor technology often requires layers that cannot be produced from the silicon substrate itself. For silicon nitride and silicon oxynitride, for example, the layer has to be created through thermal decomposition of gases that contain all the required materials, such as silicon. This is the principle underlying chemical vapor deposition, or CVD for short. The wafer merely serves as the base surface and does not react with the gases. Depending on pressure and temperature, the CVD process is divided into different methods, whose resulting layers differ in density and edge coverage. If the layer growth on vertical edges is exactly as high as on horizontal surfaces, the deposition is conformal.

The conformity K is the ratio of the layer growth on vertical surfaces R_v to the growth on horizontal surfaces R_h: $K = R_v : R_h$. If the deposition is not ideally conformal, K is noticeably less than 1 (e.g. $R_v : R_h = 1:2 \rightarrow K = 0.5$). High conformities can only be achieved through high process temperatures and low pressures. A conformity of exactly 1 is only achieved by atomic layer deposition, because there it is not surface attachment but a self-limiting reaction that determines the layer thickness.

Conformality of a Deposition

Ratio of growth on vertical and horizontal surfaces



R_h is the growth on horizontal, R_v that on vertical surfaces.
The smaller K is, the more readily the layer grows together at the edge and seals in a void beneath it.

APCVD: Atmospheric Pressure CVD

APCVD is a CVD deposition at normal pressure (atmospheric pressure) which is used for producing doped and undoped oxides. The resulting oxide has a low density, and edge coverage is moderate due to the low temperature. The major advantage is throughput: without a vacuum, wafers can be processed in a continuous flow, and the deposition rate is high. For thick, non-critical oxides, the process therefore remains economical; for thin or conformal layers, it is not an option.

The process gases used are silane SiH_4 (highly diluted with nitrogen N_2) and oxygen O_2, which are thermally decomposed at about 400 °C and react with each other.

$SiH_4 + O_2 \rightarrow SiO_2 + 2H_2$ (T = 430 °C, p = 10^5 Pa)

By adding ozone O_3, better conformity can be achieved, since it increases the mobility of the depositing particles. The oxide is porous and electrically unstable and can be densified through a high-temperature step.

To prevent the formation of edges, which could cause problems when depositing further layers, so-called phosphorus silicate glass (PSG) is used for layers serving as interlayer dielectric. For this purpose, phosphine PH_3 is additionally added to the two gases SiH_4 and O_2, so that the resulting oxide contains 4–8 % phosphorus. A high phosphorus content significantly improves flow properties;

however, phosphoric acid can form, which corrodes aluminum (interconnects).

Since annealing steps affect preceding processes (e.g. doping), only brief annealing using powerful argon lamps (several hundred kilowatts, <10 s, T = 1100 °C) is used to reflow the PSG, instead of long furnace processes.

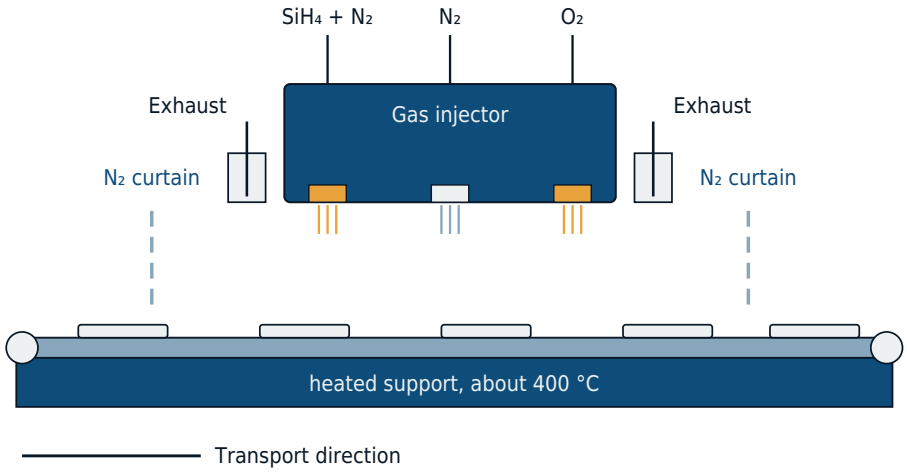
Analogous to phosphosilicate glass, boron can also be added at the same time (borophosphosilicate glass, BPSG, 4 % B and 4 % P).

The purpose of these doped glasses was to reflow at high temperature, thereby leveling out steps. This task has since been taken over by chemical mechanical polishing, which requires no thermal load and also planarizes over large areas. Doped glasses are still used today mainly as interlayer dielectric beneath the first metal layer, where the phosphorus additionally binds alkali ions and keeps them away from the transistor.

Illustration of a horizontal reactor for APCVD deposition

Horizontal Reactor for APCVD

The wafers pass under the gas injector at atmospheric pressure



At atmospheric pressure no load lock is needed: the wafers pass through continuously, which is what gives the process its high throughput.

Silane and oxygen are fed in separately and only brought together above the wafer, so they don't react inside the injector itself.

LPCVD: Low Pressure CVD

In the LPCVD process, the reaction takes place under vacuum, allowing the deposition of thin layers such as silicon nitride (Si_3N_4), silicon oxynitride (SiON), silicon dioxide (SiO_2), and tungsten (W). LPCVD processes enable very good

conformities of nearly 1; the reason for this is the low pressure of only 10–100 Pa (normal air pressure is about 100,000 Pa), which means the particles do not experience directional movement but instead spread out through mutual collisions, so that edges on the wafer surface are covered evenly by all gas particles. The process temperature of up to 900 °C also contributes to the high conformity. In contrast to the APCVD process, the density and stability are very high.

The reactions for Si_3N_4, SiON, SiO_2, and tungsten proceed according to the following reaction equations:

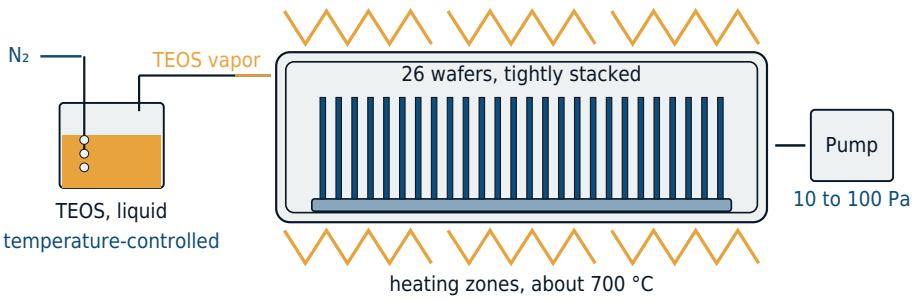
a) Si_3N_4 (850 °C): $4NH_3 + 3SiH_2Cl_2 \rightarrow Si_3N_4 + 6HCl + 6H_2$
b) SiON (900 °C): $NH_3 + SiH_2Cl_2 + N_2O \rightarrow SiO_xN_y$ + byproducts
c) SiO_2 (700 °C): $SiO_4C_8H_{20} \rightarrow SiO_2$ + byproducts
d) Tungsten (400 °C): $WF_6 + 3H_2 \rightarrow W + 6HF$

In contrast to the gaseous precursors used for Si_3N_4, SiON, and tungsten, liquid tetraethyl orthosilicate (TEOS) is used to produce SiO_2; in addition, there are other liquid sources such as ditertiarybutylsilane (DTBS, $SiH_2C_8H_{20}$) or tetramethylcyclotetrasiloxane ($Si_4O_4C_4H_{16}$).

A tungsten layer can only be produced on silicon. Therefore, if no silicon is available as a base surface, silane has to be added.

LPCVD System for TEOS Layers

TEOS is liquid at room temperature and must first be vaporized



The low pressure lets the gas molecules travel far and reach every gap. That's why the wafers can stand close together, and the edges are covered evenly: the conformality is close to 1.

PECVD: Plasma Enhanced CVD

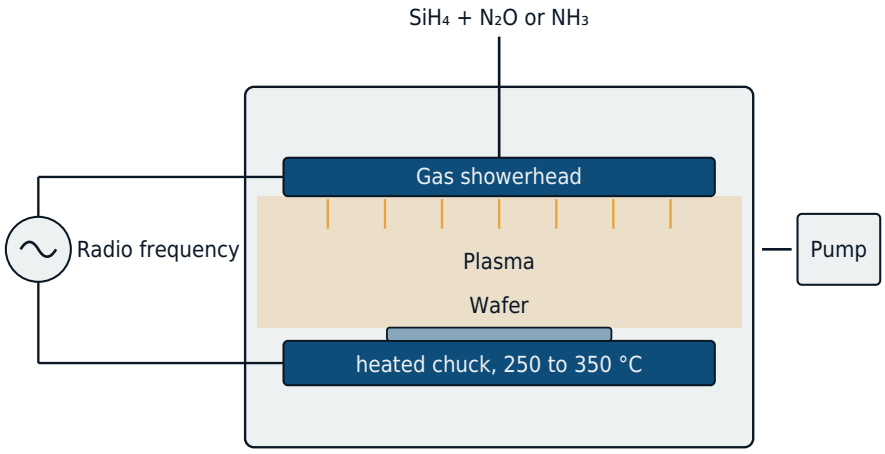
The plasma-enhanced CVD process takes place at 250–350 °C. Since the temperature is too low to decompose the gases, a high-frequency voltage is used to bring the gas into the plasma state. In the plasma state it is very energetic

and then deposits onto the wafer. In particular, aluminum interconnects cannot be exposed to high temperatures (aluminum melts at 660 °C, but the thermal budget of finished aluminum interconnects is already exhausted at about 450 °C), which is why the PECVD process is used above all here for depositing SiO_2 and Si_3N_4. Instead of dichlorosilane SiH_2Cl_2, as used in the LPCVD process, silane SiH_4 is employed, since it decomposes more readily at low temperatures. The conformity is not as ideal as with the LPCVD process (about 0.6–0.8), but the deposition rate, at 500 nm per minute, is significantly higher.

This combination – low temperature at a high rate – accounts for the process's significance today. Everything deposited after the first metal layer is produced by PECVD: the interlayer dielectrics between the wiring levels, the nitride layers that serve as etch stops and as diffusion barriers against copper, the hard masks used for patterning, and the final passivation of the finished chip. The carbon-containing low-k dielectrics are also produced this way, by adding an organosilicon compound to the process gas.

PECVD System

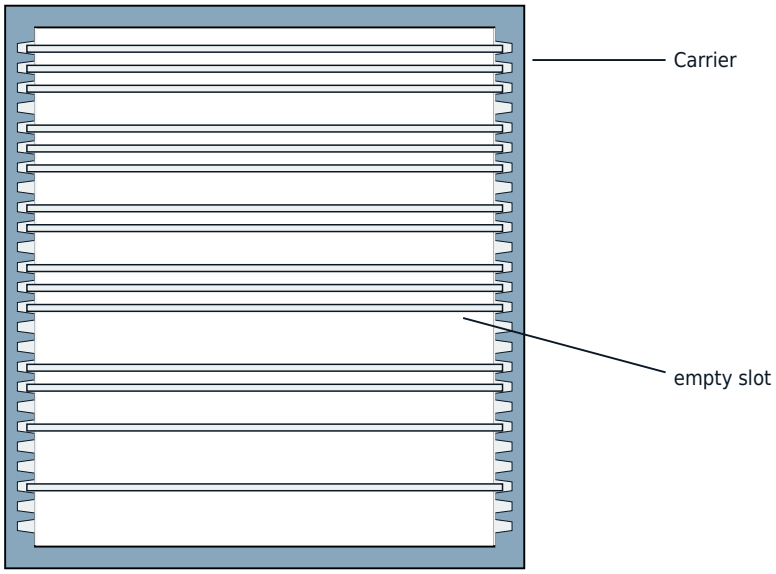
The plasma supplies the energy that would otherwise have to come from temperature



At 250 to 350 °C, the gases would not decompose on their own. That's why the process can also be used over finished aluminum interconnects – where a furnace process would already be far too hot.

Wafer Carrier

Side view: the wafers lie horizontally, stacked on top of each other



The carrier holds 25 wafers; here 15 of them are loaded. The slot is taller than the wafer is thick and widens toward the opening, so the wafer slides in easily.

ALD: Atomic Layer deposition

Atomic Layer Deposition (ALD) is a CVD deposition process for producing thin layers. It is a cyclical process in which several gases are introduced into the process chamber alternately.

Each gas reacts in such a way that the respective surface atoms become saturated, meaning the reaction is self-limiting. The other gas can then continue to react at this new surface. Alternating with the reactive gases, the chamber is purged with inert gases such as argon or nitrogen. A simple ALD cycle, which usually lasts only a few seconds, can proceed, for example, as follows:

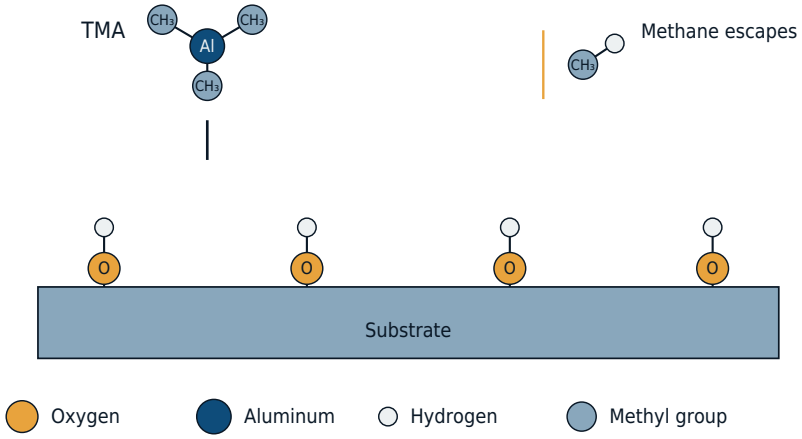
1. self-limiting reaction with the first gas, which saturates the surface atoms
2. purge step with an inert gas
3. self-limiting reaction with a second gas, which reacts at the new surface
4. purge step with an inert gas

A specific example is the deposition of an aluminum oxide layer Al_2O_3, in which trimethylaluminum C_3H_9Al (TMA) and water H_2O are used as the reactive gases.

In the first step, a methyl group CH_3 of the aluminum compound removes near-surface hydrogen atoms from OH groups, forming methane CH_4. The remaining molecules bond to the now unsaturated oxygen atoms.

1. Trimethylaluminum meets the surface

The methyl groups take the hydrogen from the OH groups

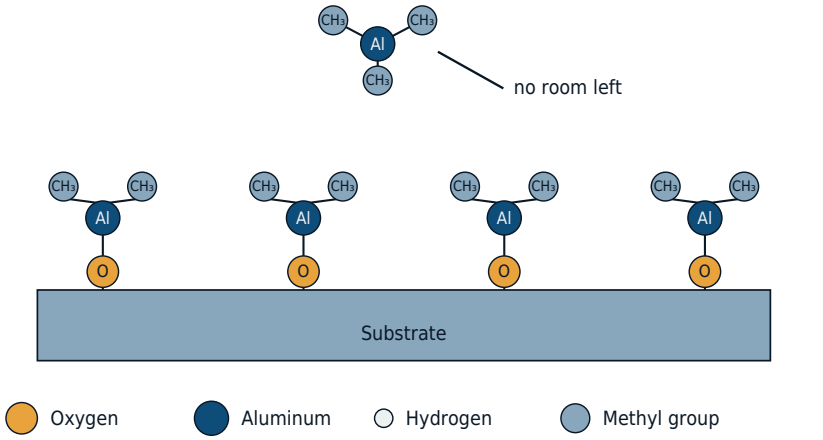


Each methyl group binds one hydrogen atom and leaves the surface as methane.

Once these atoms are saturated, no further attachment of TMA can occur.

2. The surface is saturated

No room left for more TMA – the reaction limits itself

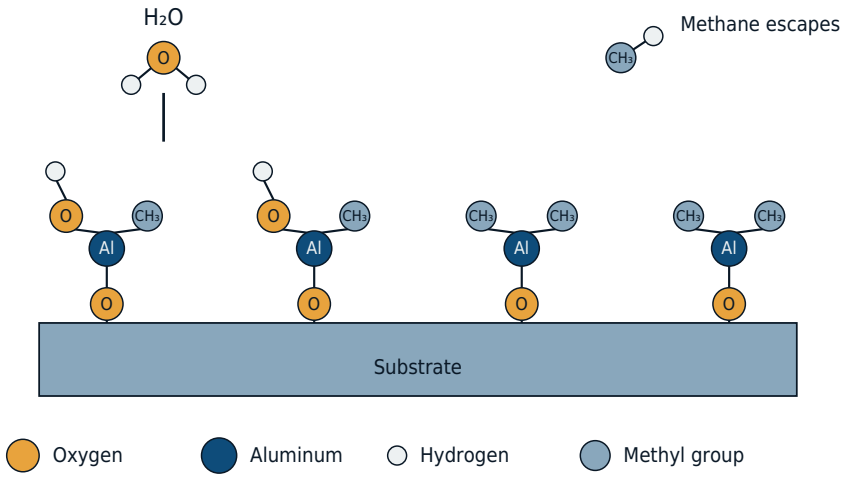


This is exactly what makes the process so precise: the layer grows only one layer at a time.

The process chamber is purged, and water vapor is subsequently introduced into the chamber. One hydrogen atom from each H_2O molecule removes the CH_3 groups of the previously deposited layer, forming methane.

3. Water vapor follows the purge

The water replaces the methyl groups, again forming methane

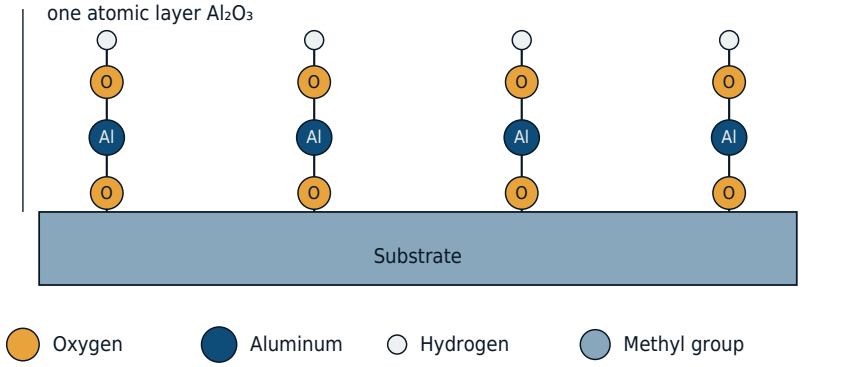


One hydrogen from the water binds the methyl group, leaving the oxygen behind.

What remains are OH^- ions, which now bond with the aluminum, so that once again near-surface hydrogen atoms are available to react with TMA.

4. One layer of aluminum oxide has grown

The new OH groups are the starting point for the next cycle



About a tenth of a nanometer grows per cycle – the thickness follows from the number of cycles.

Atomic layer deposition offers substantial advantages over other deposition techniques, which is why it is an important process for applying various layers. Even 3-dimensional structures (e.g. trenches) can be coated very uniformly. Both insulating and conductive films are possible on various substrates (semiconductors, polymers, etc.). Through the number of cycles, an exact layer thickness is easy to achieve. Since the reactive gases are not introduced into the chamber simultaneously, nucleation cannot occur before the actual deposition

takes place. The quality of the layer is therefore very high.

What ALD is used for today

The price for this precision is speed: only about one-tenth of a nanometer grows per cycle, and a cycle takes seconds. The process is therefore unsuitable for thick layers. Wherever only a few atomic layers matter, however, it is unrivaled:

- Gate dielectric: The silicon dioxide beneath the gate could not be thinned any further without current tunneling through it. It was replaced by hafnium dioxide, which, for the same effect, may be thicker. It is deposited via ALD – otherwise a layer of just a few atomic layers cannot be produced uniformly enough across the entire wafer.
- Spacer layers in multipatterning: In the self-aligned process, the thickness of a deposited layer determines the resulting feature width. It must therefore be adjustable more precisely than any exposure step could ever be.
- Barriers and seed layers: in deep, narrow contact holes, which no other process can line uniformly.
- Storage capacitors: the capacitors of a DRAM cell are narrow, deep trenches whose walls must be completely coated.

If the temperature must remain low, the second reaction step is assisted by a plasma (plasma-enhanced ALD). The radicals from the plasma react more readily than the neutral gas, allowing the cycle to proceed even well below the temperature otherwise required.

Gap fill

One task arises with every deposition process, and none of the methods described so far solves it on its own: filling narrow, deep gaps. Such gaps occur wherever an insulator has to be introduced between structures that already exist – between the trenches of trench isolation, between closely spaced interconnects, between the fins of a transistor.

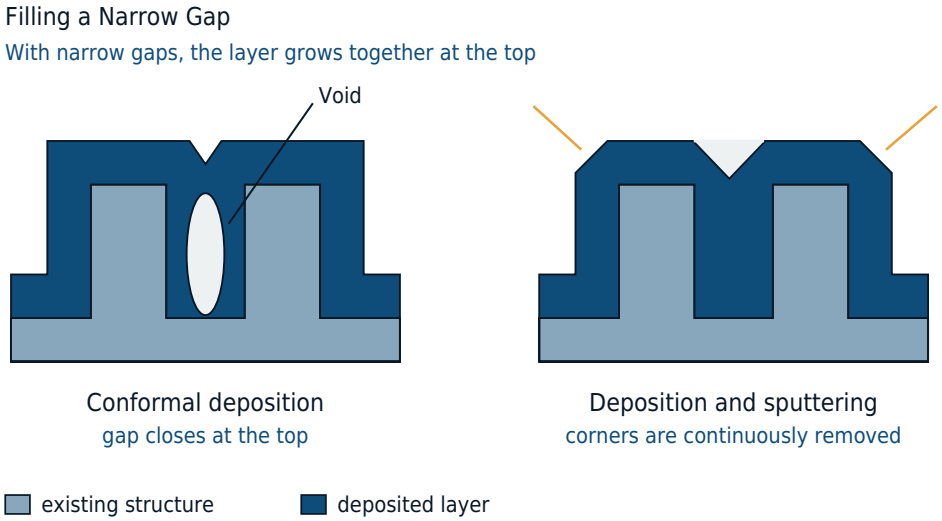
Why a void forms

A conformal layer grows at the same rate on every surface, including the two opposing gap walls. If the gap is narrower than twice the layer thickness, the two fronts meet. That alone would be unproblematic if they met everywhere at the same time. In reality, the layer grows faster at the gap opening, because the opening is accessible from more directions than the bottom. The gap therefore closes at the top first, leaving a void behind underneath.

Such a void is not merely a cosmetic problem. If it is exposed during a later step, for example during polishing or while etching a contact hole, the void fills with

etch solution or with metal – in an unfavorable case, this creates a short circuit between two interconnects.

Depositing and sputtering at the same time



The solution is to continuously remove the corners at the gap opening again while the layer grows. For this purpose, a sputter component is superimposed on the deposition: argon ions are accelerated from a dense plasma onto the wafer and knock material out there. This is the same process used in sputtering for coating – only here the ion beam does not strike a target, but the growing layer itself.

Because sputtering is most effective at oblique incidence, it attacks the edges at the gap opening much more strongly than the horizontal surface at the bottom. This creates a sloped surface at the top edge, referred to as a facet. The opening remains open, and the gap fills from the bottom upward.

Processes of this kind are known as HDP-CVD (high density plasma CVD). The decisive process parameter is the ratio of deposition rate to sputter rate, referred to as the deposition/sputter ratio, or D/S for short: too little sputtering allows the void to form, too much removes the structures themselves. The sputter component is typically 10 to 20 percent of the net deposition rate.

When even this is no longer enough

As depth-to-width ratios continue to increase, even this process reaches a limit. The next step returns to the oldest principle of all: a liquid flows into any gap by itself, regardless of its shape. For this purpose, a silicon-containing precursor is applied that is initially liquid, fills the gap completely, and is only afterward

converted into solid silicon dioxide through treatment with water vapor and temperature. The resulting layer is less dense than a deposited oxide, but it fills gaps that are inaccessible to any directional process.

Physical deposition methods

MBE: Molecular beam epitaxy

The growth behavior in molecular beam epitaxy (MBE) corresponds to that of silicon vapor phase epitaxy. However, in this case the silicon is deposited onto the wafer in a different way.

The process takes place under ultra-high vacuum (UHV, 10^{-8} Pa); the wafer to be processed is held upside down at the top of the process chamber and heated to about 600–800 °C so that the native oxide volatilizes.

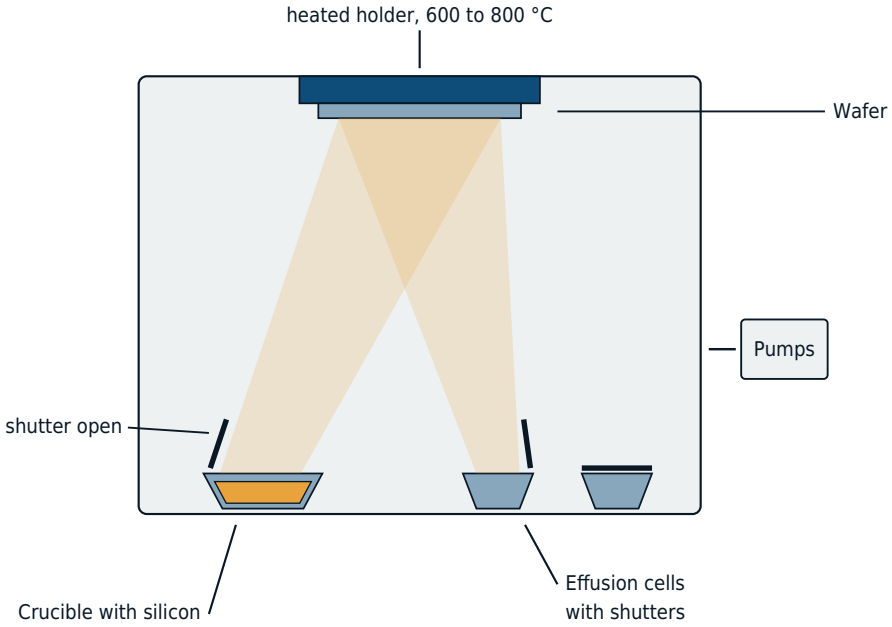
Silicon is evaporated using an electron beam in a crucible and then deposits onto the wafer. Dopants can be evaporated in an effusion source and thus reach the wafer together with the silicon vapor. Through targeted temperature control and the use of shutters, the particle beam can be precisely dosed. This process also makes it possible to grow layers of different materials on top of one another that would otherwise not be possible due to differing atomic sizes. This allows a layer of silicon and germanium to be produced, which is needed for high-frequency applications in bipolar technology.

This is precisely where the strength of the process lies: because the atoms arrive individually and without a chemical reaction, the composition can be changed from one atomic layer to the next. For layer sequences made of compound semiconductors, as needed in laser diodes and high-frequency devices, this is indispensable. In silicon manufacturing, on the other hand, MBE plays no role – there it is too slow.

However, the effort required to generate the UHV in this epitaxy process is very high. To reach this pressure, which is less than one-trillionth of normal atmospheric pressure, several pumps must be combined and the process chamber pumped down for a long time. Only one wafer can be processed at a time, and growth amounts to only about 1 µm per hour.

MBE System

The atoms fly individually through the ultra-high vacuum onto the wafer



Each source can be switched on individually; growth is about one micrometer per hour. This is the only way to change the composition from one atomic layer to the next. The pressure is around 10^{-8} Pa.

Evaporating

In evaporation, metallic layers, such as aluminum, can be produced on the wafers. The material is placed in a crucible made of a high-melting-point metal such as tantalum and heated there until it transitions into the gaseous state. The metal vapor strikes the wafer perpendicularly, so that edges are not covered well; the resulting layer is polycrystalline. As an alternative to melting in a crucible, the metal can also be evaporated using an electron beam. Compared to thermal evaporation, this electron beam method makes it possible to achieve a very precise deposition rate. Since edge coverage is not very good, these two processes are mostly used for full-area backside coating for later electrical contacting.

In the manufacturing of integrated circuits, evaporation has largely been displaced by sputtering, which provides denser layers and better edge coverage. It has been retained where poor edge coverage is precisely what is desired: in the lift-off process, a resist mask is first patterned, then the metal is evaporated, and afterward the resist together with the metal lying on top of it is removed. Because the sidewalls of the resist mask are barely covered, the film breaks cleanly at the edge. This makes it possible to pattern metals that are difficult to etch.

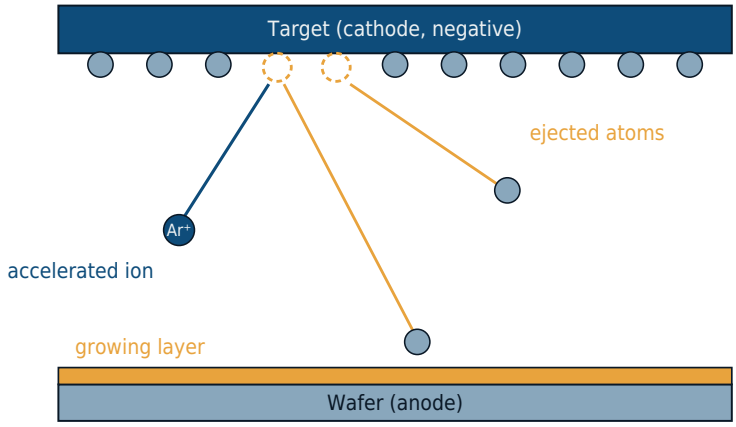
Schematic illustration of an evaporation system

Sputtering

In sputtering, ions (usually argon) are accelerated onto a target and strike out atoms or molecules there; the target consists of the material of the layer to be deposited. The mean free path of these particles is a few millimeters long, i.e. they collide frequently, which means that vertical surfaces on the wafer are also covered well. The deposited particles form a porous layer, which can be densified by annealing. Sputtering can be divided into passive (inert) and reactive sputtering.

Sputtering

Argon ions knock atoms out of the target, which deposit on the wafer



The atoms fly off in all directions and collide with gas particles along the way. That's why even vertical surfaces are well covered – the layer stays porous at first and is later densified by annealing.

In passive sputtering, only the material of the target is deposited on the wafers; depending on the target material, highly pure layers can be produced, since a precise mixing ratio of the substances in the target is possible. In reactive sputtering, a reactive gas (e.g. oxygen O_2) is added to the gas in the process chamber, which combines with the sputtered material from the target and then deposits on the wafer. This makes it possible to produce insulating layers from a metal target (e.g. aluminum Al):

$4\ Al + 3\ O_2 \rightarrow 2\ Al_2O_3$

To produce metallic layers, DC sputtering is used. Here, the ions are accelerated

onto the target with up to 3 kilovolts (kV) and discharge there. Since these charges always have to be dissipated, only a conductive material can serve as the target here. For insulating layers, reactive sputtering must be used. If an insulating layer is to be produced directly from the target, radio-frequency sputtering (RF sputtering) is used.

In RF sputtering, the voltage is applied to one electrode each behind the target (cathode) and the wafer (anode). Due to the high-frequency voltage, electrons are attracted to the target during the positive half-wave, causing the target to charge negatively. This negatively charged target attracts ions, which knock particles out of it. In magnetron sputtering, magnets are additionally placed behind the target so that electrons are deflected into circular paths and can thus ionize argon atoms more frequently, causing a considerable increase in the deposition rate. Since the anode is connected to the process chamber, its potential difference relative to the plasma, averaged over time and referenced to the surface, is substantially smaller than that of the cathode, which is why the ions migrate only toward the target and not toward the wafer.

To increase edge coverage, BIAS sputtering is used, in which a negative voltage is applied to the substrate. As a result, particles of the deposited layer are removed here as well, just as at the target, and the surface is planarized. However, care must be taken that no removal of the substrate itself occurs. This so-called back-etching is also the principle behind most plasma etching systems.

Sputtering into narrow structures

The particles knocked out are uncharged and fly off in all directions. Over a narrow, deep contact hole, this means: most of them strike the edge of the opening, and only a little reaches the bottom. The opening closes up before the bottom is covered.

A remedy is to ionize the sputtered particles themselves. In a dense plasma between the target and the wafer, they lose an electron and can then be accelerated straight down by a voltage applied to the wafer. This directs the material flow and allows it to reach the bottom of deep holes as well. In this way, the tantalum barrier and the copper seed layer are deposited in the damascene process – the seed layer must be present without gaps, on the sidewalls just as much as on the bottom, otherwise the copper will not grow continuously during the subsequent electroplating.

Sputtering is therefore well suited for producing metallic layers with good conformity and very good reproducibility. The effort involved is low, and the reduced pressure of about 5 Pa is fairly easy to generate.