Semiconductor Device Fabrication Process

Val Bradford

First Edition, 2011

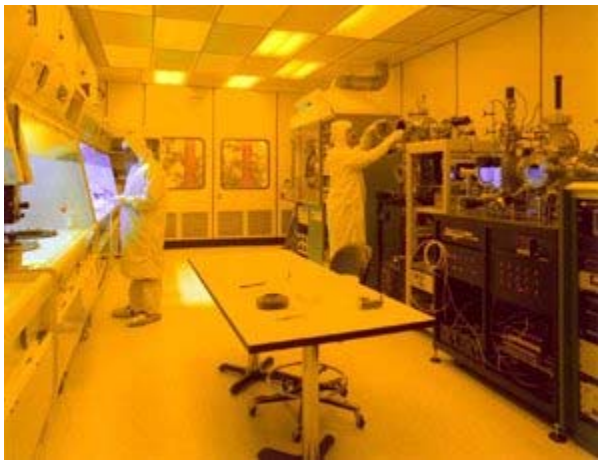
ISBN 978-93-81157-31-2

Published by:
The English Press
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter- 1

Semiconductor Device Fabrication



NASA's Glenn Research Center cleanroom.

Semiconductor device fabrication is the process used to create the integrated circuits (silicon chips) that are present in everyday electrical and electronic devices. It is a multiple-step sequence of photographic and chemical processing steps during which electronic circuits are gradually created on a wafer made of pure semiconducting material. Silicon is the most commonly used semiconductor material today, along with various compound semiconductors.

The entire manufacturing process, from start to packaged chips ready for shipment, takes six to eight weeks and is performed in highly specialized facilities referred to as fabs.

History

When feature widths were far greater than about 10 micrometres, purity was not the issue that it is today in device manufacturing. As devices became more integrated, cleanrooms became even cleaner. Today, the fabs are pressurized with filtered air to remove even the smallest particles, which could come to rest on the wafers and contribute to defects. The

workers in a semiconductor fabrication facility are required to wear cleanroom suits to protect the devices from human contamination.

In an effort to increase profits, semiconductor device manufacturing has spread from Texas and California in the 1960s to the rest of the world, such as Europe, Israel, and Asia. It is a global business today.

The leading semiconductor manufacturers typically have facilities all over the world. Intel, the world's largest manufacturer, has facilities in Europe and Asia as well as the U.S. Other top manufacturers include STMicroelectronics (Europe), Analog Devices (US), Integrated Device Technology (US), Atmel (US/Europe), Freescale Semiconductor (US), Samsung (Korea), Texas Instruments (US), GlobalFoundries (Germany, Singapore, future New York fab in construction), Toshiba (Japan), NEC Electronics (Japan), Infineon (Europe), Renesas (Japan), Taiwan Semiconductor Manufacturing Company (Taiwan), Fujitsu (Japan/US), NXP Semiconductors (Europe), Micron Technology (US), Hynix (Korea) and SMIC (China).

Wafers

A typical wafer is made out of extremely pure silicon that is grown into mono-crystalline cylindrical ingots (boules) up to 300 mm (slightly less than 12 inches) in diameter using the Czochralski process. These ingots are then sliced into wafers about 0.75 mm thick and polished to obtain a very regular and flat surface.

Once the wafers are prepared, many process steps are necessary to produce the desired semiconductor integrated circuit. In general, the steps can be grouped into two major parts:

- Front-end-of-line (FEOL) processing
- Back-end-of-line (BEOL) processing

Processing

In semiconductor device fabrication, the various processing steps fall into four general categories: deposition, removal, patterning, and modification of electrical properties.

- Deposition is any process that grows, coats, or otherwise transfers a material onto the wafer. Available technologies consist of physical vapor deposition (PVD), chemical vapor deposition (CVD), electrochemical deposition (ECD), molecular beam epitaxy (MBE) and more recently, atomic layer deposition (ALD) among others.
- Removal processes are any that remove material from the wafer either in bulk or selectively and consist primarily of etch processes, either wet etching or dry etching. Chemical-mechanical planarization (CMP) is also a removal process used between levels.

- Patterning covers the series of processes that shape or alter the existing shape of the deposited materials and is generally referred to as lithography. For example, in conventional lithography, the wafer is coated with a chemical called a *photoresist*. The photoresist is exposed by a *stepper*, a machine that focuses, aligns, and moves the mask, exposing select portions of the wafer to short wavelength light. The unexposed regions are washed away by a developer solution. After etching or other processing, the remaining photoresist is removed by plasma ashing.
- Modification of electrical properties has historically consisted of doping transistor sources and drains originally by diffusion furnaces and later by ion implantation. These doping processes are followed by furnace anneal or in advanced devices, by rapid thermal anneal (RTA) which serve to activate the implanted dopants. Modification of electrical properties now also extends to reduction of dielectric constant in low-k insulating materials via exposure to ultraviolet light in UV processing (UVP).

Many modern chips have up to eleven metal levels produced in over 300 sequenced processing steps.

Front-end-of-line (FEOL) processing

FEOL processing refers to the formation of the transistors directly in the silicon. The raw wafer is engineered by the growth of an ultrapure, virtually defect-free silicon layer through epitaxy. In the most advanced logic devices, *prior* to the silicon epitaxy step, tricks are performed to improve the performance of the transistors to be built. One method involves introducing a *straining step* wherein a silicon variant such as silicon-germanium (SiGe) is deposited. Once the epitaxial silicon is deposited, the crystal lattice becomes stretched somewhat, resulting in improved electronic mobility. Another method, called *silicon on insulator* technology involves the insertion of an insulating layer between the raw silicon wafer and the thin layer of subsequent silicon epitaxy. This method results in the creation of transistors with reduced parasitic effects.

Gate oxide and implants

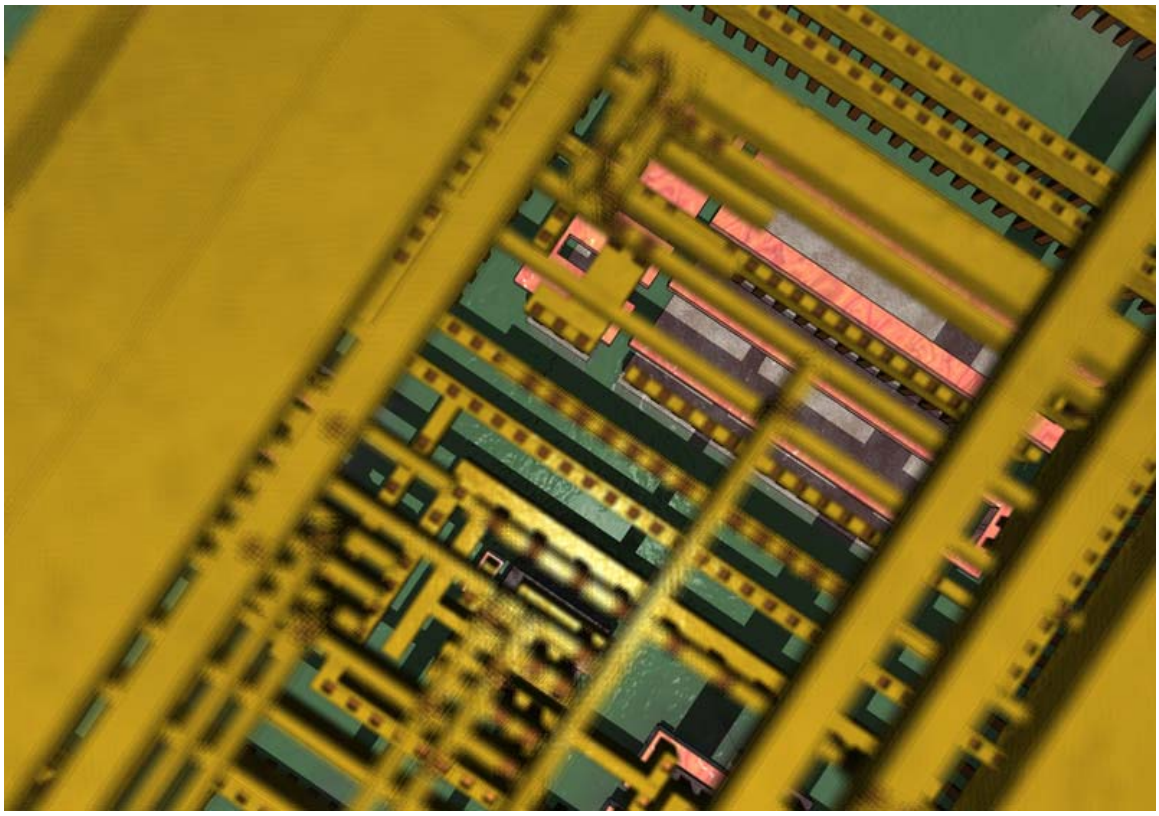
Front-end surface engineering is followed by: growth of the gate dielectric, traditionally silicon dioxide (SiO_2), patterning of the gate, patterning of the source and drain regions, and subsequent implantation or diffusion of dopants to obtain the desired complementary electrical properties. In dynamic random access memory (DRAM) devices, storage capacitors are also fabricated at this time, typically stacked above the access transistor (implementing them as trenches etched deep into the silicon surface was a technique developed by the now defunct DRAM manufacturer Qimonda).

Back-end-of-line (BEOL) processing

Metal layers

Once the various semiconductor devices have been created, they must be interconnected to form the desired electrical circuits. This occurs in a series of wafer processing steps collectively referred to as BEOL (not to be confused with *back end* of chip fabrication which refers to the packaging and testing stages). BEOL processing involves creating metal interconnecting wires that are isolated by dielectric layers. The insulating material was traditionally a form of SiO_2 or a silicate glass, but recently new low dielectric constant materials are being used. These dielectrics presently take the form of SiOC and have dielectric constants around 2.7 (compared to 3.9 for SiO_2), although materials with constants as low as 2.2 are being offered to chipmakers.

Interconnect



Synthetic detail of a standard cell through four layers of planarized copper interconnect, down to the polysilicon (pink), wells (greyish) and substrate (green).

Historically, the metal wires consisted of aluminium. In this approach to wiring often called *subtractive aluminium*, blanket films of aluminium are deposited first, patterned, and then etched, leaving isolated wires. Dielectric material is then deposited over the exposed wires. The various metal layers are interconnected by etching holes, called *vias*, in the insulating material and depositing tungsten in them with a CVD technique. This approach is still used in the fabrication of many memory chips such as dynamic random access memory (DRAM) as the number of interconnect levels is small, currently no more than four.

More recently, as the number of interconnect levels for logic has substantially increased due to the large number of transistors that are now interconnected in a modern microprocessor, the timing delay in the wiring has become significant prompting a change in wiring material from aluminium to copper and from the silicon dioxides to newer low-K material. This performance enhancement also comes at a *reduced cost* via damascene processing that eliminates processing steps. In damascene processing, in contrast to subtractive aluminium technology, the dielectric material is deposited first as a blanket film, and is patterned and etched leaving holes or trenches. In *single damascene* processing, copper is then deposited in the holes or trenches surrounded by a thin barrier film resulting in filled vias or wire *lines* respectively. In *dual damascene* technology, both the trench and via are fabricated before the deposition of copper resulting in formation of both the via and line simultaneously, further reducing the number of processing steps. The thin barrier film, called copper barrier seed (CBS), is necessary to prevent copper diffusion into the dielectric. The ideal barrier film is as thin as possible. As the presence of excessive barrier film competes with the available copper wire cross section, formation of the thinnest continuous barrier represents one of the greatest ongoing challenges in copper processing today.

As the number of interconnect levels increases, planarization of the previous layers is required to ensure a flat surface prior to subsequent lithography. Without it, the levels would become increasingly crooked and extend outside the depth of focus of available lithography, interfering with the ability to pattern. CMP (chemical mechanical planarization) is the primary processing method to achieve such planarization although dry *etch back* is still sometimes employed if the number of interconnect levels is no more than three.

Wafer test

The highly serialized nature of wafer processing has increased the demand for metrology in between the various processing steps. Wafer test metrology equipment is used to verify that the wafers haven't been damaged by previous processing steps up until testing. If the number of dies—the integrated circuits that will eventually become chips— etched on a wafer exceeds a failure threshold (ie. too many failed dies on one wafer), the wafer is scrapped rather than investing in further processing.

Device test

Once the front-end process has been completed, the semiconductor devices are subjected to a variety of electrical tests to determine if they function properly. The proportion of devices on the wafer found to perform properly is referred to as the yield.

The fab tests the chips on the wafer with an electronic tester that presses tiny probes against the chip. The machine marks each bad chip with a drop of dye. The fab charges for test time; the prices are on the order of cents per second. Chips are often designed

with “testability features” such as "built-in self-test" to speed testing, and reduce test costs.

Good designs try to test and statistically manage *corners*: extremes of silicon behavior caused by operating temperature combined with the extremes of fab processing steps. Most designs cope with more than 64 corners.

Die preparation

Once tested, a wafer is typically reduced in thickness

before the wafer is scored and then broken into individual die -- wafer dicing.

Only the good, unmarked chips go on to be packaged.

Packaging

Plastic or ceramic packaging involves mounting the die, connecting the die pads to the pins on the package, and sealing the die. Tiny wires are used to connect pads to the pins. In the old days, wires were attached by hand, but now purpose-built machines perform the task. Traditionally, the wires to the chips were gold, leading to a “lead frame” (pronounced “leed frame”) of copper, that had been plated with solder, a mixture of tin and lead. Lead is poisonous, so lead-free “lead frames” are now mandated by ROHS.

Chip-scale package (CSP) is another packaging technology. A plastic dual in-line package, like most packages, is many times larger than the actual die hidden inside, whereas CSP chips are nearly the size of the die. CSP can be constructed for each die *before* the wafer is diced.

The packaged chips are retested to ensure that they were not damaged during packaging and that the die-to-pin interconnect operation was performed correctly. A laser etches the chip’s name and numbers on the package.

List of steps

This is a list of processing techniques that are employed numerous times in a modern electronic device and do not necessarily imply a specific order.

- Wafer processing
    - Wet cleans
    - Photolithography
    - Ion implantation (in which dopants are embedded in the wafer creating regions of increased (or decreased) conductivity)
    - Dry etching
    - Wet etching

    - Plasma ashing
    - Thermal treatments
        - Rapid thermal anneal
        - Furnace anneals
        - Thermal oxidation
    - Chemical vapor deposition (CVD)
    - Physical vapor deposition (PVD)
    - Molecular beam epitaxy (MBE)
    - Electrochemical Deposition (ECD).
    - Chemical-mechanical planarization (CMP)
    - Wafer testing (where the electrical performance is verified)
    - Wafer backgrinding (to reduce the thickness of the wafer so the resulting chip can be put into a thin device like a smartcard or PCMCIA card.)
- Die preparation
    - Wafer mounting
    - Die cutting
- IC packaging
    - Die attachment
    - IC Bonding
        - Wire bonding
        - Thermosonic Bonding
        - Flip chip
        - Tab bonding
    - IC encapsulation
        - Baking
        - Plating
        - Lasermarking
        - Trim and form
- IC testing

Hazardous materials

Many toxic materials are used in the fabrication process. These include:

- poisonous elemental dopants such as arsenic, antimony and phosphorus
- poisonous compounds like arsine, phosphine and silane
- highly reactive liquids, such as hydrogen peroxide, fuming nitric acid, sulfuric acid and hydrofluoric acid

It is vital that workers not be directly exposed to these dangerous substances. The high degree of automation common in the IC fabrication industry helps to reduce the risks of exposure of this sort. Most fabrication facilities employ exhaust management systems, such as wet scrubbers, combustors, heated absorber cartridges etc, to control the risk to workers and also the environment if these toxic materials are released into the atmosphere.

Chapter- 2

Photolithography

Photolithography (or "optical lithography") is a process used in microfabrication to selectively remove parts of a thin film or the bulk of a substrate. It uses light to transfer a geometric pattern from a photo mask to a light-sensitive chemical "photoresist", or simply "resist," on the substrate. A series of chemical treatments then engraves the exposure pattern into the material underneath the photo resist. In complex integrated circuits, for example a modern CMOS, a wafer will go through the photolithographic cycle up to 50 times.

Photolithography shares some fundamental principles with photography in that the pattern in the etching resist is created by exposing it to light, either using a projected image or an optical mask. This procedure is comparable to a high precision version of the method used to make printed circuit boards. Subsequent stages in the process have more in common with etching than to lithographic printing. It is used because it affords exact control over the shape and size of the objects it creates, and because it can create patterns over an entire surface simultaneously. Its main disadvantages are that it requires a flat substrate to start with, it is not very effective at creating shapes that are not flat, and it can require extremely clean operating conditions.

Basic procedure

The wafer track portion of an aligner that uses 365 nm ultraviolet light.

A single iteration of photolithography combines several steps in sequence. Modern cleanrooms use automated, robotic wafer track systems to coordinate the process. The procedure described here omits some advanced treatments, such as thinning agents or edge-bead removal.

Cleaning

If organic or inorganic contaminations are present on the wafer surface, they are usually removed by wet chemical treatment, e.g. the RCA clean procedure based on solutions containing hydrogen peroxide.

Preparation

The wafer is initially heated to a temperature sufficient to drive off any moisture that may be present on the wafer surface. Wafers that have been in storage must be chemically cleaned to remove contamination. A liquid or gaseous "adhesion promoter", such as Bis(trimethylsilyl)amine ("hexamethyldisilazane", HMDS), is applied to promote adhesion of the photoresist to the wafer. The phrase "adhesion promoter" is a misnomer, as the surface layer of silicon dioxide on the wafer reacts with the agent to form Methylated Silicon-hydroxide, a highly water repellent layer not unlike the layer of wax

on a car's paint. This water repellent layer prevents the aqueous developer from penetrating between the photoresist layer and the wafer's surface, thus preventing so-called lifting of small photoresist structures in the (developing) pattern.

Photoresist application

The wafer is covered with photoresist by spin coating. A viscous, liquid solution of photoresist is dispensed onto the wafer, and the wafer is spun rapidly to produce a uniformly thick layer. The spin coating typically runs at 1200 to 4800 rpm for 30 to 60 seconds, and produces a layer between 0.5 and 2.5 micrometres thick. The spin coating process results in a uniform thin layer, usually with uniformity of within 5 to 10 nanometres. This uniformity can be explained by detailed fluid-mechanical modelling, which shows that the resist moves much faster at the top of the layer than at the bottom, where viscous forces bind the resist to the wafer surface. Thus, the top layer of resist is quickly ejected from the wafer's edge while the bottom layer still creeps slowly radially along the wafer. In this way, any 'bump' or 'ridge' of resist is removed, leaving a very flat layer. Final thickness is also determined by the evaporation of liquid solvents from the resist. For very small, dense features (<125 or so nm), thinner resist thicknesses (<0.5 micrometres) are needed to overcome collapse effects at high aspect ratios; typical aspect ratios are <4:1.

The photo resist-coated wafer is then prebaked to drive off excess photoresist solvent, typically at 90 to 100 °C for 30 to 60 seconds on a hotplate.

Exposure and developing

After prebaking, the photoresist is exposed to a pattern of intense light. Optical lithography typically uses ultraviolet light (see below). Positive photoresist, the most common type, becomes soluble in the basic developer when exposed; exposed negative photoresist becomes insoluble in the (organic) developer. This chemical change allows some of the photoresist to be removed by a special solution, called "developer" by analogy with photographic developer.

A PEB (post-exposure bake) is performed before developing, typically to help reduce standing wave phenomena caused by the destructive and constructive interference patterns of the incident light. In DUV (deep ultraviolet, or 248 nm exposure wavelength) lithography, CAR (chemically amplified resist) chemistry is used. This process is much more sensitive to PEB time, temperature, and delay, as most of the "exposure" reaction (creating acid, making the polymer soluble in the basic developer) actually occurs in the PEB.

The develop chemistry is delivered on a spinner, much like photoresist. Developers originally often contained sodium hydroxide (NaOH). However, sodium is considered an extremely undesirable contaminant in MOSFET fabrication because it degrades the insulating properties of gate oxides (specifically, sodium ions can migrate in and out of the gate, changing the threshold voltage of the transistor and making it harder or easier to

turn the transistor on over time). Metal-ion-free developers such as tetramethylammonium hydroxide (TMAH) are now used.

The resulting wafer is then "hard-baked" if a non-chemically amplified resist was used, typically at 120 to 180 °C for 20 to 30 minutes. The hard bake solidifies the remaining photoresist, to make a more durable protecting layer in future ion implantation, wet chemical etching, or plasma etching.

Etching

In etching, a liquid ("wet") or plasma ("dry") chemical agent removes the uppermost layer of the substrate in the areas that are not protected by photoresist. In semiconductor fabrication, dry etching techniques are generally used, as they can be made anisotropic, in order to avoid significant undercutting of the photoresist pattern. This is essential when the width of the features to be defined is similar to or less than the thickness of the material being etched (i.e. when the aspect ratio approaches unity). Wet etch processes are generally isotropic in nature, which is often indispensable for microelectromechanical systems, where suspended structures must be "released" from the underlying layer.

The development of low-defectivity anisotropic dry-etch process has enabled the ever-smaller features defined photolithographically in the resist to be transferred to the substrate material.

Photoresist removal

After a photoresist is no longer needed, it must be removed from the substrate. This usually requires a liquid "resist stripper", which chemically alters the resist so that it no longer adheres to the substrate. Alternatively, photoresist may be removed by a plasma containing oxygen, which oxidizes it. This process is called ashing, and resembles dry etching.

Exposure ("printing") systems

Exposure systems typically produce an image on the wafer using a photomask. The light shines through the photomask, which blocks it in some areas and lets it pass in others. (Maskless lithography projects a precise beam directly onto the wafer without using a mask, but it is not widely used in commercial processes.) Exposure systems may be classified by the optics that transfer the image from the mask to the wafer.

Contact and proximity

A contact printer, the simplest exposure system, puts a photomask in direct contact with the wafer and exposes it to a uniform light. A proximity printer puts a small gap between the photomask and wafer. In both cases, the mask covers the entire wafer, and simultaneously patterns every die.

Contact printing is liable to damage both the mask and the wafer, and this was the primary reason it was abandoned for high volume production. Both contact and proximity lithography require the light intensity to be uniform across an entire wafer, and the mask to align precisely to features already on the wafer. As modern processes use increasingly large wafers, these conditions become increasingly difficult.

Research and prototyping processes often use contact lithography, because it uses inexpensive hardware and can achieve high optical resolution. The resolution is approximately the square root of the product of the wavelength and the gap distance. Hence, contact printing offers the best resolution, because its gap distance is approximately zero (neglecting the thickness of the photoresist itself). In addition, nanoimprint lithography may revive interest in this familiar technique, especially since the cost of ownership is expected to be low.

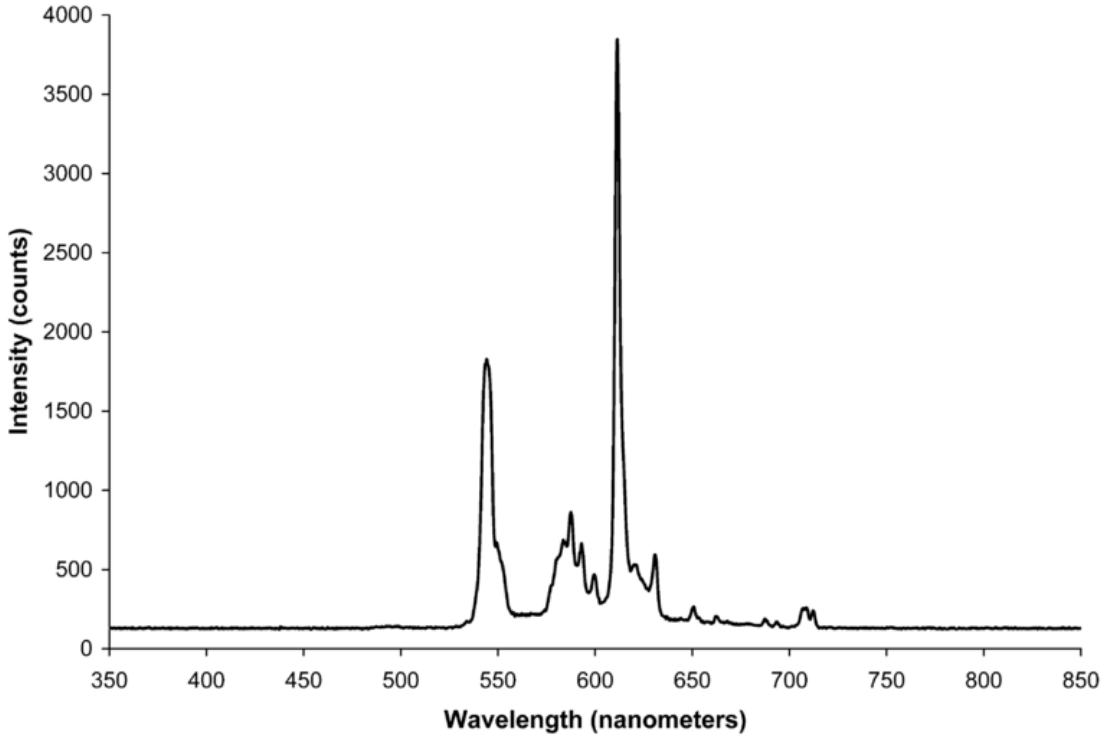
Projection

Very-large-scale integration lithography uses projection systems. Unlike contact or proximity masks, which cover an entire wafer, projection masks (known as "reticles") show only one die or an array of die (known as a "field"). Projection exposure systems (steppers) project the mask onto the wafer many times to create the complete pattern.

Photomasks

The image for the mask originates from a computerized data file. This data file is converted to a series of polygons and written onto a square fused quartz substrate covered with a layer of chrome using a photolithographic process. A laser beam (laser writer) or a beam of electrons (e-beam writer) is used to expose the pattern defined in the data file and travels over the surface of the substrate in either a vector or raster scan manner. Where the photoresist on the mask is exposed, the chrome can be etched away, leaving a clear path for the light in the stepper/scanner systems to travel through.

Resolution in projection systems



The filtered fluorescent lighting in photolithography cleanrooms contains no ultraviolet or blue light in order to avoid exposing photoresists. The spectrum of light emitted by such fixtures gives virtually all such spaces a bright yellow color.

The ability to project a clear image of a small feature onto the wafer is limited by the wavelength of the light that is used, and the ability of the reduction lens system to capture enough diffraction orders from the illuminated mask. Current state-of-the-art photolithography tools use deep ultraviolet (DUV) light with wavelengths of 248 and 193 nm, which allow minimum feature sizes down to 50 nm.

The minimum feature size that a projection system can print is given approximately by:

$$CD = k_1 \cdot \frac{\lambda}{NA}$$

where

CD is the **minimum feature size** (also called the **critical dimension**, *target design rule*). It is also common to write 2 *times* the *half-pitch*.

k_1 (commonly called *k1 factor*) is a coefficient that encapsulates process-related factors, and typically equals 0.4 for production. The minimum feature size can be reduced by decreasing this coefficient through Computational lithography.

λ is the wavelength of light used

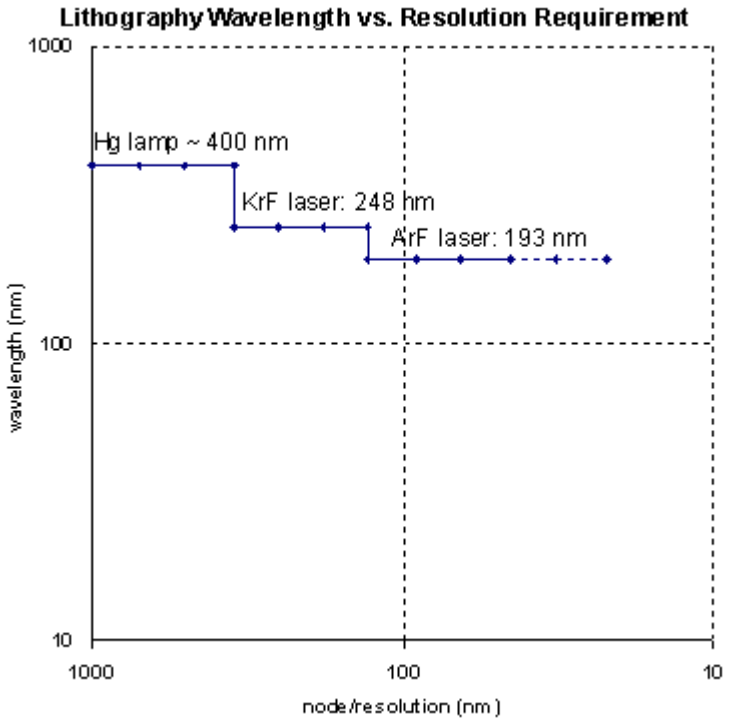
NA is the numerical aperture of the lens as seen from the wafer

According to this equation, minimum feature sizes can be decreased by decreasing the wavelength, and increasing the numerical aperture (to achieve a tighter focused beam and a smaller spot size). However, this design method runs into a competing constraint. In modern systems, the depth of focus is also a concern:

$$D_F = k_2 \cdot \frac{\lambda}{NA^2}$$

Here, k_2 is another process-related coefficient. The depth of focus restricts the thickness of the photoresist and the depth of the topography on the wafer. Chemical mechanical polishing is often used to flatten topography before high-resolution lithographic steps.

Light sources



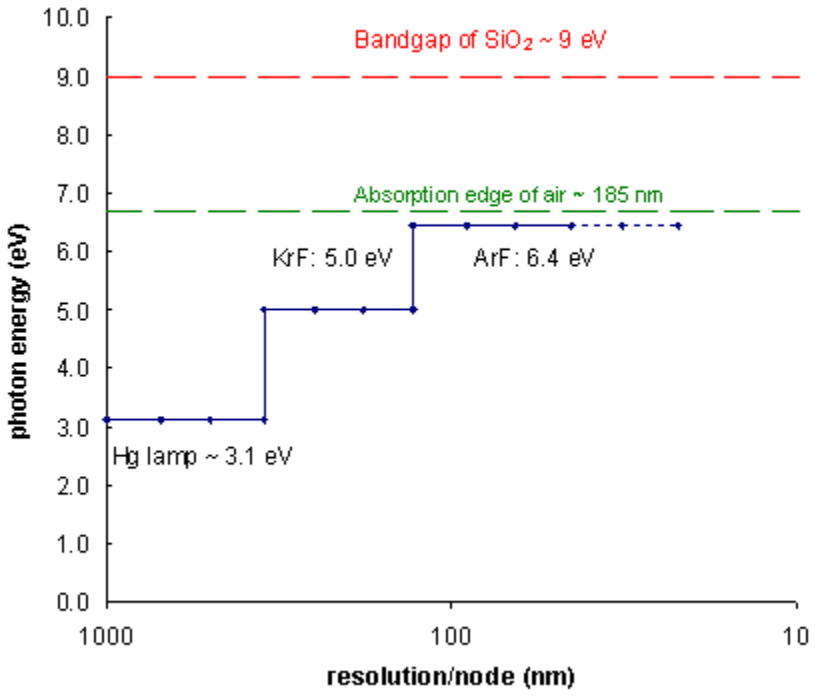
The evolution of lithography wavelength corresponding to different light sources. It is worth noting that the same light source may be used for several technology generations.

Historically, photolithography has used ultraviolet light from gas-discharge lamps using mercury, sometimes in combination with noble gases such as xenon. These lamps

produce light across a broad spectrum with several strong peaks in the ultraviolet range. This spectrum is filtered to select a single spectral line, usually the "g-line" (436 nm) or "i-line" (365 nm).

More recently, lithography has moved to "deep ultraviolet", produced by excimer lasers. (In lithography, wavelengths below 300 nm are called "deep UV".) Krypton fluoride produces a 248-nm spectral line, and argon fluoride a 193-nm line. Excimer laser light sources for photolithography applications were first introduced by Cymer Inc. in the late 1980s for 248-nm wavelength lithography (KrF), and in 2001 for 193-nm wavelength lithography (ArF). Generally, changing wavelength is not a trivial matter, as the method of generating the new wavelength is completely different, and the absorption characteristics of materials change. For example, air begins to absorb significantly around the 193 nm wavelength; moving to sub-193 nm wavelengths would require installing vacuum pump and purge equipment on the lithography tools (a significant challenge). Furthermore, insulating materials such as silicon dioxide (SiO_2), when exposed to photons with energy greater than the band gap, release free electrons and holes which subsequently cause adverse charging.

Optical lithography has been extended to feature sizes below 50 nm using 193 nm and liquid immersion techniques. Also termed immersion lithography, this enables the use of optics with numerical apertures exceeding 1.0. The liquid used is typically ultra-pure, deionised water, which provides for a refractive index above that of the usual air gap between the lens and the wafer surface. The water is continually circulated to eliminate thermally-induced distortions. Water will only allow *NA*'s of up to ~1.4, but materials with higher refractive indices will allow the effective *NA* to be increased further.



Changing the lithography wavelength is significantly limited by absorption. Air absorbs below ~ 185 nm.

Experimental tools using 157 nm wavelength DUV in a manner similar to current exposure systems have been built. These were once targeted to succeed 193 nm at the 65 nm feature size node but have now all but been eliminated by the introduction of immersion lithography. This was due to persistent technical problems with the 157 nm technology and economic considerations that provided strong incentives for the continued use of 193 nm technology. High-index immersion lithography is the newest extension of 193 nm lithography to be considered. In 2006, features less than 30 nm were demonstrated by IBM using this technique.

Experimental methods

Photolithography has been defeating predictions of its demise for many years. For instance, it was predicted that features smaller than 1 micrometre could not be printed optically. Modern techniques already print features with dimensions a fraction of the wavelength of light used - an amazing optical feat. Current research is exploring new tricks in the ultraviolet regime, as well as alternatives to conventional UV, such as electron beam lithography, X-ray lithography, extreme ultraviolet lithography, ion projection lithography, and immersion lithography.

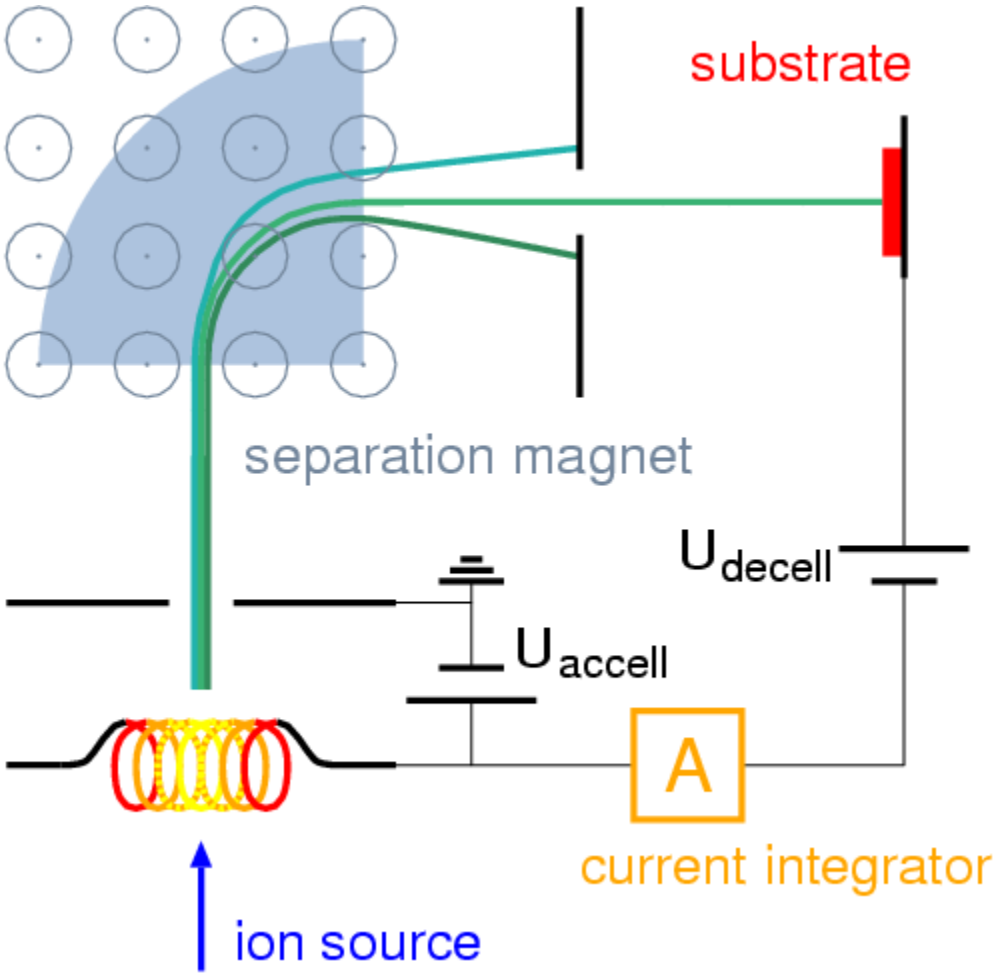
Chapter- 3

Ion Implantation

An ion implantation system at LAAS technological facility in Toulouse, France.

Ion implantation is a materials engineering process by which ions of a material can be implanted into another solid, thereby changing the physical properties of the solid. Ion implantation is used in semiconductor device fabrication and in metal finishing, as well as various applications in materials science research. The ions introduce both a chemical change in the target, in that they can be a different element than the target or induce a nuclear transmutation, and a structural change, in that the crystal structure of the target can be damaged or even destroyed by the energetic collision cascades.

General principle



Ion implantation setup with mass separator

Ion implantation equipment typically consists of an ion source, where ions of the desired element are produced, an accelerator, where the ions are electrostatically accelerated to a high energy, and a target chamber, where the ions impinge on a target, which is the material to be implanted. Thus ion implantation is a special case of particle radiation. Each ion is typically a single atom or molecule, and thus the actual amount of material implanted in the target is the integral over time of the ion current. This amount is called the dose. The currents supplied by implanters are typically small (microamperes), and

thus the dose which can be implanted in a reasonable amount of time is small. Therefore, ion implantation finds application in cases where the amount of chemical change required is small.

Typical ion energies are in the range of 10 to 500 keV (1,600 to 80,000 aJ). Energies in the range 1 to 10 keV (160 to 1,600 aJ) can be used, but result in a penetration of only a few nanometers or less. Energies lower than this result in very little damage to the target, and fall under the designation ion beam deposition. Higher energies can also be used: accelerators capable of 5 MeV (800,000 aJ) are common. However, there is often great structural damage to the target, and because the depth distribution is broad, the net composition change at any point in the target will be small.

The energy of the ions, as well as the ion species and the composition of the target determine the depth of penetration of the ions in the solid: A monoenergetic ion beam will generally have a broad depth distribution. The average penetration depth is called the range of the ions. Under typical circumstances ion ranges will be between 10 nanometers and 1 micrometer. Thus, ion implantation is especially useful in cases where the chemical or structural change is desired to be near the surface of the target. Ions gradually lose their energy as they travel through the solid, both from occasional collisions with target atoms (which cause abrupt energy transfers) and from a mild drag from overlap of electron orbitals, which is a continuous process. The loss of ion energy in the target is called stopping and can be simulated with the binary collision approximation method.

Application in semiconductor device fabrication

Doping

The introduction of dopants in a semiconductor is the most common application of ion implantation. Dopant ions such as boron, phosphorus or arsenic are generally created from a gas source, so that the purity of the source can be very high. These gases tend to be very hazardous. When implanted in a semiconductor, each dopant atom can create a charge carrier in the semiconductor after annealing. A hole can be created for a p-type dopant, and an electron for an n-type dopant. This modifies the conductivity of the semiconductor in its vicinity. The technique is used, for example, for adjusting the threshold of a MOSFET.

Ion implantation was developed as a method of producing the p-n junction of photovoltaic devices in the late 1970s and early 1980s, along with the use of pulsed-electron beam for rapid annealing, although it has not to date been used for commercial production.

Silicon on insulator

One prominent method for preparing silicon on insulator (SOI) substrates from conventional silicon substrates is the *SIMOX* (**S**eparation by **IM**plantation of **OX**ygen)

process, wherein a buried high dose oxygen implant is converted to silicon oxide by a high temperature annealing process.

Mesotaxy

Mesotaxy is the term for the growth of a crystallographically matching phase underneath the surface of the host crystal (compare to epitaxy, which is the growth of the matching phase on the surface of a substrate). In this process, ions are implanted at a high enough energy and dose into a material to create a layer of a second phase, and the temperature is controlled so that the crystal structure of the target is not destroyed. The crystal orientation of the layer can be engineered to match that of the target, even though the exact crystal structure and lattice constant may be very different. For example, after the implantation of nickel ions into a silicon wafer, a layer of nickel silicide can be grown in which the crystal orientation of the silicide matches that of the silicon.

Application in metal finishing

Tool steel toughening

Nitrogen or other ions can be implanted into a tool steel target (drill bits, for example). The structural change caused by the implantation produces a surface compression in the steel, which prevents crack propagation and thus makes the material more resistant to fracture. The chemical change can also make the tool more resistant to corrosion.

Surface finishing

In some applications, for example prosthetic devices such as artificial joints, it is desired to have surfaces very resistant to both chemical corrosion and wear due to friction. Ion implantation is used in such cases to engineer the surfaces of such devices for more reliable performance. As in the case of tool steels, the surface modification caused by ion implantation includes both a surface compression which prevents crack propagation and an alloying of the surface to make it more chemically resistant to corrosion.

Other applications

Ion beam mixing

Ion implantation can be used to achieve ion beam mixing, i.e. mixing up atoms of different elements at an interface. This may be useful for achieving graded interfaces or strengthening adhesion between layers of immiscible materials.

Problems with ion implantation

Crystallographic damage

Each individual ion produces many point defects in the target crystal on impact such as vacancies and interstitials. Vacancies are crystal lattice points unoccupied by an atom: in this case the ion collides with a target atom, resulting in transfer of a significant amount of energy to the target atom such that it leaves its crystal site. This target atom then itself becomes a projectile in the solid, and can cause successive collision events. Interstitials result when such atoms (or the original ion itself) come to rest in the solid, but find no vacant space in the lattice to reside. These point defects can migrate and cluster with each other, resulting in dislocation loops and other defects.

Damage recovery

Because ion implantation causes damage to the crystal structure of the target which is often unwanted, ion implantation processing is often followed by a thermal annealing. This can be referred to as damage recovery.

Amorphization

The amount of crystallographic damage can be enough to completely amorphize the surface of the target: i.e. it can become an amorphous solid (such a solid produced from a melt is called a glass). In some cases, complete amorphization of a target is preferable to a highly defective crystal: An amorphized film can be regrown at a lower temperature than required to anneal a highly damaged crystal.

Sputtering

Some of the collision events result in atoms being ejected (sputtered) from the surface, and thus ion implantation will slowly etch away a surface. The effect is only appreciable for very large doses.

Ion channelling

A diamond cubic crystal viewed from the <110> direction, showing hexagonal ion channels.

If there is a crystallographic structure to the target, and especially in semiconductor substrates where the crystal structure is more open, particular crystallographic directions offer much lower stopping than other directions. The result is that the range of an ion can be much longer if the ion travels exactly along a particular direction, for example the <110> direction in silicon and other diamond cubic materials. This effect is called *ion channelling*, and, like all the channelling effects, is highly nonlinear, with small

variations from perfect orientation resulting in extreme differences in implantation depth. For this reason, most implantation is carried out a few degrees off-axis, where tiny alignment errors will have more predictable effects.

Ion channelling can be used directly in Rutherford backscattering and related techniques as an analytical method to determine the amount and depth profile of damage in crystalline thin film materials.

Hazardous Materials Note

In the ion implantation semiconductor fabrication process of wafers, it is important for the workers to minimize their exposure to the toxic materials used in the ion implanter process. Such hazardous elements, solid source and gasses are used, such as Arsine and Phosphine. For this reason, the semiconductor fabrication facilities are highly automated, and may feature negative pressure gas bottles safe delivery system (SDS). Other elements may include Antimony, Arsenic, Phosphorus, and Boron. Residue of these elements show up when the machine is opened to atmosphere, and can also be accumulated and found concentrated in the vacuum pumps hardware. It is important not to expose yourself to these carcinogenic, corrosive, flammable, and toxic elements. Many overlapping safety protocols must be used when handling these deadly compounds. Use safety, and read MSDS's.

High Voltage Safety

High voltage power supplies in ion implantation equipment can pose a risk of electrocution. In addition, high-energy atomic collisions can, in some cases, generate radionuclides. Operators and Maintenance personnel should learn and follow the safety advice of the manufacturer and/or the institution responsible for the equipment. Prior to entry to high voltage area, terminal components must be grounded using a grounding stick. Next, power supplies should be locked in the off state and tagged to prevent unauthorized energizing.

Chapter- 4

Dry Etching, Etching & Plasma Ashing

Dry etching

Dry etching refers to the removal of material, typically a masked pattern of semiconductor material, by exposing the material to a bombardment of ions (usually a plasma of reactive gases such as fluorocarbons, oxygen, chlorine, boron trichloride; sometimes with addition of nitrogen, argon, helium and other gases) that dislodge portions of the material from the exposed surface. Unlike with many of the wet chemical etchants used in wet etching, the dry etching process typically etches directionally or anisotropically.

Explanation

Dry etching is used in conjunction with photolithographic techniques to attack certain areas of a semiconductor surface in order to form recesses in material, such as contact holes (which are contacts to the underlying semiconductor substrate) or via holes (which are holes that are formed to provide an interconnect path between conductive layers in the layered semiconductor device) or to otherwise remove portions of semiconductor layers where predominantly vertical sides are desired. In conjunction with semiconductor manufacturing, micromachining and display production the removal of organic residues by oxygen plasmas is sometimes correctly described as a dry etch process. However, also the term plasma ashing may be used.

Dry etching is particularly useful for materials and semiconductors which are chemically resistant and could not be wet etched, such as silicon carbide or gallium nitride

Etching (microfabrication)

Etching tanks used to perform Piranha, Hydrofluoric acid or RCA clean on 4-inch wafer batches at LAAS technological facility in Toulouse, France.

Etching is used in microfabrication to chemically remove layers from the surface of a wafer during manufacturing. Etching is a critically important process module, and every wafer undergoes many etching steps before it is complete.

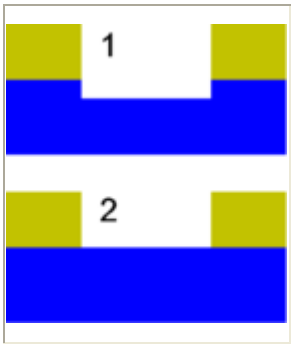
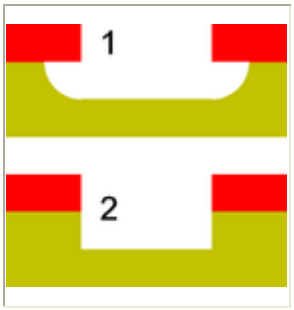
For many etch steps, part of the wafer is protected from the etchant by a "masking" material which resists etching. In some cases, the masking material is a photoresist which has been patterned using photolithography. Other situations require a more durable mask, such as silicon nitride.

Figures of merit

If the etch is intended to make a cavity in a material, the depth of the cavity may be controlled approximately using the etching time and the known etch rate. More often, though, etching must entirely remove the top layer of a multilayer structure, without damaging the underlying or masking layers. The etching system's ability to do this depends on the ratio of etch rates in the two materials (*selectivity*).

Some etches undercut the masking layer and form cavities with sloping sidewalls. The distance of undercutting is called *bias*. Etchants with large bias are called *isotropic*,

because they erode the substrate equally in all directions. Modern processes greatly prefer anisotropic etches, because they produce sharp, well-controlled features.

Selectivity		Yellow: layer to be removed; blue: layer to remain 1. A poorly selective etch removes the top layer, but also attacks the underlying material. 2. A highly selective etch leaves the underlying material unharmed.
Isotropy		Red: masking layer; yellow: layer to be removed 1. A perfectly isotropic etch produces round sidewalls. 2. A perfectly anisotropic etch produces vertical sidewalls.

Etching media and technology

The two fundamental types of etchants are liquid-phase ("wet") and plasma-phase ("dry"). Each of these exists in several varieties.

Wet etching

The first etching processes used liquid-phase ("wet") etchants. The wafer can be immersed in a bath of etchant, which must be agitated to achieve good process control. For instance, buffered hydrofluoric acid (BHF) is used commonly to etch silicon dioxide over a silicon substrate.

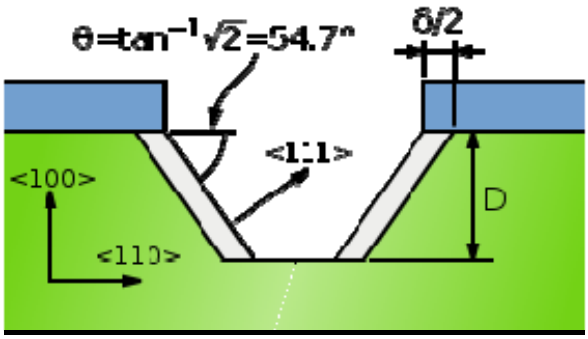
Different specialised etchants can be used to characterise the surface etched.

Wet etchants are usually isotropic, which leads to large bias when etching thick films. They also require the disposal of large amounts of toxic waste. For these reasons, they are seldom used in state-of-the-art processes. However, the photographic developer used for photoresist resembles wet etching.

As an alternative to immersion, single wafer machines use the Bernoulli principle to employ a gas (usually, pure nitrogen) to cushion and protect one side of the wafer while etchant is applied to the other side. It can be done to either the front side or back side. The etch chemistry is dispensed on the top side when in the machine and the bottom side is not affected. This etch method is particularly effective just before "backend"

processing (BEOL), where wafers are normally very much thinner after wafer backgrinding, and very sensitive to thermal or mechanical stress. Etching a thin layer of even a few micrometres will remove microcracks produced during backgrinding resulting in a the wafer having dramatically increased strength and flexibility without breaking.

Anisotropic wet etching



An anisotropic wet etch on a silicon wafer creates a cavity with a trapezoidal cross-section. The bottom of the cavity is a <100> plane, and the sides are <111> planes. The blue material is an etch mask, and the green material is silicon.

Some wet etchants etch crystalline materials at very different rates depending upon which crystal face is exposed. In single-crystal materials (e.g. silicon wafers), this effect can allow very high anisotropy, as shown in the figure.

Several anisotropic wet etchants are available for silicon. For instance, potassium hydroxide (KOH) can achieve selectivity of 400 between <100> and <111> planes. Another option is EDP (an aqueous solution of ethylene diamine and pyrocatechol), which also displays high selectivity for p-type doping. Neither of these etchants may be used on wafers that contain CMOS integrated circuits. Both of them etch aluminum, commonly used as a metallization (wiring) material. KOH introduces mobile potassium ions into silicon dioxide, and EDP is highly corrosive and carcinogenic. Tetramethylammonium hydroxide (TMAH) presents a safer alternative, although it has even worse selectivity between <100> and <111> planes in silicon than does EDP.

Etching a rectangular hole in a (100)-Si wafer will result in a pyramid shaped etch pit. The wall will be flat and angled (as opposed to curved in isotropic etching), and have an angle to the surface of the wafer of:

$$\tan^{-1}\sqrt{2} = 54.7^{\circ}$$

If the etching is stopped before the pyramid is formed, a frustum will be formed. The undercut, δ, under the resist mask is given by:

$$\delta - \frac{\sqrt{6}D}{S} - \frac{\sqrt{6}R_{100}T}{R_{100}/R_{111}} - \sqrt{6}TR_{111},$$

where R_{xxx} is the etch rate in the <xxx> direction, T is the etch time, D is the etch depth and S is the anisotropy of the material and etchant.

Different etchants have different anisotropies. Below is a table of common anisotropic etchants for silicon:

Etchant	Operating temp (°C)	R_{100} (µm/min)	$S=R_{100}/R_{111}$	Mask materials
Ethylenediamine pyrocatechol (EDP)	110	0.47	17	SiO_2, Si_3N_4, Au, Cr, Ag, Cu
Potassium hydroxide/Isopropyl alcohol (KOH/IPA)	50	1.0	400	Si_3N_4, SiO_2 (etches at 140nm/min)
Tetramethylammonium hydroxide (TMAH)	80	0.6	37	Si_3N_4, SiO_2

Plasma etching

Modern VLSI processes avoid wet etching, and use *plasma etching* instead. Plasma etchers can operate in several modes by adjusting the parameters of the plasma. Ordinary plasma etching operates between 0.1 and 5 Torr. (This unit of pressure, commonly used in vacuum engineering, equals approximately 133.3 pascals.) The plasma produces energetic free radicals, neutrally charged, that react at the surface of the wafer. Since neutral particles attack the wafer from all angles, this process is isotropic.

The source gas for the plasma usually contains small molecules rich in chlorine or fluorine. For instance, carbon tetrachloride (CCl_4) etches silicon and aluminium, and trifluoromethane etches silicon dioxide and silicon nitride. A plasma containing oxygen is used to oxidize ("ash") photoresist and facilitate its removal.

Ion milling, or *sputter etching*, uses lower pressures, often as low as 10^{-4} Torr (10 mPa). It bombards the wafer with energetic ions of noble gases, often Ar^+, which knock atoms from the substrate by transferring momentum. Because the etching is performed by ions, which approach the wafer approximately from one direction, this process is highly anisotropic. On the other hand, it tends to display poor selectivity. *Reactive-ion etching* (RIE) operates under conditions intermediate between sputter and plasma etching (between 10^{-3} and 10^{-1} Torr). *Deep reactive-ion etching* (DRIE) modifies the RIE technique to produce deep, narrow features.

Common etch processes used in microfabrication

Etchants for common microfabrication materials

Material to be etched	Wet etchants	Plasma etchants
Aluminium (Al)	80% phosphoric acid (H_3PO_4) + 5% acetic acid + 5% nitric acid (HNO_3) + 10% water (H_2O) at 35–45 °C	Cl_2, CCl_4, $SiCl_4$, BCl_3
Indium tin oxide [ITO] (In_2O_3:SnO_2)	Hydrochloric acid (HCl) + nitric acid (HNO_3) + water (H_2O) (1:0.1:1) at 40 °C	
Chromium (Cr)	• "Chrome etch": ceric ammonium nitrate ($(NH_4)_2Ce(NO_3)_6$) + nitric acid (HNO_3) • Hydrochloric acid (HCl)	
Gold (Au)	Aqua regia	
Molybdenum (Mo)		CF_4
Organic residues and photoresist	Piranha etch: sulfuric acid (H_2SO_4) + hydrogen peroxide (H_2O_2)	O_2 (ashing)
Platinum (Pt)	Aqua regia	
Silicon (Si)	Nitric acid (HNO_3) + hydrofluoric acid (HF)	• CF_4, SF_6, NF_3 • Cl_2, CCl_2F_2
Silicon dioxide (SiO_2)	• Hydrofluoric acid (HF) • Buffered oxide etch [BOE]: ammonium fluoride (NH_4F)	CF_4, SF_6, NF_3

	and hydrofluoric acid (HF)	
Silicon nitride (Si_3N_4)	• 85% Phosphoric acid (H_3PO_4) at 180 °C (Requires SiO_2 etch mask)	CF_4, SF_6, NF_3
Tantalum (Ta)		CF_4
Titanium (Ti)	Hydrofluoric acid (HF)	BCl_3
Titanium nitride (TiN)	• Nitric acid (HNO_3) + hydrofluoric acid (HF) • SC1	
Tungsten (W)	• Nitric acid (HNO_3) + hydrofluoric acid (HF) • Hydrogen Peroxide (H_2O_2)	• CF_4 • SF_6

Plasma ashing

In semiconductor manufacturing **plasma ashing** is the process of removing the photoresist from an etched wafer. Using a plasma source, a monatomic reactive species is generated. Oxygen or fluorine are the most common reactive species. The reactive species combines with the photoresist to form ash which is removed with a vacuum pump.

Typically, monatomic (single atom) oxygen plasma is created by exposing oxygen gas (O_2) to non-ionizing radiation. This process is done under vacuum in order to create a plasma. As the plasma is formed, many free radicals are created which could damage the wafer. Newer, smaller circuitry is increasingly susceptible to these particles. Originally, plasma was generated in the process chamber, but as the need to get rid of free radicals has increased, many machines now use a downstream plasma configuration, where plasma is formed remotely and the desired particles are channeled to the wafer. This allows electrically charged particles time to recombine before they reach the wafer surface, and prevents damage to the wafer surface.

Two forms of **plasma ashing** are typically performed on wafers. High temp ashing, or stripping, is performed to remove as much photo resist as possible, while the "descum" process is used to remove residual photo resist in trenches. The main difference between the two processes is the temperature the wafer is exposed to while in an ashing chamber.

Monatomic oxygen is electrically neutral and although it does recombine during the channeling, it does so at a slower rate than the positively or negatively charged free

radicals, which attract one another. Effectively, this means that when all of the free radicals have recombined, there is still a portion of the active species available for process. Because a large portion of the active species is lost to recombination, process times may take longer. To some extent, these longer process times can be mitigated by increasing the temperature of the reaction area.

Chapter- 5

Thermal Treatments

Rapid thermal processing

Rapid Thermal Processing (or **RTP**) refers to a semiconductor manufacturing process which heats silicon wafers to high temperatures (up to 1200 C or greater) on a timescale of several seconds or less. During cooling, however, wafer temperatures must be brought down slowly so they do not break due to thermal shock. Such rapid heating rates are often attained by high intensity lamps or lasers. These processes are used for a wide variety of applications in semiconductor manufacturing including dopant activation, thermal oxidation, metal reflow and chemical vapor deposition.

Temperature Control

One of the key challenges in rapid thermal processing is accurate measurement and control of the wafer temperature. Monitoring the ambient with a thermocouple has only recently become feasible, in that the high temperature ramp rates prevent the wafer from coming to thermal equilibrium with the process chamber. One temperature control strategy involves *in situ* pyrometry to effect real time control.

Rapid thermal anneal

Rapid thermal anneal (RTA) is a subset of Rapid Thermal Processing. It is a process used in semiconductor device fabrication which consists of heating a single wafer at a time in order to affect its electrical properties. Unique heat treatments are designed for different effects. Wafers can be heated in order to activate dopants, change film-to-film or film-to-wafer substrate interfaces, densify deposited films, change states of grown films, repair damage from ion implantation, move dopants or drive dopants from one film into another or from a film into the wafer substrate.

Rapid thermal anneals are performed by equipment that heats a single wafer at a time using either lamp based heating, a hot chuck, or a hot plate that a wafer is brought near.

Unlike furnace anneals they are short in duration, processing each wafer in several minutes.

To achieve short time annealing time trade off is made in temperature and process uniformity, temperature measurement and control and wafer stress as well as throughput.

Recently, RTP-like processing has found applications in another rapidly growing field — solar cell fabrication. RTP-like processing, in which an increase in the temperature of the semiconductor sample is produced by the absorption of the optical flux, is now used for a host of solar cell fabrication steps, including phosphorus diffusion for N/P junction formation and impurity gettering, hydrogen diffusion for impurity and defect passivation, and formation of screen-printed contacts using Ag-ink for the front and Al-ink for back contacts, respectively.

Furnace anneal

Furnace annealing is a process used in semiconductor device fabrication which consist of heating multiple semiconductor wafers in order to affect their electrical properties. Heat treatments are designed for different effects. Wafers can be heated in order to activate dopants, change film to film or film to wafer substrate interfaces, densify deposited films, change states of grown films, repair damage from implants, move dopants or drive dopants from one film into another or from a film into the wafer substrate.

Furnace anneals may be integrated into other furnace processing steps, such as oxidations, or may be processed on their own.

Furnace anneals are performed by equipment especially built to heat semiconductor wafers. Furnaces are capable of processing lots of wafers at a time but each process can last between several hours and a day.

Increasingly, furnace anneals are being supplanted by Rapid Thermal Anneal (RTA) or Rapid Thermal Processing (RTP). The reason for this is the relatively long thermal cycles of furnaces causes dopants that are being activated, especially boron, to diffuse farther than is intended. RTP or RTA fixes this by having thermal cycles for each wafer that is of the order of minutes rather than hours for furnace anneals.

Thermal oxidation

Furnaces used for diffusion and thermal oxidation at LAAS technological facility in Toulouse, France.

In microfabrication, **thermal oxidation** is a way to produce a thin layer of oxide (usually silicon dioxide) on the surface of a wafer (semiconductor). The technique forces an oxidizing agent to diffuse into the wafer at high temperature and react with it. The rate of oxide growth is often predicted by the Deal-Grove model.

The chemical reaction

Thermal oxidation of silicon is usually performed at a temperature between 800 and 1200°C, resulting in so called **High Temperature Oxide** layer (HTO). It may use either

water vapor (steam) or molecular oxygen as the oxidant; it is consequently called either *wet* or *dry* oxidation. The reaction is one of the following:

$$\mathrm{Si + 2H_2O \rightarrow SiO_2 + 2H_2\ (g)}$$
$$\mathrm{Si + O_2 \rightarrow SiO_2}$$

The oxidizing ambient may also contain several percent of hydrochloric acid (HCl). The chlorine removes metal ions that may occur in the oxide.

Thermal oxide incorporates silicon consumed from the substrate and oxygen supplied from the ambient. Thus, it grows both down into the wafer and up out of it. For every unit thickness of silicon consumed, 2.27 unit thicknesses of oxide will appear . Conversely, if a bare silicon surface is oxidized, 46% of the oxide thickness will lie below the original surface, and 54% above it.

Deal-Grove model

According to the commonly-used Deal-Grove model, the time *t* required to grow an oxide of thickness X_o, at a constant temperature, on a bare silicon surface, is:

$$t = \frac{X_o^2}{B} + \frac{X_o}{B/A}$$

where the constants A and B encapsulate the properties of the reaction and the oxide layer, respectively.

If a wafer that already contains oxide is placed in an oxidizing ambient, this equation must be modified by adding a corrective term τ, the time that would have been required to grow the pre-existing oxide under current conditions. This term may be found using the equation for *t* above.

Solving the quadratic equation for X_o yields:

$$X_o(t) = A/2 \cdot \left[\sqrt{1 + \frac{4B}{A^2}(t+\tau)} - 1\right]$$

Oxidation technology

Most thermal oxidation is performed in furnaces, at temperatures between 800 and 1200°C. A single furnace accepts many wafers at the same time, in a specially designed quartz rack (called a "boat"). Historically, the boat entered the oxidation chamber from the side (this design is called "horizontal"), and held the wafers vertically, beside each other. However, many modern designs hold the wafers horizontally, above and below each other, and load them into the oxidation chamber from below.

Vertical furnaces stand higher than horizontal furnaces, so they may not fit into some microfabrication facilities. However, they help to prevent dust contamination. Unlike horizontal furnaces, in which falling dust can contaminate any wafer, vertical furnaces only allow it to fall on the top wafer in the boat.

Vertical furnaces also eliminate an issue that plagued horizontal furnaces: non-uniformity of grown oxide across the wafer. Horizontal furnaces typically have convection currents inside the tube which causes the bottom of the tube to be slightly colder than the top of the tube. As the wafers lie vertically in the tube the convection and the temperature gradient with it causes the top of the wafer to have a thicker oxide than the bottom of the wafer. Vertical furnaces solve this problem by having wafer sitting horizontally, and then having the gas flow in the furnace flowing from top to bottom, significantly dampening any thermal convections.

Vertical furnaces also allow the use of load locks to purge the wafers with nitrogen before oxidation to limit the growth of native oxide on the Si surface.

Oxide quality

Wet oxidation is preferred to dry oxidation for growing thick oxides, because of the higher growth rate. However, fast oxidation leaves more dangling bonds at the silicon interface, which produce quantum states for electrons and allow current to leak along the interface. (This is called a "dirty" interface.) Wet oxidation also yields a lower-density oxide, with lower dielectric strength.

The long time required to grow a thick oxide in dry oxidation makes this process impractical. Thick oxides are usually grown with a long wet oxidation bracketed by short dry ones (a *dry-wet-dry* cycle). The beginning and ending dry oxidations produce films of high-quality oxide at the outer and inner surfaces of the oxide layer, respectively.

Mobile metal ions can degrade performance of MOSFETs (sodium is of particular concern). However, chlorine can immobilize sodium by forming sodium chloride. Chlorine is often introduced by adding hydrogen chloride or trichloroethylene to the oxidizing medium. Its presence also increases the rate of oxidation.

Other notes

- Thermal oxidation can be performed on selected areas of a wafer, and blocked on others. Areas which are not to be oxidized are covered with a film of silicon nitride, which blocks diffusion of oxygen and water vapor. The nitride is removed after oxidation is complete. This process cannot produce sharp features, because lateral (parallel to the surface) diffusion of oxidant molecules under the nitride mask causes the oxide to protrude into the masked area.

- Because impurities dissolve differently in silicon and oxide, a growing oxide will selectively take up or reject dopants. This redistribution is governed by the segregation coefficient, which determines how strongly the oxide absorbs or rejects the dopant, and the diffusivity.

- The orientation of the silicon crystal affects oxidation. A <100> wafer oxidizes more slowly than a <111> wafer, but produces an electrically cleaner oxide interface.

- Thermal oxidation of any variety produces a higher-quality oxide, with a much cleaner interface, than chemical vapor deposition of oxide resulting in **Low Temperature Oxide** layer (reaction of TEOS at about 600 °C). However, the high temperatures required to produce High Temperature Oxide (HTO) restrict its usability. For instance, in MOSFET processes, thermal oxidation is never performed after the doping for the source and drain terminals is performed, because it would disturb the placement of the dopants.

Chapter- 6

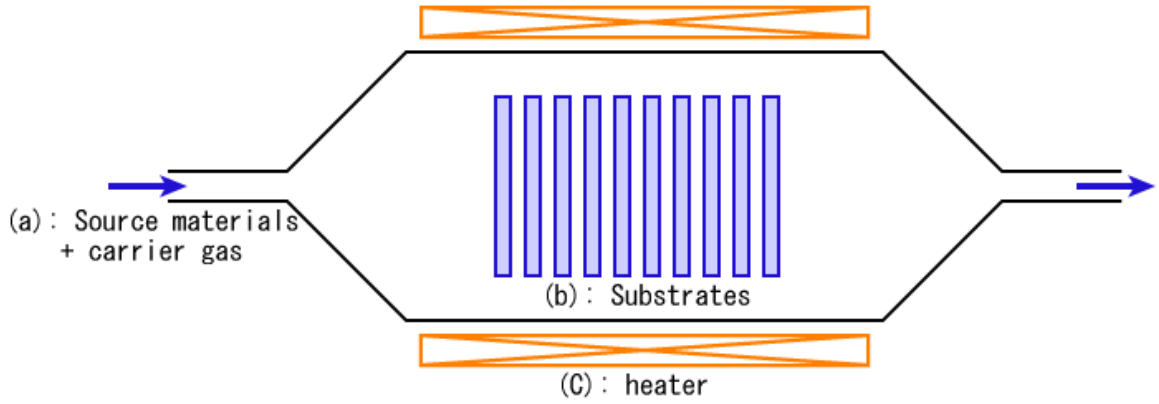
Chemical Vapor Deposition

DC plasma (violet) enhances the growth of carbon nanotubes in this laboratory-scale PECVD apparatus.

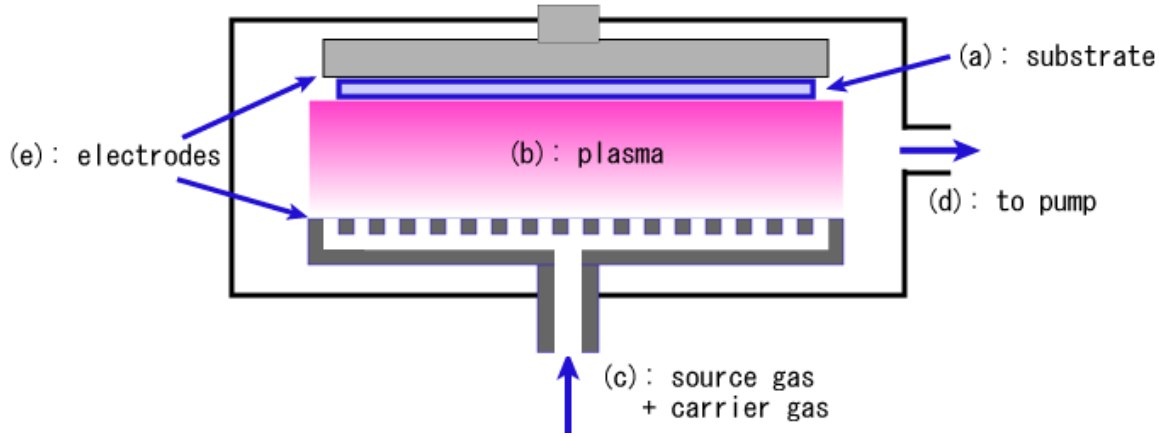
Chemical vapor deposition (CVD) is a chemical process used to produce high-purity, high-performance solid materials. The process is often used in the semiconductor industry to produce thin films. In a typical CVD process, the wafer (substrate) is exposed to one or more volatile precursors, which react and/or decompose on the substrate surface to produce the desired deposit. Frequently, volatile by-products are also produced, which are removed by gas flow through the reaction chamber.

Microfabrication processes widely use CVD to deposit materials in various forms, including: monocrystalline, polycrystalline, amorphous, and epitaxial. These materials include: silicon, carbon fiber, carbon nanofibers, filaments, carbon nanotubes, SiO_2, silicon-germanium, tungsten, silicon carbide, silicon nitride, silicon oxynitride, titanium nitride, and various high-k dielectrics. The CVD process is also used to produce synthetic diamonds.

Types of chemical vapor deposition



Hot-wall thermal CVD (batch operation type)



Plasma assisted CVD

A number of forms of CVD are in wide use and are frequently referenced in the literature. These processes differ in the means by which chemical reactions are initiated (e.g., activation process) and process conditions.

- Classified by operating pressure
  - *Atmospheric pressure CVD* (APCVD) - CVD processes at atmospheric pressure.
  - *Low-pressure CVD* (LPCVD) - CVD processes at subatmospheric pressures . Reduced pressures tend to reduce unwanted gas-phase reactions and improve film uniformity across the wafer. Most modern CVD processes are either LPCVD or UHVCVD.
  - *Ultrahigh vacuum CVD* (UHVCVD) - CVD processes at a very low pressure, typically below 10^{-6} Pa (~10^{-8} torr). Note that in other fields, a lower division between high and ultra-high vacuum is common, often 10^{-7} Pa.
- Classified by physical characteristics of vapor
  - *Aerosol assisted CVD* (AACVD) - A CVD process in which the precursors are transported to the substrate by means of a liquid/gas aerosol, which can be generated ultrasonically. This technique is suitable for use with non-volatile precursors.
  - *Direct liquid injection CVD* (DLICVD) - A CVD process in which the precursors are in liquid form (liquid or solid dissolved in a convenient solvent). Liquid solutions are injected in a vaporization chamber towards injectors (typically car injectors). Then the precursor vapors are transported to the substrate as in classical CVD process. This technique is suitable for use on liquid or solid precursors. High growth rates can be reached using this technique.
- Plasma methods
  - *Microwave plasma-assisted CVD* (MPCVD)
  - *Plasma-Enhanced CVD* (PECVD) - CVD processes that utilize plasma to enhance chemical reaction rates of the precursors . PECVD processing allows deposition at lower temperatures, which is often critical in the manufacture of semiconductors.
  - *Remote plasma-enhanced CVD* (RPECVD) - Similar to PECVD except that the wafer substrate is not directly in the plasma discharge region. Removing the wafer from the plasma region allows processing temperatures down to room temperature.
- *Atomic layer CVD* (ALCVD) – Deposits successive layers of different substances to produce layered, crystalline films.
- *Combustion Chemical Vapor Deposition* (CCVD) - nGimat's proprietary Combustion Chemical Vapor Deposition process is an open-atmosphere, flame-based technique for depositing high-quality thin films and nanomaterials.
- *Hot wire CVD* (HWCVD) - also known as catalytic CVD (Cat-CVD) or hot filament CVD (HFCVD). Uses a hot filament to chemically decompose the source gases.

- *Metalorganic chemical vapor deposition* (MOCVD) - CVD processes based on metalorganic precursors.
- *Hybrid Physical-Chemical Vapor Deposition* (HPCVD) - Vapor deposition processes that involve both chemical decomposition of precursor gas and vaporization of a solid source.
- *Rapid thermal CVD* (RTCVD) - CVD processes that use heating lamps or other methods to rapidly heat the wafer substrate. Heating only the substrate rather than the gas or chamber walls helps reduce unwanted gas phase reactions that can lead to particle formation.
- *Vapor phase epitaxy* (VPE)

Substances commonly deposited for ICs

Polysilicon

Polycrystalline silicon is deposited from silane (SiH_4), using the following reaction:

$$SiH_4 \rightarrow Si + 2\ H_2$$

This reaction is usually performed in LPCVD systems, with either pure silane feedstock, or a solution of silane with 70-80% nitrogen. Temperatures between 600 and 650 °C and pressures between 25 and 150 Pa yield a growth rate between 10 and 20 nm per minute. An alternative process uses a hydrogen-based solution. The hydrogen reduces the growth rate, but the temperature is raised to 850 or even 1050 °C to compensate.

Polysilicon may be grown directly with doping, if gases such as phosphine, arsine or diborane are added to the CVD chamber. Diborane increases the growth rate, but arsine and phosphine decrease it.

Silicon dioxide

Silicon dioxide (usually called simply "oxide" in the semiconductor industry) may be deposited by several different processes. Common source gases include silane and oxygen, dichlorosilane ($SiCl_2H_2$) and nitrous oxide (N_2O), or tetraethylorthosilicate (TEOS; $Si(OC_2H_5)_4$). The reactions are as follows:

$$SiH_4 + O_2 \rightarrow SiO_2 + 2\ H_2$$
$$SiCl_2H_2 + 2\ N_2O \rightarrow SiO_2 + 2\ N_2 + 2\ HCl$$
$$Si(OC_2H_5)_4 \rightarrow SiO_2 + \text{byproducts}$$

The choice of source gas depends on the thermal stability of the substrate; for instance, aluminium is sensitive to high temperature. Silane deposits between 300 and 500 °C, dichlorosilane at around 900 °C, and TEOS between 650 and 750 °C, resulting in a layer of *low- temperature oxide* (LTO). However, silane produces a lower-quality oxide than the other methods (lower dielectric strength, for instance), and it deposits nonconformally. Any of these reactions may be used in LPCVD, but the silane reaction is

also done in APCVD. CVD oxide invariably has lower quality than thermal oxide, but thermal oxidation can only be used in the earliest stages of IC manufacturing.

Oxide may also be grown with impurities (alloying or "doping"). This may have two purposes. During further process steps that occur at high temperature, the impurities may diffuse from the oxide into adjacent layers (most notably silicon) and dope them. Oxides containing 5–15% impurities by mass are often used for this purpose. In addition, silicon dioxide alloyed with phosphorus pentoxide ("P-glass") can be used to smooth out uneven surfaces. P-glass softens and reflows at temperatures above 1000 °C. This process requires a phosphorus concentration of at least 6%, but concentrations above 8% can corrode aluminium. Phosphorus is deposited from phosphine gas and oxygen:

$$4\ PH_3 + 5\ O_2 \rightarrow 2\ P_2O_5 + 6\ H_2$$

Glasses containing both boron and phosphorus (borophosphosilicate glass, BPSG) undergo viscous flow at lower temperatures; around 850 °C is achievable with glasses containing around 5 weight % of both constituents, but stability in air can be difficult to achieve. Phosphorus oxide in high concentrations interacts with ambient moisture to produce phosphoric acid. Crystals of BPO_4 can also precipitate from the flowing glass on cooling; these crystals are not readily etched in the standard reactive plasmas used to pattern oxides, and will result in circuit defects in integrated circuit manufacturing.

Besides these intentional impurities, CVD oxide may contain byproducts of the deposition process. TEOS produces a relatively pure oxide, whereas silane introduces hydrogen impurities, and dichlorosilane introduces chlorine.

Lower temperature deposition of silicon dioxide and doped glasses from TEOS using ozone rather than oxygen has also been explored (350 to 500 °C). Ozone glasses have excellent conformality but tend to be hygroscopic – that is, they absorb water from the air due to the incorporation of silanol (Si-OH) in the glass. Infrared spectroscopy and mechanical strain as a function of temperature are valuable diagnostic tools for diagnosing such problems.

Silicon nitride

Silicon nitride is often used as an insulator and chemical barrier in manufacturing ICs. The following two reactions deposit nitride from the gas phase:

$$3\ SiH_4 + 4\ NH_3 \rightarrow Si_3N_4 + 12\ H_2$$
$$3\ SiCl_2H_2 + 4\ NH_3 \rightarrow Si_3N_4 + 6\ HCl + 6\ H_2$$

Silicon nitride deposited by LPCVD contains up to 8% hydrogen. It also experiences strong tensile stress, which may crack films thicker than 200 nm. However, it has higher resistivity and dielectric strength than most insulators commonly available in microfabrication (10^{16} Ω·cm and 10 MV/cm, respectively).

Another two reactions may be used in plasma to deposit SiNH:

$$2\ SiH_4 + N_2 \rightarrow 2\ SiNH + 3\ H_2$$
$$SiH_4 + NH_3 \rightarrow SiNH + 3\ H_2$$

These films have much less tensile stress, but worse electrical properties (resistivity 10^6 to 10^{15} Ω·cm, and dielectric strength 1 to 5 MV/cm).

Metals

Some metals (notably aluminium and copper) are seldom or never deposited by CVD. As of 2010, a commercially, cost effective, viable CVD process for copper did not exist, though copper formate, copper(hfac)2, Cu(II) ethyl acetoacetate, and other precursors have been used. Copper deposition of the metal has been done mostly by electroplating, in order to reduce the cost. Aluminum can be deposited from tri-isobutyl aluminium (TIBAL), tri ethyl/methyl aluminum (TEA,TMA), or dimethylaluminum hydride (DMAH), but physical vapor deposition methods are usually preferred.

However, CVD processes for molybdenum, tantalum, titanium, nickel, and tungsten are widely used . These metals can form useful silicides when deposited onto silicon. Mo, Ta and Ti are deposited by LPCVD, from their pentachlorides. Nickel, molybdenum, and tungsten can be deposited at low temperatures from their carbonyl precursors. In general, for an arbitrary metal *M*, the reaction is as follows:

$$2\ MCl_5 + 5\ H_2 \rightarrow 2\ M + 10\ HCl$$

The usual source for tungsten is tungsten hexafluoride, which may be deposited in two ways:

$$WF_6 \rightarrow W + 3\ F_2$$
$$WF_6 + 3\ H_2 \rightarrow W + 6\ HF$$

Chapter- 7

Physical Vapor Deposition

Physical vapor deposition (**PVD**) is a variety of vacuum deposition and is a general term used to describe any of a variety of methods to deposit thin films by the condensation of a vaporized form of the material onto various surfaces (e.g., onto semiconductor wafers). The coating method involves purely physical processes such as high temperature vacuum evaporation or plasma sputter bombardment rather than involving a chemical reaction at the surface to be coated as in chemical vapor deposition. The term physical vapor deposition appears originally in the 1966 book *Vapor Deposition* by CF Powell, JH Oxley and JM Blocher Jr, but Michael Faraday was using PVD to deposit coatings as far back as 1838.

Variants of PVD include, in order of increasing novelty:

Cathodic arc deposition

Cathodic arc deposition or **Arc-PVD** is a physical vapor deposition technique in which an electric arc is used to vaporize material from a cathode target. The vaporized material then condenses on a substrate, forming a thin film. The technique can be used to deposit metallic, ceramic, and composite films.

History

Industrial use of modern cathodic arc deposition technology originated in Soviet Union around 1960–1970. By the late 70's Soviet government released the use of this technology to the West. Among many designs in USSR at that time the design by L. P. Sablev, et al., was allowed to be used outside the USSR.

Process

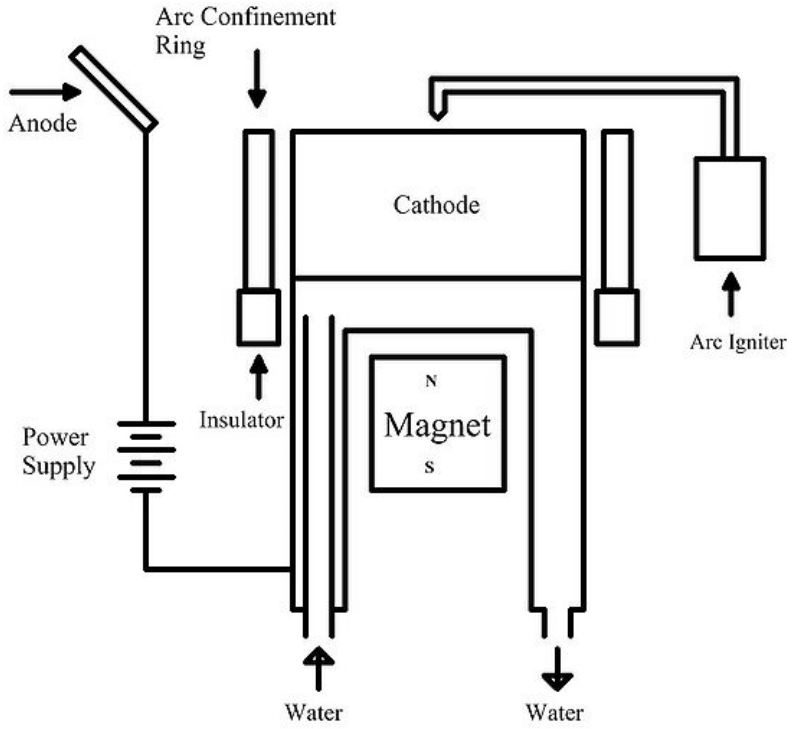
The arc evaporation process begins with the striking of a high current, low voltage arc on the surface of a cathode (known as the target) that gives rise to a small (usually a few micrometres wide), highly energetic emitting area known as a cathode spot. The localised temperature at the cathode spot is extremely high (around 15000 °C), which results in a high velocity (10 km/s) jet of vapourised cathode material, leaving a crater behind on the cathode surface. The cathode spot is only active for a short period of time, then it self-extinguishes and re-ignites in a new area close to the previous crater. This behaviour causes the apparent motion of the arc.

As the arc is basically a current carrying conductor it can be influenced by the application of an electromagnetic field, which in practice is used to rapidly move the arc over the entire surface of the target, so that the total surface is eroded over time.

The arc has an extremely high power density resulting in a high level of ionization (30-100%), multiple charged ions, neutral particles, clusters and macro-particles (droplets). If a reactive gas is introduced during the evaporation process, dissociation, ionization and excitation can occur during interaction with the ion flux and a compound film will be deposited.

One downside of the arc evaporation process is that if the cathode spot stays at an evaporative point for too long it can eject a large amount of macro-particles or droplets. These droplets are detrimental to the performance of the coating as they are poorly adhered and can extend through the coating. Worse still if the cathode target material has a low melting point such as aluminium the cathode spot can evaporate through the target resulting in either the target backing plate material being evaporated or cooling water entering the chamber. Therefore magnetic fields as mentioned previously are used to control the motion of the arc. If cylindrical cathodes are used the cathodes can also be rotated during deposition. By not allowing the cathode spot to remain in one position too long aluminium targets can be used and the number of droplets is reduced. Some companies also use filtered arcs that use magnetic fields to separate the droplets from the coating flux.

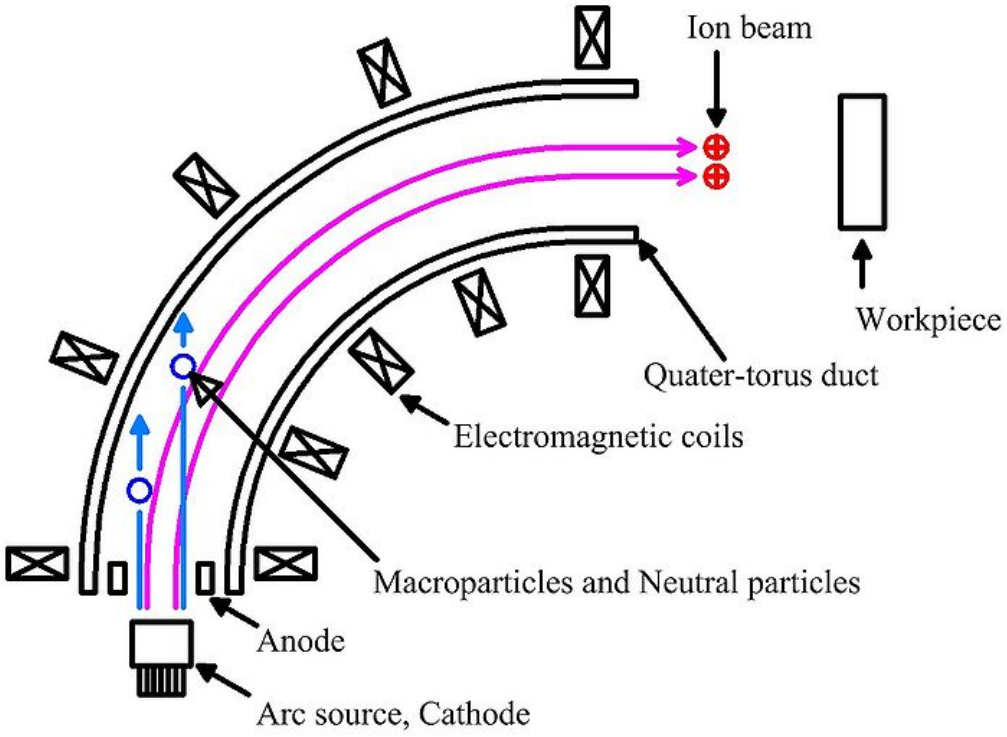
Equipment design



Cathodic arc source (Sablev type)

Sablev type Cathodic arc source with magnet to steer the movement of arc spot

Sablev type Cathodic arc source, which is the most widely used in the West, consists of a short cylindrical shape electrical conductive target at cathode with one open end. This target has an electrically float metal ring surrounded working as an arc confinement ring. The anode for the system can be either the vacuum chamber wall or a discrete anode. Arc spots are generated by mechanical trigger (or igniter) striking on open end of the target making a temporarily short circuit between the cathode and anode. After the arc spots being generated they can be steered by magnetic field or move randomly in absence of magnetic field.



Aksenov's quater-torus macroparticle filter

Aksenov Quarter-torus duct macroparticle filter using plasma optical principles which was developed by A. I. Morozov

The plasma beam from Cathodic Arc source contains some larger clusters of atoms or molecules (so called macro-particles), which prevent it from being useful for some applications without some kind of filtering. There are many designs for macro-particle filters and the most studied design is based on the work by I. I. Aksenov et al. in 70's. It consists of a quarter-torus duct bent at 90 degrees from the arc source and the plasma is guided out of the duct by principle of plasma optics.

There are also other interesting designs such as a design which incorporates a straight duct filter built-in with truncated cone shape cathode as reported by D. A. Karpov in the 90's. This design became quite popular among both the thin hard-film coaters and researchers in Russia and former USSR countries until now. Cathodic arc source can be made into the long tubular shape (extended-arc) or long rectangular shape but both designs are less popular.

Applications

Titanium Nitride (TiN) coated punches using Cathodic arc deposition technique

Aluminium Titanium Nitride (AlTiN) coated Endmills using Cathodic arc deposition technique

Aluminium Chromium Titanium Nitride (AlCrTiN) coated Hob using Cathodic arc deposition technique

Cathodic arc deposition is actively used to synthesize extremely hard film to protect the surface of cutting tools and extend their life significantly. A wide variety of thin hard-film, Superhard coatings and nanocomposite coatings can be synthesized by this technology including TiN, TiAlN, CrN, ZrN, AlCrTiN and TiAlSiN.

This is also used quite extensively particularly for carbon ion deposition to create diamond-like carbon films. Because the ions are blasted from the surface ballistically, it is common for not only single atoms, but larger clusters of atoms to be ejected. Thus, this kind of system requires a filter to remove atom clusters from the beam before deposition. The DLC film from filtered-arc contains extremely high percentage of sp^3 diamond which is known as tetrahedral amorphous carbon, or ta-C.

Filtered Cathodic arc can be used as metal ion/plasma source for *Ion implantation* and *Plasma Immersion Ion Implantation and Deposition* (PIII&D).

Electron beam physical vapor deposition

Electron Beam Physical Vapor Deposition or **EBPVD** is a form of physical vapor deposition in which a target anode is bombarded with an electron beam given off by a charged tungsten filament under high vacuum. The electron beam causes atoms from the target to transform into the gaseous phase. These atoms then precipitate into solid form, coating everything in the vacuum chamber (within line of sight) with a thin layer of the anode material.

Introduction

Thin film deposition is a process applied in the semiconductor industry to grow electronic materials, in the aerospace industry to form thermal and chemical barrier coatings to protect surfaces against corrosive environments, in optics to impart the desired reflective and transmissive properties to a substrate and elswhere in industry to modify surfaces to have a variety desired properties. The deposition process can be broadly classified into physical vapor deposition (PVD) and chemical vapor deposition (CVD). In CVD, the film growth takes place at high temperatures, leading to the formation of corrosive gaseous products, and it may leave impurities in the film. The PVD process can be carried out at lower deposition temperatures and without corrosive products, but deposition rates are typically lower. Electron beam physical vapor deposition, however, yields a high deposition rate from 0.1 μm / min to 100 μm / min at relatively low substrate temperatures, with very high material utilization efficiency. The schematic of an EBPVD system is shown in Fig 1.

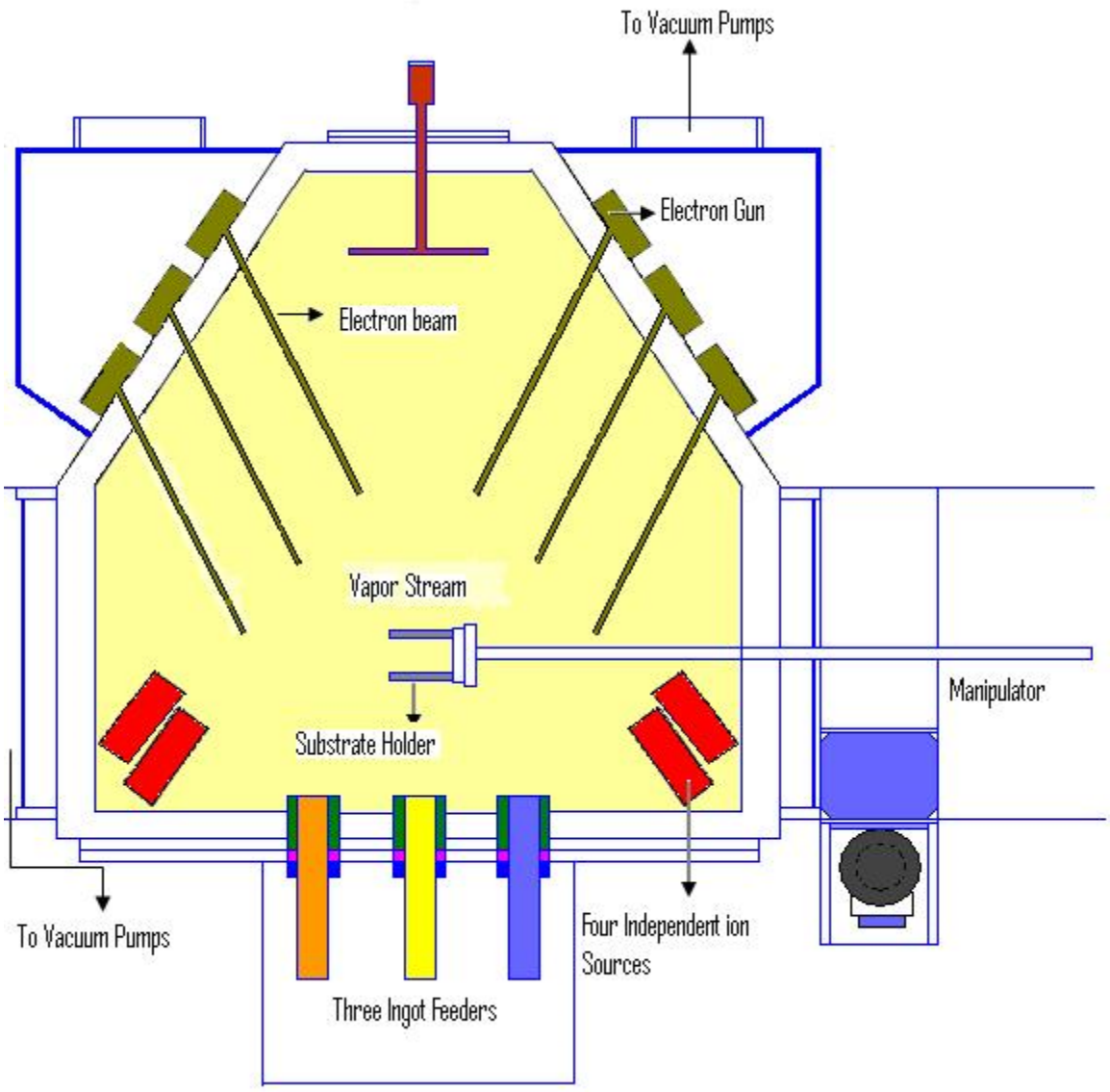


Fig 1 Electron Beam Physical Vapor Deposition

Thin film deposition process

In an EBPVD system, the deposition chamber is evacuated to a pressure of 10^{-4} Torr. The material to be evaporated is in the form of ingots. There are as many as six electron guns, each having a power from tens to hundreds of kW. Electron beams can be generated by thermionic emission, field electron emission or the anodic arc method. The generated electron beam is accelerated to a high kinetic energy and focused towards the ingot. When the accelerating voltage is between 20 kV – 25 kV and the beam current is a few amperes, 85% of the kinetic energy of the electrons is converted into thermal energy as the beam bombards the surface of the ingot. The surface temperature of the ingot increases resulting in the formation of a liquid melt. Although some of incident electron energy is lost in the excitation of X-rays and secondary emission, the liquid ingot material evaporates under vacuum.

The ingot itself is enclosed in a copper crucible, which is cooled by water circulation. The level of molten liquid pool on the surface of the ingot is kept constant by vertical displacement of the ingot. The number of ingot feeders depends upon the material to be deposited. The evaporation rate may be of the order of 10^{-2} g/cm2 sec.

Material evaporation methods

Refractory carbides like titanium carbide and borides like titanium boride and zirconium boride can evaporate without undergoing decomposition in the vapor phase. These compounds are deposited by direct evaporation. In this process these compounds, compacted in the form of an ingot, are evaporated in vacuum by the focused high energy electron beam and the vapors are directly condensed over the substrate.

Certain refractory oxides and carbides undergo fragmentation during their evaporation by the electron beam, resulting in a stoichiometry that is different from the initial material. For example, alumina, when evaporated by electron beam, dissociates into aluminum, AlO_3 and Al_2O. Some refractory carbides like silicon carbide and tungsten carbide decompose upon heating and the dissociated elements have different volatilities. These compounds can be deposited on the substrate either by reactive evaporation or by co-evaporation. In the reactive evaporation process, the metal is evaporated from the ingot by the electron beam. The vapors are carried by the reactive gas, which is oxygen in case of metal oxides or acetylene in case of metal carbides. When the thermodynamic conditions are met, the vapors react with the gas in the vicinity of the substrate to form films. Metal carbide films can also be deposited by co-evaporation. In this process, two ingots are used, one for metal and the other for carbon. Each ingot is heated with a different beam energy so that their evaporation rate can be controlled. As the vapors arrive at the surface, they chemically combine under proper thermodynamic conditions to form a metal carbide film.

The substrate

The substrate on which the film deposition takes place is ultrasonically cleaned and fastened to the substrate holder. The substrate holder is attached to the manipulator shaft. The manipulator shaft moves translationally to adjust the distance between the ingot source and the substrate. The shaft also rotates the substrate at a particular speed so that the film is uniformly deposited on the substrate. A negative bias D.C. voltage of 200 V – 400 V can be applied to the substrate. Often, focused high energy electrons from one of the electron guns or infrared light from heater lamps is used to preheat the substrate.

Ion beam assisted deposition

EBPVD systems are equipped with ion sources. These ion sources are used for substrate etching and cleaning, sputtering the target and controlling the microstructure of the substrate. The ion beams bombard the surface and alter the microstructure of the film. When the deposition reaction takes place on the hot substrate surface, the films can

develop an internal tensile stress due to the mismatch in the coefficient of thermal expansion between the substrate and the film. High energy ions can be used to bombard these ceramic thermal barrier coatings and change the tensile stress into compressive stress. Ion bombardment also increases the density of the film, changes the grain size and modifies amorphous films to polycrystalline films. Low energy ions are used for the surfaces of semiconductor films.

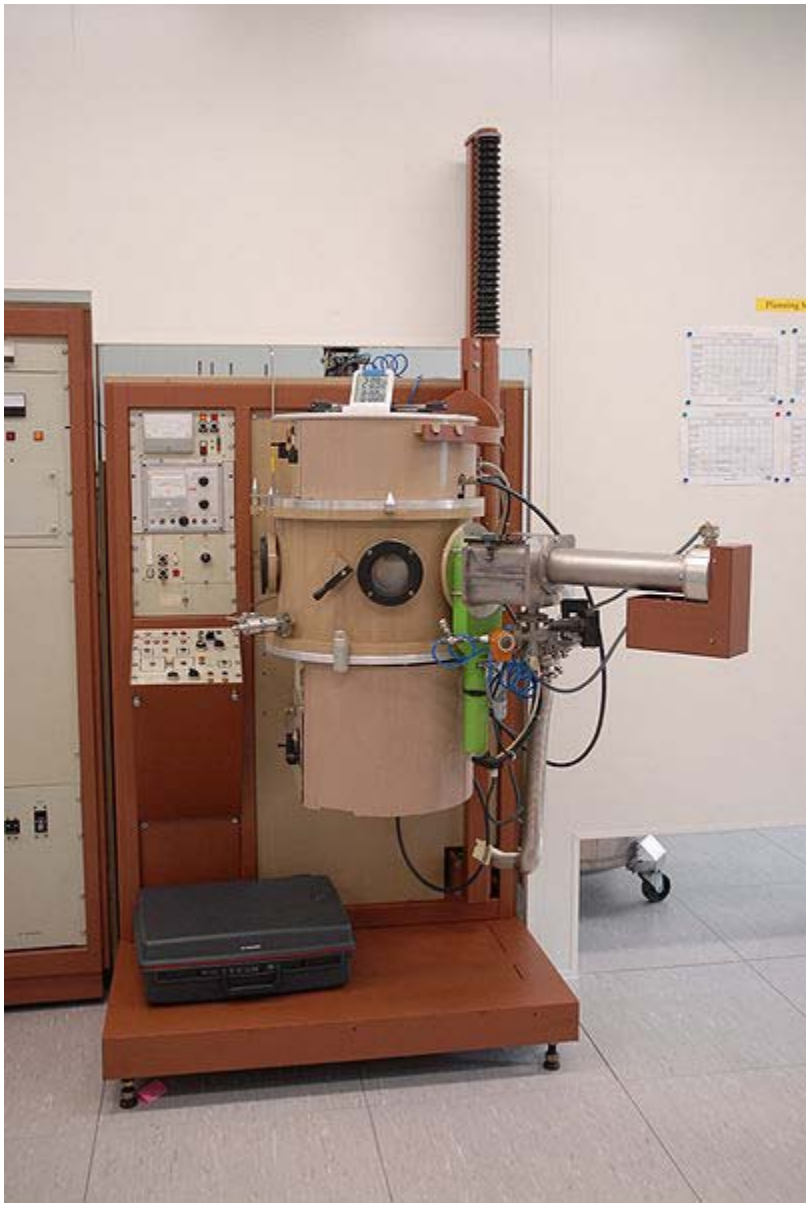
Advantages of EBPVD

The deposition rate in this process can be as low as 1 nm per minute to as high as few micrometers per minute. The material utilization efficiency is high relative to other methods and the process offers structural and morphological control of films. Due to the very high deposition rate, this process has potential industrial application for wear resistant and thermal barrier coatings in aerospace industries, hard coatings for cutting and tool industries, and electronic and optical films for semiconductor industries.

Disadvantages of EBPVD

EBPVD is a line-of-sight of deposition process. The translational and rotational motion of the shaft helps for coating the outer surface of complex geometries, but this process cannot be used to coat the inner surface of complex geometries. Another potential problem is that filament degradation in the electron gun results in a non-uniform evaporation rate.

Evaporation (deposition)



Evaporation machine used for metallization at LAAS technological facility in Toulouse, France.

Evaporation is a common method of thin film deposition. The source material is evaporated in a vacuum. The vacuum allows vapor particles to travel directly to the target object (substrate), where they condense back to a solid state. Evaporation is used in microfabrication, and to make macro-scale products such as metallized plastic film.

Physical principle

Evaporation involves two basic processes: a hot source material evaporates and condenses on the substrate. It resembles the familiar process by which liquid water appears on the lid of a boiling pot. However, the gaseous environment and heat source are different.

Evaporation takes place in a vacuum, i.e. vapors other than the source material are almost entirely removed before the process begins. In high vacuum (with a long mean free path), evaporated particles can travel directly to the deposition target without colliding with the background gas. (By contrast, in the boiling pot example, the water vapor pushes the air out of the pot before it can reach the lid.) At a typical pressure of 10^{-4} Pa, an 0.4-nm particle has a mean free path of 60 m. Hot objects in the evaporation chamber, such as heating filaments, produce unwanted vapors that limit the quality of the vacuum.

Evaporated atoms that collide with foreign particles may react with them; for instance, if aluminium is deposited in the presence of oxygen, it will form aluminium oxide. They also reduce the amount of vapor that reaches the substrate, which makes the thickness difficult to control.

Evaporated materials deposit nonuniformly if the substrate has a rough surface (as integrated circuits often do). Because the evaporated material attacks the substrate mostly from a single direction, protruding features block the evaporated material from some areas. This phenomenon is called "shadowing" or "step coverage."

When evaporation is performed in poor vacuum or close to atmospheric pressure, the resulting deposition is generally non-uniform and tends not to be a continuous or smooth film. Rather, the deposition will appear fuzzy.

Equipment

Any evaporation system includes a vacuum pump. It also includes an energy source that evaporates the material to be deposited. Many different energy sources exist:

- In the *thermal* method, metal wire is fed onto heated semimetal (ceramic) evaporators known as "boats" due to their shape. A pool of melted metal forms in the boat cavity and evaporates into a cloud above the source. Alternatively the source material is placed in a crucible, which is radiatively heated by an electric filament, or the source material may be hung from the filament itself (*filament evaporation*).
    - Molecular beam epitaxy is an advanced form of thermal evaporation.
- In the *electron-beam* method, the source is heated by an electron beam with an energy up to 15 keV.
- In *flash evaporation*, a fine wire of source material is fed continuously onto a hot ceramic bar, and evaporates on contact.

- *Resistive evaporation* is accomplished by passing a large current through a resistive wire or foil containing the material to be deposited.

Some systems mount the substrate on an out-of-plane planetary mechanism. The mechanism rotates the substrate simultaneously around two axes, to reduce shadowing.

Optimization

- Purity of the deposited film depends on the quality of the vacuum, and on the purity of the source material.
- The thickness of the film will vary due to the geometry of the evaporation chamber. Collisions with residual gases aggravate nonuniformity of thickness.
- Filament and resistive evaporation cannot deposit thick films, because the size of the filament limits the amount of material that can be deposited. However, flash evaporation and methods that use crucibles can deposit thick films.
- In order to deposit a material, the evaporation system must be able to melt it. This makes refractory materials such as tungsten hard to deposit by methods that do not use electron-beam heating.
- Electron-beam evaporation allows tight control of the evaporation rate. Thus, an electron-beam system with multiple beams and multiple sources can deposit a chemical compound or composite material of known composition.
- Step coverage

Applications

An important example of an evaporative process is the production of aluminized PET film packaging film in a roll-to-roll web system. Often, the aluminum layer in this material is not thick enough to be entirely opaque since a thinner layer can be deposited more cheaply than a thick one. The main purpose of the aluminum is to isolate the product from the external environment by creating a barrier to the passage of light, oxygen, or water vapor.

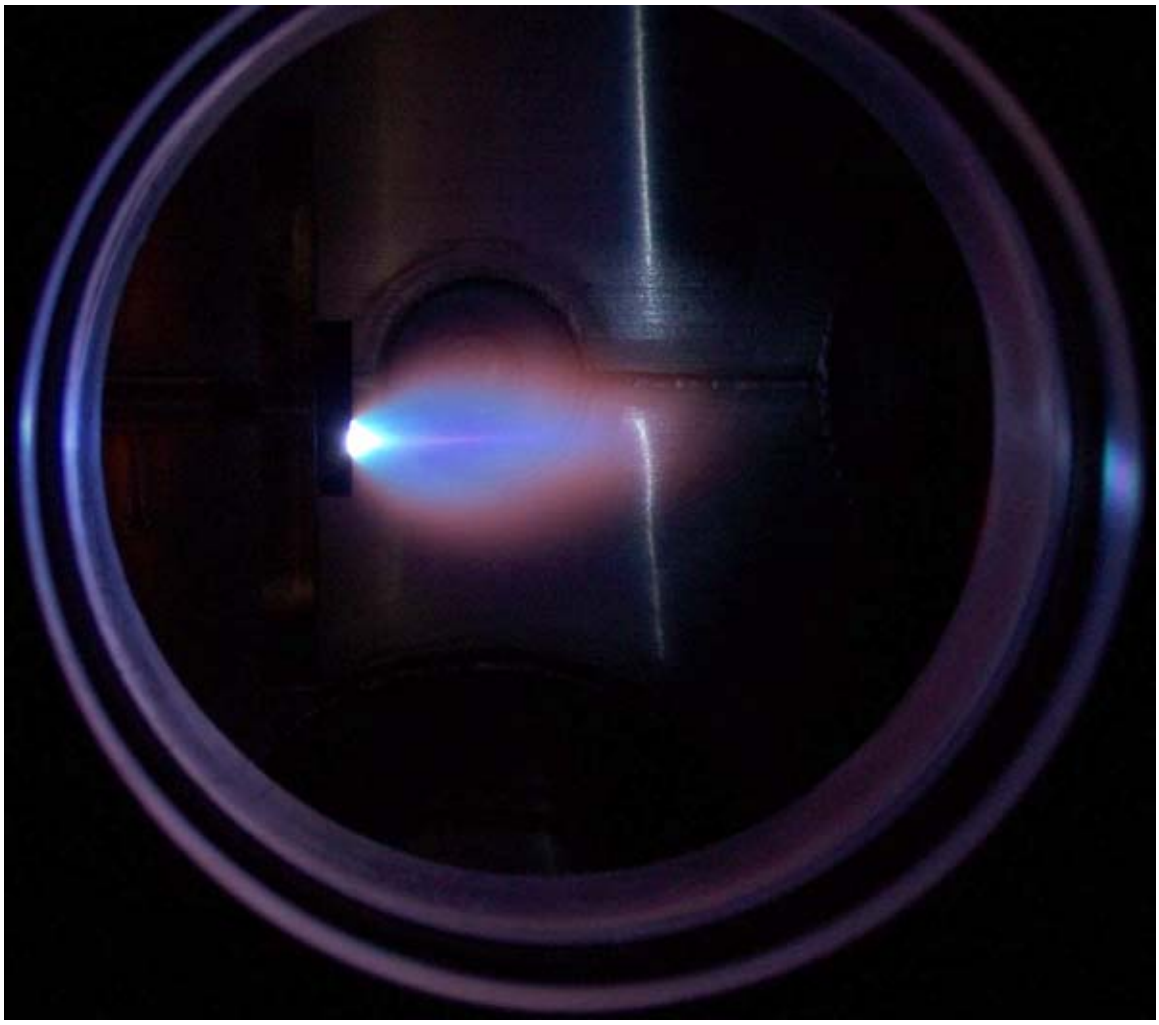
Evaporation is commonly used in microfabrication to deposit metal films.

Comparison to other deposition methods

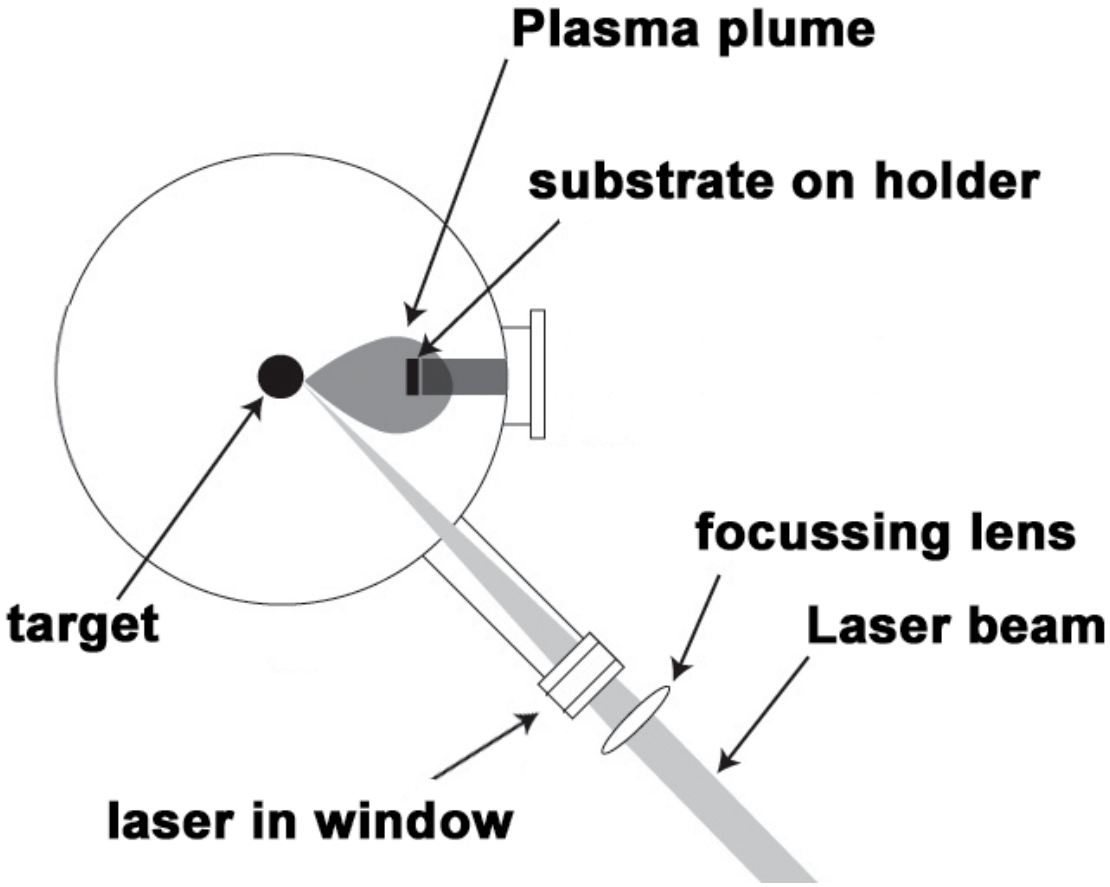
- Alternatives to evaporation, such as sputtering and chemical vapor deposition, have better step coverage. This may be an advantage or disadvantage, depending on the desired result.
- Sputtering tends to deposit material more slowly than evaporation.
- Sputtering uses a plasma, which produces many high-speed atoms that bombard the substrate and may damage it. Evaporated atoms have a Maxwellian energy distribution, determined by the temperature of the source, which reduces the number of high-speed atoms. However, electron beams tend to produce X-rays

(Bremsstrahlung) and stray electrons, each of which can also damage the substrate.

Pulsed laser deposition



A plume ejected from a $SrRuO_3$ target during pulsed laser deposition.



One possible configuration of a PLD deposition chamber.

Pulsed laser deposition (PLD) is a thin film deposition (specifically a physical vapor deposition, PVD) technique where a high power pulsed laser beam is focused inside a vacuum chamber to strike a target of the material that is to be deposited. This material is vaporized from the target (in a plasma plume) which deposits it as a thin film on a substrate (such as a silicon wafer facing the target). This process can occur in ultra high vacuum or in the presence of a background gas, such as oxygen which is commonly used when depositing oxides to fully oxygenate the deposited films.

While the basic-setup is simple relative to many other deposition techniques, the physical phenomena of laser-target interaction and film growth are quite complex. When the laser pulse is absorbed by the target, energy is first converted to electronic excitation and then into thermal, chemical and mechanical energy resulting in evaporation, ablation, plasma formation and even exfoliation . The ejected species expand into the surrounding vacuum in the form of a plume containing many energetic species including atoms, molecules, electrons, ions, clusters, particulates and molten globules, before depositing on the typically hot substrate.

Process

The detailed mechanisms of PLD are very complex including the ablation process of the target material by the laser irradiation, the development of a plasma plume with high energetic ions, electrons as well as neutrals and the crystalline growth of the film itself on the heated substrate. The process of PLD can generally be divided into four stages:

- Laser ablation of the target material and creation of a plasma
- Dynamic of the plasma
- Deposition of the ablation material on the substrate
- Nucleation and growth of the film on the substrate surface

Each of these steps is crucial for the crystallinity, uniformity and stoichiometry of the resulting film.

Laser ablation of the target material and creation of a plasma

The ablation of the target material upon laser irradiation and the creation of plasma are very complex processes. The removal of atoms from the bulk material is done by vaporization of the bulk at the surface region in a state of non-equilibrium and is caused by a Coulomb explosion. In this the incident laser pulse penetrates into the surface of the material within the penetration depth. This dimension is dependent on the laser wavelength and the index of refraction of the target material at the applied laser wavelength and is typically in the region of 10 nm for most materials. The strong electrical field generated by the laser light is sufficiently strong to remove the electrons from the bulk material of the penetrated volume. This process occurs within 10 ps of a ns laser pulse and is caused by non-linear processes such as multiphoton ionization which are enhanced by microscopic cracks at the surface, voids, and nodules, which increase the electric field. The free electrons oscillate within the electromagnetic field of the laser light and can collide with the atoms of the bulk material thus transferring some of their energy to the lattice of the target material with in the surface region. The surface of the target is then heated up and the material is vaporized.

Dynamic of the plasma

In the second stage the material expands in a plasma parallel to the normal vector of the target surface towards the substrate due to Coulomb repulsion and recoil from the target surface. The spatial distribution of the plume is dependent on the background pressure inside the PLD chamber. The density of the plume can be described by a $\cos^n(x)$ law with a shape similar to a Gaussian curve. The dependency of the plume shape on the pressure can be described in three stages:

- The vacuum stage, where the plume is very narrow and forward directed; almost no scattering occurs with the background gases.
- The intermediate region where a splitting of the high energetic ions from the less energetic species can be observed. The time-of-flight (TOF) data can be fitted to a shock wave model; however, other models could also be possible.

- High pressure region where we find a more diffusion-like expansion of the ablated material. Naturally this scattering is also dependent on the mass of the background gas and can influence the stoichiometry of the deposited film.

The most important consequence of increasing the background pressure is the slowing down of the high energetic species in the expanding plasma plume. It has been shown that particles with kinetic energies around 50 eV can resputter the film already deposited on the substrate. This results in a lower deposition rate and can furthermore result in a change in the stoichiometry of the film.

Deposition of the ablation material on the substrate

The third stage is important to determine the quality of the deposited films. The high energetic species ablated from the target are bombarding the substrate surface and may cause damage to the surface by sputtering off atoms from the surface but also by causing defect formation in the deposited film. The sputtered species from the substrate and the particles emitted from the target form a collision region, which serves as a source for condensation of particles. When the condensation rate is high enough, a thermal equilibrium can be reached and the film grows on the substrate surface at the expense of the direct flow of ablation particles and the thermal equilibrium obtained..

Nucleation and growth of the film on the substrate surface

The nucleation process and growth kinetics of the film depend on several growth parameters including:

- *Laser parameters* – several factors such as the laser fluence [Joule/cm^2], laser energy, and ionization degree of the ablated material will affect the film quality, the stoichiometry , and the deposition flux. Generally, the nucleation density increases when the deposition flux is increased.
- *Surface temperature* – The surface temperature has a large affect on the nucleation density. Generally, the nucleation density decreases as the temperature is increased .
- *Substrate surface* – The nucleation and growth can be affected by the surface preparation (such as chemical etching), the miscut of the substrate, as well as the roughness of the substrate.
- *Background pressure* – Common in oxide deposition, an oxygen background is needed to ensure stoichiometric transfer from the target to the film. If, for example, the oxygen background is too low, the film will grow off stoichiometry which will affect the nucleation density and film quality .

In PLD, a large supersaturation occurs on the substrate during the pulse duration. The pulse lasts around 10–40 microseconds depending on the laser parameters. This high supersaturation causes a very large nucleation density on the surface as compared to Molecular Beam Epitaxy or Sputtering Deposition. This nucleation density increases the smoothness of the deposited film.

In PLD, [depending on the deposition parameters above] three growth modes are possible:

- *Step-flow growth* – All substrates have a miscut associated with the crystal. These miscuts give rise to atomic steps on the surface. In step-flow growth, atoms land on the surface and diffuse to a step edge before they have a chance to nucleated a surface island. The growing surface is viewed as steps traveling across the surface. This growth mode is obtained by deposition on a high miscut substrate, or depositing at elevated temperatures
- *Layer-by-layer growth* – In this growth mode, islands nucleate on the surface until a critical island density is reached. As more material is added, the islands continue to grow until the islands begin to run into each other. This is know a coalescence. Once coalescence is reached, the surface has a large density of pits. When additional material is added to the surface the atoms diffuse into these pits to complete the layer. This process is repeated for each subsequent layer.
- *3D growth* – This mode is similar to the layer-by-layer growth, except that once an island is formed an additional island will nucleate on top of the 1st island. Therefore the growth does not persist in a layer by layer fashion, and the surface roughens each time material is added.

History

Pulsed laser deposition is only one of many thin film deposition techniques. Other methods include molecular beam epitaxy (MBE), chemical vapor deposition (CVD), sputter deposition (RF, Magnetron, and ion beam). The history of laser-assisted film growth started soon after the technical realization of the first laser in 1960 by Maiman. Smith and Turner utilized a ruby laser to deposit the first thin films in 1965, three years after Breech and Cross studied the laser-vaporization and excitation of atoms from solid surfaces. However, the deposited films were still inferior to those obtained by other techniques such as chemical vapor deposition and molecular beam epitaxy. In the early 1980s, a few research groups (mainly in the former USSR) achieved remarkable results on manufacturing of thin film structures utilizing laser technology. The breakthrough came in 1987 when Dijkkamp and Venkatesan were able to laser deposit a thin film of $YBa_2Cu_3O_7$, a high temperature superconductive material, which was of more superior quality than films deposited with alternative techniques. Since then, the technique of Pulsed Laser Deposition has been utilized to fabricate high quality crystalline films. The deposition of ceramic oxides, nitride films, metallic multilayers and various superlattices has been demonstrated. In the 1990s the development of new laser technology, such as lasers with high repetition rate and short pulse durations, made PLD a very competitive tool for the growth of thin, well defined films with complex stoichiometry.

Technical aspects

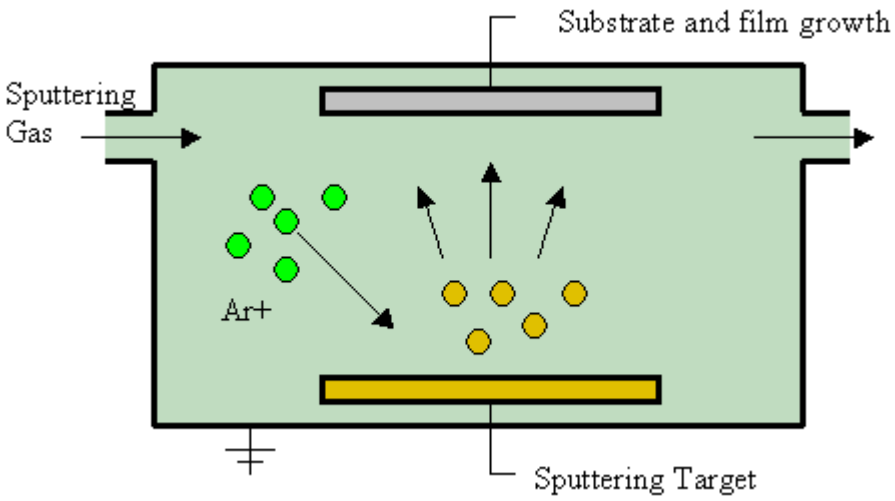
There are many different arrangements to build a deposition chamber for PLD. The target material which is evaporated by the laser is normally found as a rotating disc attached to

a support. However, it can also be sintered into a cylindrical rod with rotational motion and a translational up and down movement along its axis. This special configuration allows not only the utilization of a synchronized reactive gas pulse but also of a multicomponent target rod with which films of different multilayers can be created.

Some factors that influence deposition thickness:

- Target material
- Pulse energy of laser
- Distance from target to substrate
- Type of gas and pressure in chamber (oxygen, argon, etc.)

Sputter deposition



Sputter deposition is a physical vapor deposition (PVD) method of depositing thin films by sputtering, that is ejecting, material from a "target," that is source, which then deposits onto a "substrate," such as a silicon wafer. Resputtering is re-emission of the deposited material during the deposition process by ion or atom bombardment. Sputtered atoms ejected from the target have a wide energy distribution, typically up to tens of eV (100000 K). The sputtered ions (typically only a small fraction — order 1% — of the ejected particles are ionized) can ballistically fly from the target in straight lines and impact energetically on the substrates or vacuum chamber (causing resputtering). Alternatively, at higher gas pressures, the ions collide with the gas atoms that act as a moderator and move diffusively, reaching the substrates or vacuum chamber wall and condensing after undergoing a random walk. The entire range from high-energy ballistic impact to low-energy thermalized motion is accessible by changing the background gas pressure. The sputtering gas is often an inert gas such as argon. For efficient momentum transfer, the atomic weight of the sputtering gas should be close to the atomic weight of

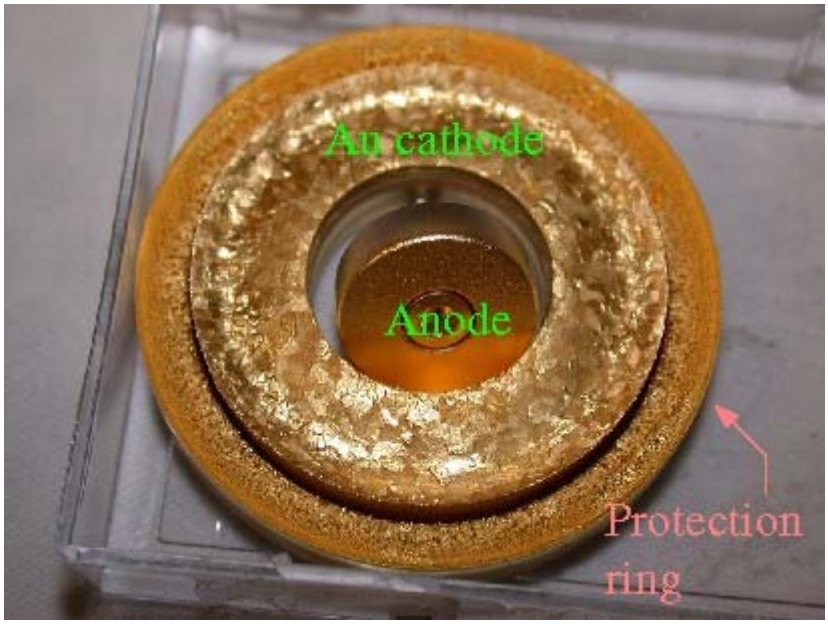
the target, so for sputtering light elements neon is preferable, while for heavy elements krypton or xenon are used. Reactive gases can also be used to sputter compounds. The compound can be formed on the target surface, in-flight or on the substrate depending on the process parameters. The availability of many parameters that control sputter deposition make it a complex process, but also allow experts a large degree of control over the growth and microstructure of the film.

Uses of sputtering

Sputtering is used extensively in the semiconductor industry to deposit thin films of various materials in integrated circuit processing. Thin antireflection coatings on glass for optical applications are also deposited by sputtering. Because of the low substrate temperatures used, sputtering is an ideal method to deposit contact metals for thin-film transistors. Perhaps the most familiar products of sputtering are low-emissivity coatings on glass, used in double-pane window assemblies. The coating is a multilayer containing silver and metal oxides such as zinc oxide, tin oxide, or titanium dioxide. Sputtering is also used to metalize plastics such as potato chip bags. A large industry has developed around tool bit coating using sputtered nitrides, such as titanium nitride, creating the familiar gold colored hard coat. Sputtering is also used as the process to deposit the metal (e.g. aluminium) layer during the fabrication of CD and DVD discs.

Hard disk surfaces use sputtered CrO_x and other sputtered materials. Sputtering is one of the main processes of manufacturing optical waveguides and is another way for making efficient photovoltaic solar cells.

Comparison with other deposition methods



A typical ring-geometry sputter target, here gold showing the cathode made of the material to be deposited, the anode counter-electrode and an outer ring meant to prevent sputtering of the hearth that holds the target.

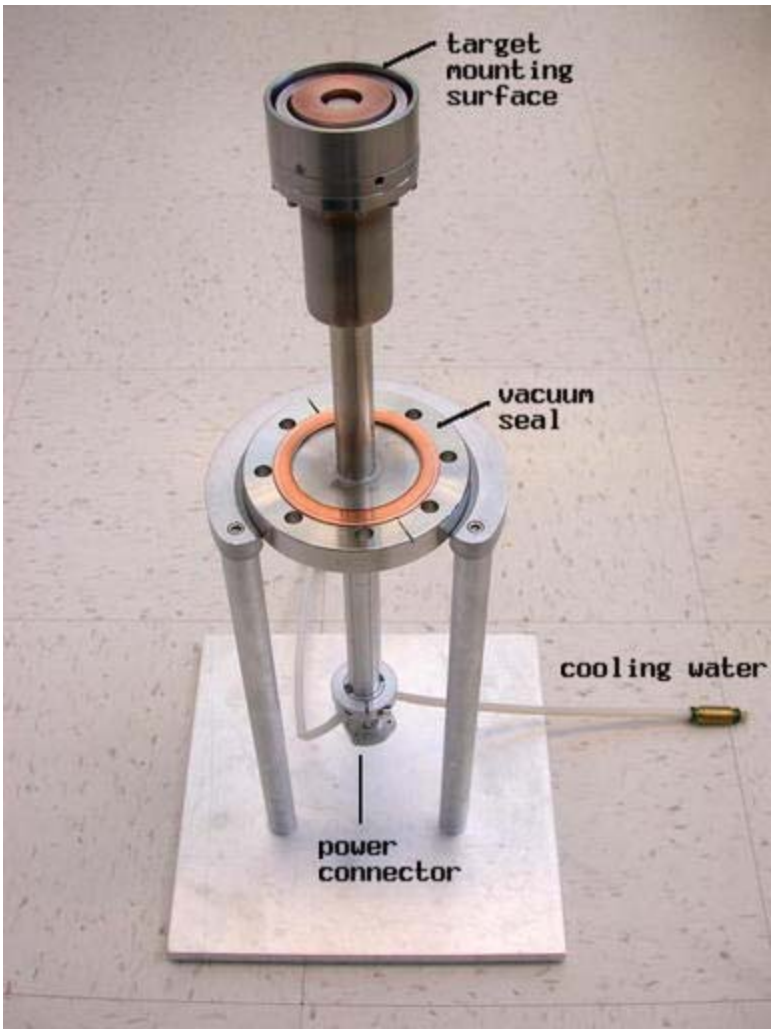
An important advantage of sputter deposition is that even materials with very high melting points are easily sputtered while evaporation of these materials in a resistance evaporator or Knudsen cell is problematic or impossible. Sputter deposited films have a composition close to that of the source material. The difference is due to different elements spreading differently because of their different mass (light elements are deflected more easily by the gas) but this difference is constant. Sputtered films typically have a better adhesion on the substrate than evaporated films. A target contains a large amount of material and is maintenance free making the technique suited for ultrahigh vacuum applications. Sputtering sources contain no hot parts (to avoid heating they are typically water cooled) and are compatible with reactive gases such as oxygen. Sputtering can be performed top-down while evaporation must be performed bottom-up. Advanced processes such as epitaxial growth are possible.

Some disadvantages of the sputtering process are that the process is more difficult to combine with a lift-off process for structuring the film. This is because the diffuse transport, characteristic of sputtering, makes a full shadow impossible. Thus, one cannot fully restrict where the atoms go, which can lead to contamination problems. Also, active control for layer-by-layer growth is difficult compared to pulsed laser deposition and inert sputtering gases are built into the growing film as impurities.

Types of sputter deposition

Sputtering sources are usually magnetrons that utilize strong electric and magnetic fields to trap electrons close to the surface of the magnetron, which is known as the target. The electrons follow helical paths around the magnetic field lines undergoing more ionizing collisions with gaseous neutrals near the target surface than would otherwise occur. The sputter gas is inert, typically argon. The extra argon ions created as a result of these collisions leads to a higher deposition rate. It also means that the plasma can be sustained at a lower pressure. The sputtered atoms are neutrally charged and so are unaffected by the magnetic trap. Charge build-up on insulating targets can be avoided with the use of **RF sputtering** where the sign of the anode-cathode bias is varied at a high rate. RF sputtering works well to produce highly insulating oxide films but only with the added expense of RF power supplies and impedance matching networks. Stray magnetic fields leaking from ferromagnetic targets also disturb the sputtering process. Specially designed sputter guns with unusually strong permanent magnets must often be used in compensation.

Ion-beam sputtering



A magnetron sputter gun showing the target-mounting surface, the vacuum feedthrough, the power connector and the water lines. This design uses a disc target as opposed to the ring geometry illustrated above.

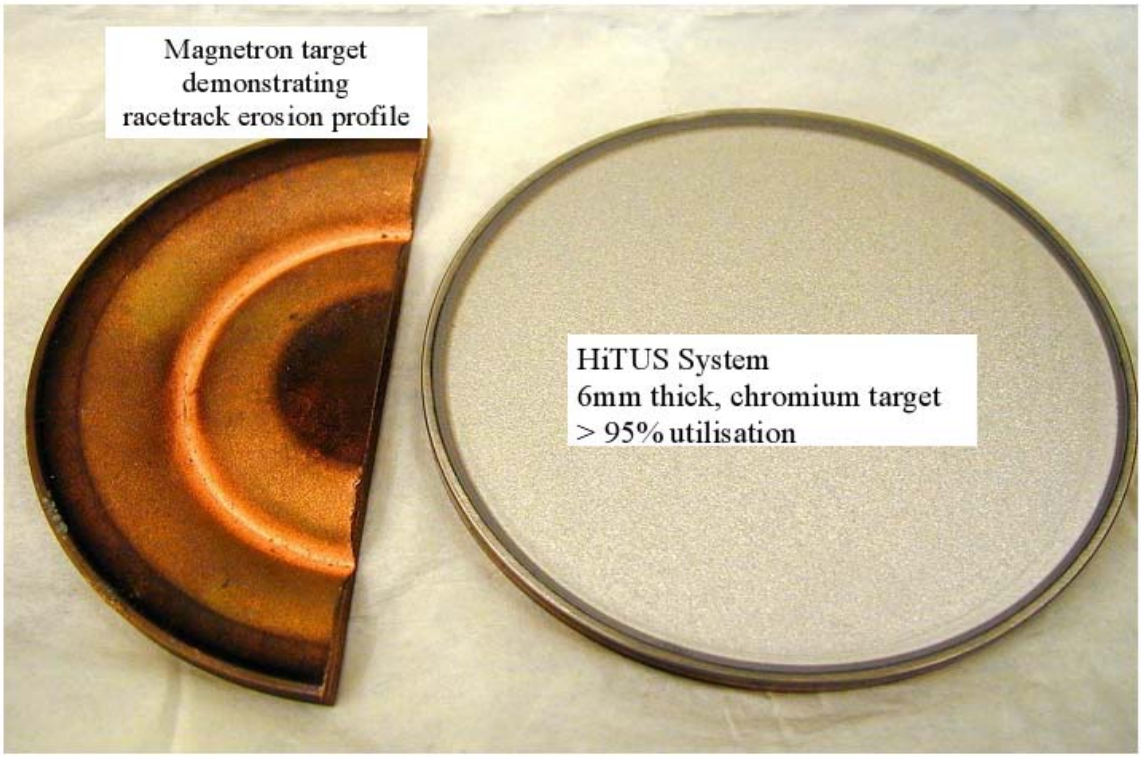
Ion-beam sputtering (IBS) is a method in which the target is external to the ion source. A source can work without any magnetic field like in a hot filament ionization gauge. In a Kaufman source ions are generated by collisions with electrons that are confined by a magnetic field as in a magnetron. They are then accelerated by the electric field emanating from a grid toward a target. As the ions leave the source they are neutralized by electrons from a second external filament. IBS has an advantage in that the energy and flux of ions can be controlled independently. Since the flux that strikes the target is composed of neutral atoms, either insulating or conducting targets can be sputtered. IBS has found application in the manufacture of thin-film heads for disk drives. A pressure gradient between the ion source and the sample chamber is generated by placing the gas inlet at the source and shooting through a tube in into the sample chamber. This saves gas and reduces contamination in UHV applications. The principal drawback of IBS is the large amount of maintenance required to keep the ion source operating.

Reactive sputtering

In reactive sputtering, the deposited film is formed by chemical reaction between the target material and a gas which is introduced into the vacuum chamber. Oxide and nitride films are often fabricated using reactive sputtering. The composition of the film can be controlled by varying the relative pressures of the inert and reactive gases. Film stoichiometry is an important parameter for optimizing functional properties like the stress in SiN_x and the index of refraction of SiO_x. The transparent indium tin oxide conductor that is used in optoelectronics and solar cells is made by reactive sputtering.

Ion-assisted deposition

In ion-assisted deposition (IAD), the substrate is exposed to a secondary ion beam operating at a lower power than the sputter gun. Usually a Kaufman source like that used in IBS supplies the secondary beam. IAD can be used to deposit carbon in diamond-like form on a substrate. Any carbon atoms landing on the substrate which fail to bond properly in the diamond crystal lattice will be knocked off by the secondary beam. NASA used this technique to experiment with depositing diamond films on turbine blades in the 1980s. IAS is used in other important industrial applications such as creating tetrahedral amorphous carbon surface coatings on hard disk platters and hard transition metal nitride coatings on medical implants.



Comparison of target utilization via HiTUS process - 95%

High-target-utilization sputtering

Sputtering may also be performed by remote generation of a high density plasma. The plasma is generated in a side chamber opening into the main process chamber, containing the target and the substrate to be coated. As the plasma is generated remotely, and not from the target itself (as in conventional magnetron sputtering), the ion current to the target is independent of the voltage applied to the target.

High-power impulse magnetron sputtering (HIPIMS)

HIPIMS is a method for physical vapor deposition of thin films which is based on magnetron sputter deposition. HIPIMS utilizes extremely high power densities of the order of kW/cm^2 in short pulses (impulses) of tens of microseconds at low duty cycle of < 10%.

Gas flow sputtering

The process makes use of the hollow cathode effect, by which also hollow cathode lamp are operated. In gas flow sputtering a working gas like argon is led through an opening in a metal subjected to a negative electrical potential . Enhanced plasma densities occur in the hollow cathode, if the pressure in the chamber p and a characteristic dimension L of the hollow cathode obey the Paschen relation 0.5 Pa·m < p·L < 5 Pa·m. This causes a high flux of ions on the surrounding surfaces and a large sputter effect. The hollow-cathode based gas flow sputtering may thus be associated with large deposition rates up to values of a few μm/min .

Structure and morphology

In 1974 J. A. Thornton applied the structure zone model for the description of thin film morphologies to sputter deposition. In a study on metallic layers prepared by DC sputtering , he extended the structure zone concept initially introduced by Movchan and Demchishin for evaporated films . Thornton introduced a further structure zone T, which was observed at low argon pressures and characterized by densely packed fibrous grains. The most important point of this extension was to emphasize the pressure p as a decisive process parameter. In particular, if hyperthermal techniques like sputtering etc. are used for the sublimation of source atoms, the pressure governs via the mean free path the energy distribution with which they impinge on the surface of the growing film. Next to the deposition temperature T_d the chamber pressure or mean free path should thus always be specified when considering a deposition process.

Since sputter deposition belongs to the group of plasma-assisted processes, next to neutral atoms also charged species (like argon ions) hit the surface of the growing film, and this component may exert a large effect. Denoting the fluxes of the arriving ions and atoms by J_i and J_a, it turned out that the magnitude of the J_i/J_a ratio plays a decisive role on the microstructure and morphology obtained in the film . The effect of ion

bombardment may quantitatively be derived from structural parameters like preferred orientation of crystallites or texture and from the state of residual stress. It has been shown recently that textures and residual stresses may arise in gas-flow sputtered Ti layers that compare to those obtained in macroscopic Ti work pieces subjected to a severe plastic deformation by shot peening.

PVD is used in the manufacture of items including semiconductor devices, aluminized PET film for balloons and snack bags, and coated cutting tools for metalworking. Besides PVD tools for fabrication special smaller tools mainly for scientific purposes have been developed. They mainly serve the purpose of extreme thin films like atomic layers and are used mostly for small substrates. A good example are mini e-beam evaporators which can deposit monolayers of virtually all materials with melting points up to 3500°C.

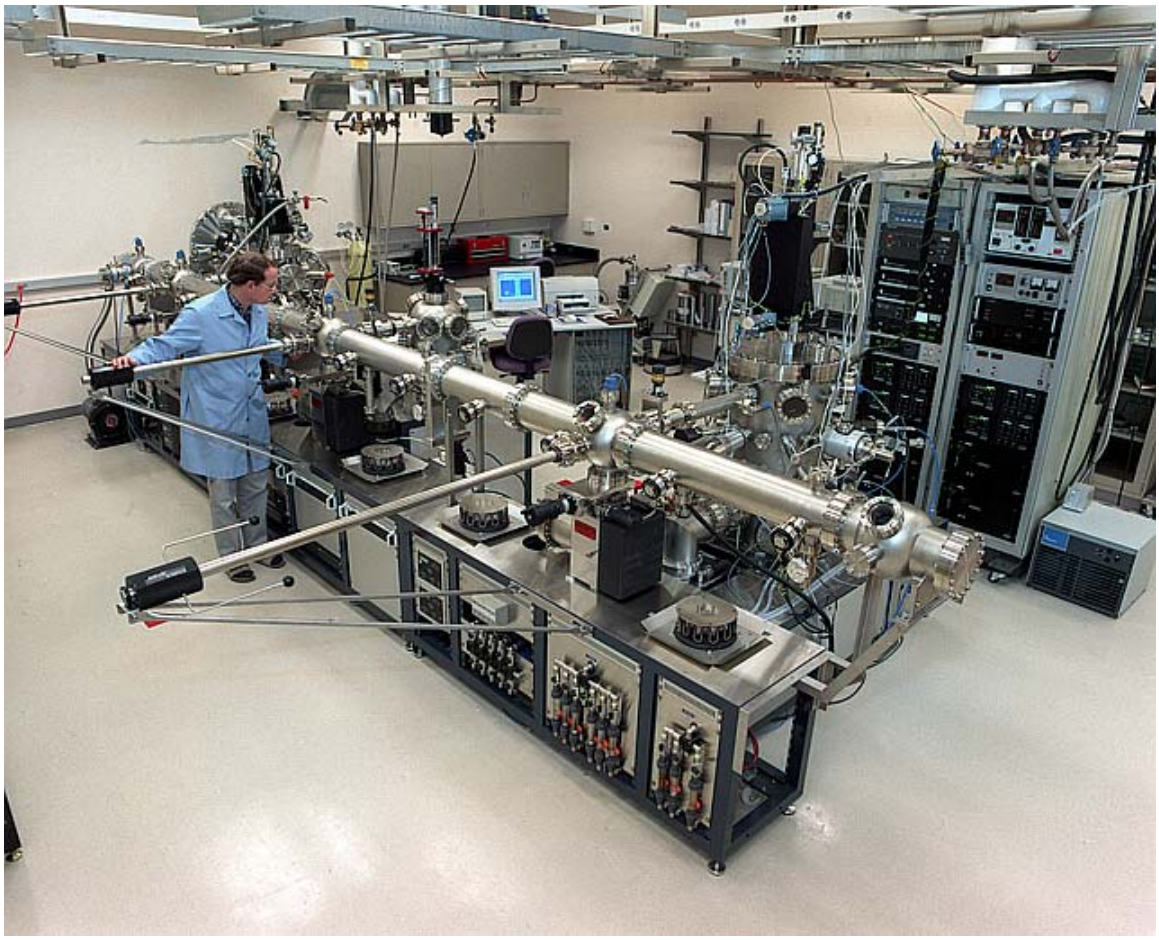
Some of the techniques used to measure the physical properties of PVD coatings are:

- Calo tester: coating thickness test
- Nanoindentation: hardness test for thin-film coatings
- Pin on disc tester: wear and friction coefficient test
- Scratch tester: coating adhesion test

Chapter- 8

Molecular Beam Epitaxy & Chemical-Mechanical Planarization

Molecular beam epitaxy



The Molecular Beam Epitaxy System in the William R. Wiley Environmental Molecular Sciences Laboratory is used to grow and characterize thin crystalline films of oxides and ceramics to understand in detail the chemistry that occurs on oxides and ceramic surfaces.

Molecular beam epitaxy (MBE) is one of several methods of depositing single crystals. It was invented in the late 1960s at Bell Telephone Laboratories by J. R. Arthur and Alfred Y. Cho.

Method

Molecular beam epitaxy takes place in high vacuum or ultra high vacuum (10^{-8} Pa). The most important aspect of MBE is the slow deposition rate (typically less than 1000 nm per hour), which allows the films to grow epitaxially. The slow deposition rates require proportionally better vacuum to achieve the same impurity levels as other deposition techniques.

In solid-source MBE, ultra-pure elements such as gallium and arsenic are heated in separate quasi-Knudsen effusion cells until they begin to slowly sublimate. The gaseous elements then condense on the wafer, where they may react with each other. In the example of gallium and arsenic, single-crystal gallium arsenide is formed. The term "beam" means that evaporated atoms do not interact with each other or vacuum chamber gases until they reach the wafer, due to the long mean free paths of the atoms.

During operation, reflection high energy electron diffraction (RHEED) is often used for monitoring the growth of the crystal layers. A computer controls shutters in front of each furnace, allowing precise control of the thickness of each layer, down to a single layer of atoms. Intricate structures of layers of different materials may be fabricated this way. Such control has allowed the development of structures where the electrons can be confined in space, giving quantum wells or even quantum dots. Such layers are now a critical part of many modern semiconductor devices, including semiconductor lasers and light-emitting diodes.

In systems where the substrate needs to be cooled, the ultra-high vacuum environment within the growth chamber is maintained by a system of cryopumps, and cryopanels, chilled using liquid nitrogen or cold nitrogen gas to a temperature close to 77 Kelvin (−196 degrees Celsius). Cryogenic temperatures act as a sink for impurities in the vacuum, so vacuum levels need to be several orders of magnitude better to deposit films under these conditions. In other systems, the wafers on which the crystals are grown may be mounted on a rotating platter which can be heated to several hundred degrees Celsius during operation.

Molecular beam epitaxy is also used for the deposition of some types of organic semiconductors. In this case, molecules, rather than atoms, are evaporated and deposited onto the wafer. Other variations include gas-source MBE, which resembles chemical vapor deposition.

ATG instability

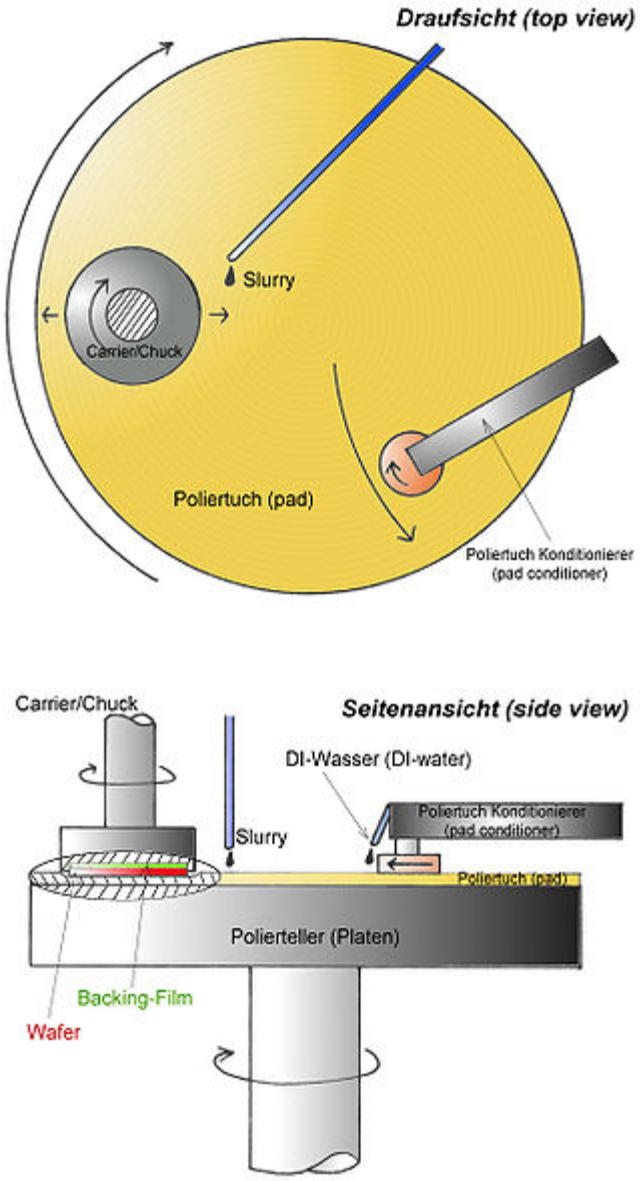
The **ATG** (Asaro-Tiller-Grinfeld) instability, also known as the Grinfeld instability, is an elastic instability often encountered during molecular beam epitaxy. If there is a mismatch between the lattice sizes of the growing film and the supporting crystal, elastic energy will be accumulated in the growing film. At some critical height, the free energy of the film can be lowered if the film breaks into isolated islands, where the tension can be relaxed laterally. The critical height depends on Young's moduli, mismatch size, and surface tensions.

Some applications for this instability have been researched, such as the self-assembly of quantum dot. This community uses the name of Stranski–Krastanov growth for ATG.

Chemical-mechanical planarization

Chemical Mechanical Polishing/Planarization is a process of smoothing surfaces with the combination of chemical and mechanical forces. It can be thought of as a hybrid of chemical etching and free abrasive polishing.

Description



functional principle of CMP

The process uses an abrasive and corrosive chemical slurry (commonly a colloid) in conjunction with a polishing pad and retaining ring, typically of a greater diameter than the wafer. The pad and wafer are pressed together by a dynamic polishing head and held in place by a plastic retaining ring. The dynamic polishing head is rotated with different axes of rotation (i.e., not concentric). This removes material and tends to even out any irregular topography, making the wafer flat or planar. This may be necessary in order to set up the wafer for the formation of additional circuit elements. For example, this might be necessary in order to bring the entire surface within the depth of field of a photolithography system, or to selectively remove material based on its position. Typical depth-of-field requirements are down to Angstrom levels for the latest 65 nm technology.

Working Principles

Typical CMP tools, such as the ones seen on the next page (Figures 1) consist of a rotating and extremely flat platen which is covered by a pad. The wafer that is being polished is mounted upside-down in a carrier/spindle on a backing film. The retaining ring (Figure 1) keeps the wafer in the correct horizontal position. During the process of loading and unloading the wafer onto the tool, the wafer is held by vacuum by the carrier to prevent unwanted particles from building up on the wafer surface. A slurry introduction mechanism deposits the slurry on the pad, this is represented by the slurry supply in Figure 1. Both the platen and the carrier are then rotated and the carrier is kept oscillating as well; this can be better seen in the top view of Figure 2. A downward pressure/down force is applied to the carrier, pushing it against the pad; typically the down force is an average force, but local pressure is needed for the removal mechanisms. Down force depends on the contact area which, in turn, is dependent on the structures of both the wafer and the pad. Typically the pads have a roughness of 50µm; contact is made by asperities (which typically are the high points on the wafer) and, as a result, the contact area is only a fraction of the wafer area. In CMP, the mechanical properties of the wafer itself must be considered too. If the wafer has a slightly bowed structure, the pressure will be greater on the edges than it would on the center, which causes non-uniform polishing. In order to compensate for the wafer bow, pressure can be applied to the wafer's backside which, in turn, will equalize the centre-edge differences. The pads used in the CMP tool should be rigid in order to uniformly polish the wafer surface. However, these rigid pads must be kept in alignment with the wafer at all times. Therefore, real pads are often just stacks of soft and hard materials that conform to wafer topography to some extent. Generally, these pads are made from porous polymeric materials with a pore size between 30-50µm, are consumed in the process, and must be regularly reconditioned. In most cases the pads are very much proprietary, and are usually referred to by their trademark names rather than their chemical or other properties.

Usage in semiconductor fabrication

Before about 1990 CMP was looked on as too "dirty" to be included in high-precision fabrication processes, since abrasion tends to create particles and the abrasives themselves are not without impurities. Since that time, the integrated circuit industry has moved from aluminium to copper conductors. This required the development of an *additive patterning* process, which relies on the unique abilities of CMP to remove material in a planar and uniform fashion and to stop repeatably at the interface between copper and oxide insulating layers. Adoption of this process has made CMP processing much more widespread. In addition to aluminum and copper, CMP processes have been developed for polishing tungsten, silicon dioxide, and (recently) carbon nanotubes.

Limitations of CMP

Nowadays, several limitations of CMP appear during the polishing process. A new technology needs to be optimized. Consequently, this process is immaturity requiring an increase of wafer metrology to attain the desired results. In addition, it was discovered that CMP process has several potential defects, such as, stress cracking, delaminating at weak interfaces, and corrosive attacks from slurry chemicals. The oxide polishing process, which is the oldest and most used in today's industry, has one problem: a lack of end points creates blind spot polishing, making it hard to determine when the desired amount of material has been removed or the desired degree of planarization has been obtained. If the oxide layer has not been sufficiently thinned and/or the desired degree of planarity has not been obtained during this process, then the wafer must be repolished. If the oxide thickness is too thin or too non-uniform, then the wafer must be reworked. Obviously, this method is time-consuming and costly since technicians have to be more attentive while performing this process.

Application

Shallow trench isolation, a process used to fabricate semiconductor devices, is a technique used to enhance the isolation between devices and active areas. Moreover, STI has a higher degree of planarity making it essential in photolithographic applications, depth of focus budget by decreasing minimum line width. To planarize shallow trenches, a common method should be used such as the combination of resist etching-back (REB) and chemical mechanical polishing (CMP). This process comes in a sequence pattern as follows. First, the isolation trench pattern is transferred to the silicon wafer. Oxide is deposited on the wafer in the shape of trenches. A photo mask, composed of silicon nitrate, is patterned on the top of this sacrificial oxide. A second layer is added to the wafer to create a planar surface. After that, the silicon is thermally oxidized, so the oxide grows in regions where there is no SiN_4 and the growth is between 0.5 and 1.0 μm thick. Since the oxidizing species such as water or oxygen are unable to diffuse through the mask, the nitride prevents the oxidation. Next, the etching process is used to etch the wafer and leave a small amount of oxide in the active areas. In the end, CMP is used to polish the SiO_2 overburden with an oxide on the active area.

Chapter- 9

Wafer Testing, Wafer Backgrinding & Die Preparation

Wafer testing

Wafer testing is a step performed during semiconductor device fabrication. During this step, performed before a wafer is sent to die preparation, all individual integrated circuits that are present on the wafer are tested for functional defects by applying special test patterns to them. The wafer testing is performed by a piece of test equipment called a wafer prober. The process of wafer testing can be referred to in several ways: Wafer Sort (WS), Wafer Final Test (WFT), Electronic Die Sort (EDS) and Circuit Probe (CP) are probably the most common.

The wafer prober also exercises any test circuitry on the wafer scribe lines. Some companies get most of their information about device performance from these scribe line test structures.

When all test patterns pass for a specific die, its position is remembered for later use during IC packaging. Sometimes a die has internal spare resources available for repairing (i.e. flash memory IC); if it does not pass some test patterns these spare resources can be used. If redundancy of failed die is not possible the die is considered faulty and is discarded. Non-passing circuits are typically marked with a small dot of ink in the middle of the die, or the information of passing/non-passing is stored in a file, named a wafermap. This map categorizes the passing and non-passing dies by making use of bins. A bin is then defined as a good or bad die. This wafermap is then sent to the die attachment process which then only picks up the passing circuits by selecting the bin number of good dies. The process where no ink dot is used to mark the bad dies is named substrate mapping. When ink dots are used, vision systems on subsequent die handling equipment can disqualify the die by recognizing the ink dot.

In some very specific cases, a die that passes some but not all test patterns can still be used as a product, typically with limited functionality. The most common example of this is a microprocessor for which only one part of the on-die cache memory is functional. In this case, the processor can sometimes still be sold as a lower cost part with a smaller amount of memory and thus lower performance. Additionally when bad dies have been identified, the die from the bad bin can be used by production personnel for assembly line setup.

The contents of all test patterns and the sequence by which they are applied to an integrated circuit are called the test program.

After IC packaging, a packaged chip will be tested again during the IC testing phase, usually with the same or very similar test patterns. For this reason, one might think that wafer testing is an unnecessary, redundant step. In reality this is not usually the case, since the removal of defective dies saves the considerable cost of packaging faulty devices. However, when the production yield is so high that wafer testing is more expensive than the packaging cost of defect devices, the wafer testing step can be skipped altogether and dies will undergo blind assembly.

Wafer backgrinding

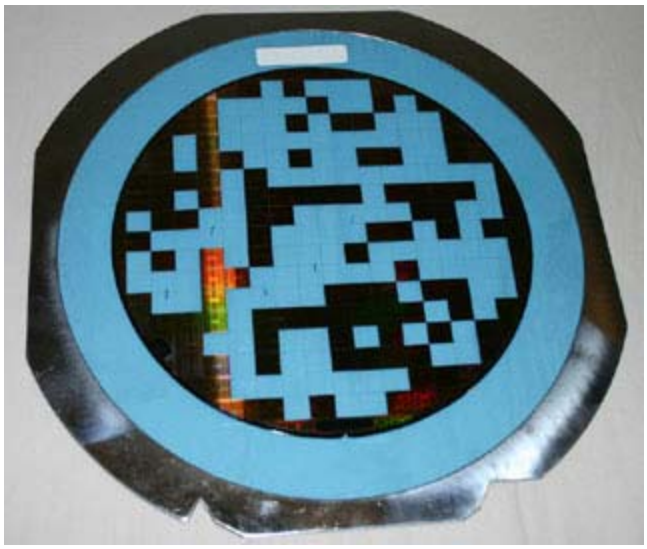
Wafer backgrinding is a semiconductor device fabrication step during which wafer thickness is reduced to allow for stacking and high density packaging of integrated circuits (IC).

ICs are being produced on semiconductor wafers that undergo a multitude of processing steps. The silicon wafers predominantly being used today have diameters of 20 and 30 cm. They are roughly 750 µm thick to ensure a minimum of mechanical stability and to avoid warping during high-temperature processing steps.

Smartcards, USB memory sticks, smartphones, handheld music players, and other ultra compact electronic products would not be feasible in their present form without minimizing the size of their various components along all dimensions. The backside of the wafers are thus ground prior to wafer dicing (where the individual microchips are being singulated). Wafers thinned down to 75 to 50 µm are common today.

The process is also known as 'Backlap' or 'Wafer thinning'.

Die preparation



Wafer glued on blue tape and cut into pieces

Die preparation is a step of semiconductor device fabrication during which a wafer is prepared for IC packaging and IC testing. The process of die preparation typically consists of 2 steps: wafer mounting and wafer dicing.

Wafer mounting

Wafer mounting is a step that is performed during the die preparation of a wafer as part of the process of semiconductor fabrication. During this step, the wafer is mounted on a plastic tape that is attached to a ring. Wafer mounting is performed right before the wafer is cut into separate dies. The adhesive film on which the wafer is mounted ensures that the individual dies remain firmly in place during 'dicing', as the process of cutting the wafer is called.

The picture on the right shows a 300 mm wafer after it was mounted and diced. The blue plastic is the adhesive tape. The wafer is the round disc in the middle. In this case, a large number of dies were already removed.

Semiconductor-die cutting

In the manufacturing of micro-electronic devices, **die cutting**, **dicing** or **singulation** is a process of reducing a wafer containing multiple identical integrated circuits to individual dies each containing one of those circuits.

During this process, a wafer with up to thousands of circuits is cut into rectangular pieces, each called a die. In between those functional parts of the circuits, a thin non-functional spacing is foreseen where a saw can safely cut the wafer without damaging the circuits. This spacing is called *scribe line* or *saw street*. The width of the scribe is very small, typically around 100 μm. A very thin and accurate saw is therefore needed to cut

the wafer into pieces. Usually the dicing is performed with a water-cooled circular saw with diamond-tipped teeth.

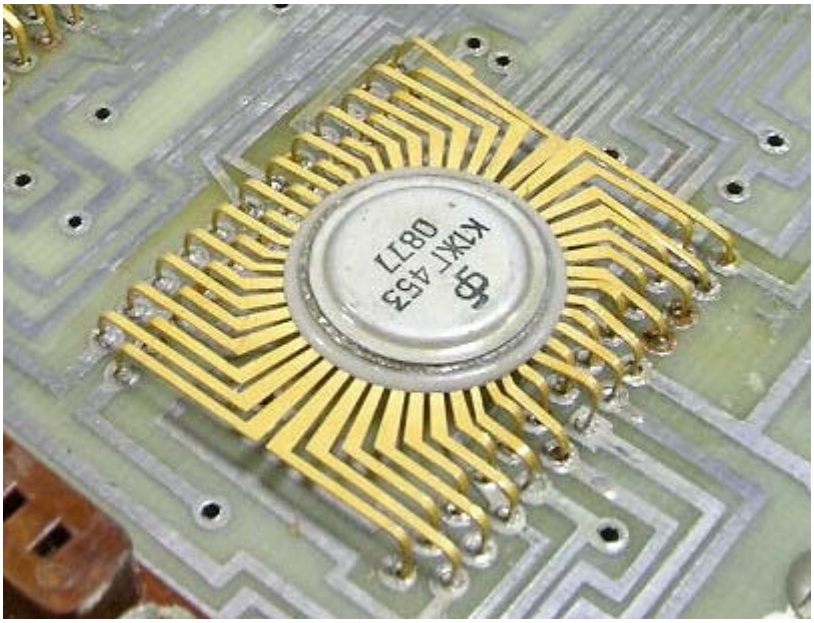
Types of blades

The most common make up of blade used is either a metal or resin bond containing abrasive grit of natural or more commonly synthetic diamond, or borazon in various forms. Alternatively, the bond and grit may be applied as a coating to a metal former. Link to diamond tools.

Chapter- 10

Other Processes in Semiconductor Device Fabrication

Integrated Circuit Packaging



Early USSR made integrated circuit

Integrated circuit packaging is the final stage of semiconductor device fabrication *per se*, followed by IC testing.

In the integrated circuit industry it is called simply **packaging** and sometimes **semiconductor device assembly**, or simply **assembly**. Also, sometimes it is called **encapsulation** or **seal**, by the name of its last step, which might lead to confusion,

because the term packaging generally comprises the steps or the technology of mounting and interconnecting of devices.

Approaches

The earliest integrated circuits were packaged in ceramic flat packs, which continued to be used by the military for their reliability and small size for many years. Commercial circuit packaging quickly moved to the dual in-line package (DIP), first in ceramic and later in plastic. In the 1980s pin counts of VLSI circuits exceeded the practical limit for DIP packaging, leading to pin grid array (PGA) and leadless chip carrier (LCC) packages. Surface mount packaging appeared in the early 1980s and became popular in the late 1980s, using finer lead pitch with leads formed as either gull-wing or J-lead, as exemplified by small-outline integrated circuit — a carrier which occupies an area about 30 – 50% less than an equivalent DIP, with a typical thickness that is 70% less. This package has "gull wing" leads protruding from the two long sides and a lead spacing of 0.050 inches.

Small-outline integrated circuit (SOIC) and Plastic leaded chip carrier (PLCC) packages. In the late 1990s, plastic quad flat pack(PQFP) and thin small-outline packages (TSOP) became the most common for high pin count devices, though PGA packages are still often used for high-end microprocessors. Intel and AMD are currently transitioning from PGA packages on high-end microprocessors to land grid array (LGA) packages.

Ball grid array (BGA) packages have existed since the 1970s. Flip-chip Ball Grid Array packages, which allow for much higher pin count than other package types, were developed in the 1990s. In an FCBGA package the die is mounted upside-down (flipped) and connects to the package balls via a package substrate that is similar to a printed-circuit board rather than by wires. FCBGA packages allow an array of input-output signals (called Area-I/O) to be distributed over the entire die rather than being confined to the die periphery.

Traces out of the die, through the package, and into the printed circuit board have very different electrical properties, compared to on-chip signals. They require special design techniques and need much more electric power than signals confined to the chip itself.

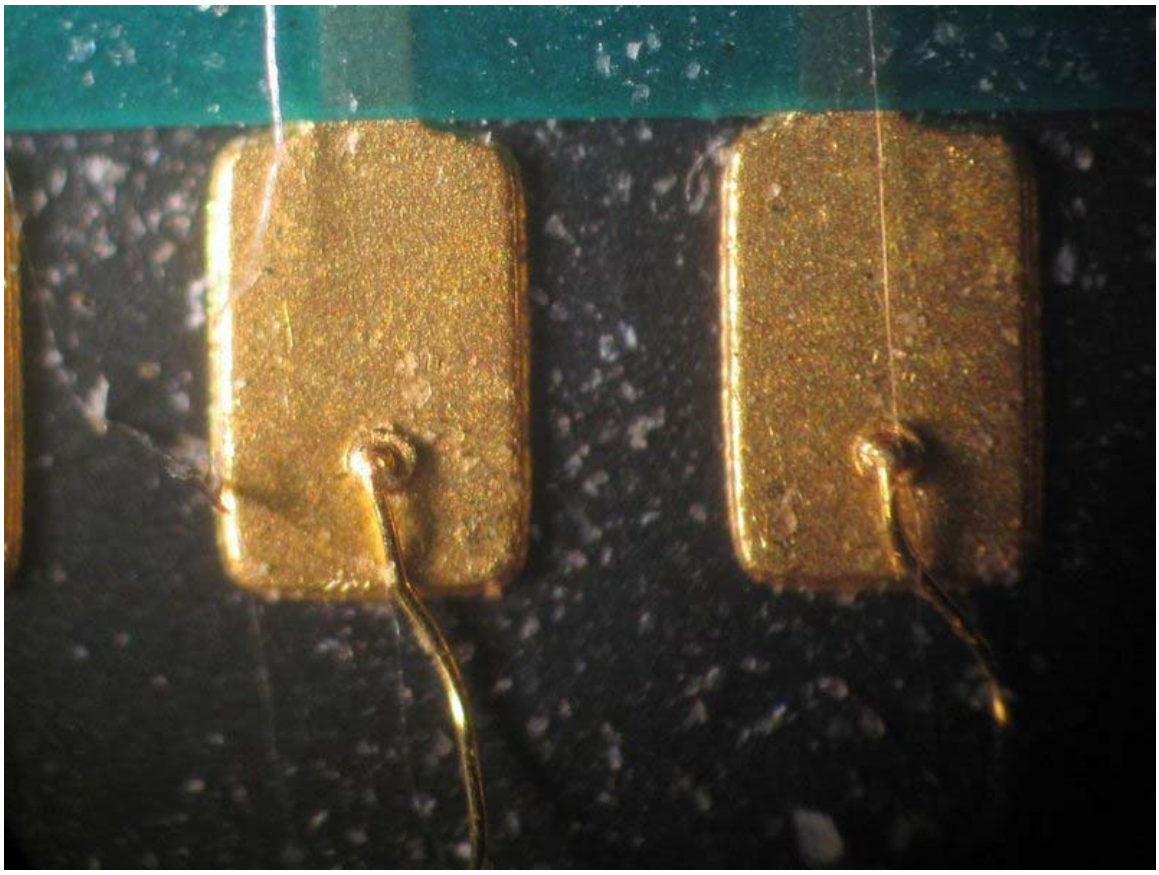
When multiple dies are stacked in one package, it is called SiP, for *System In Package*, or three-dimensional integrated circuit. When multiple dies are combined on a small substrate, often ceramic, it's called an MCM, or Multi-Chip Module. The boundary between a big MCM and a small printed circuit board is sometimes fuzzy.

Die attachment

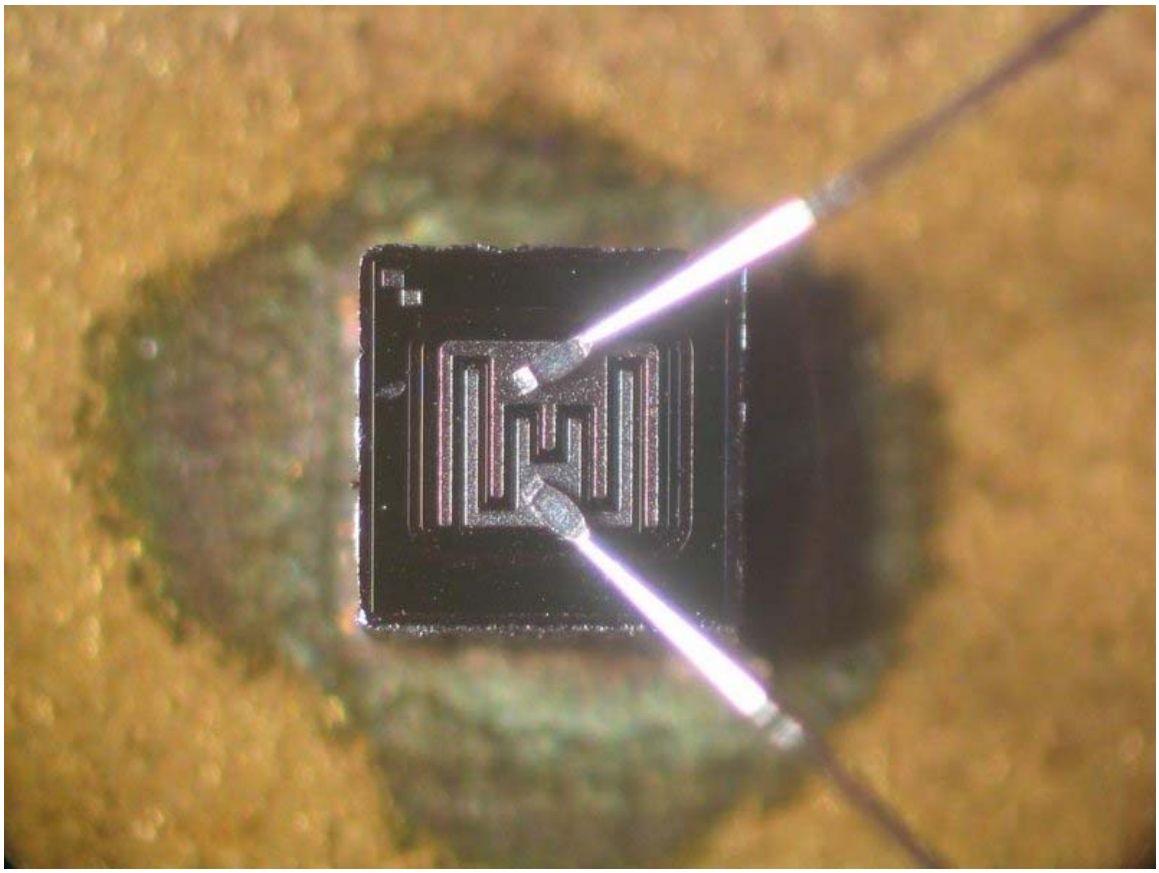
Die attachment is the step during the integrated circuit packaging phase of semiconductor device fabrication during which a die is mounted and fixed to the package or support structure.

For high-powered applications, the die is usually eutectic bonded onto the package, using e.g. gold-tin or gold-silicon solder (for good heat conduction). For low-cost, low-powered applications, the die is often glued directly onto a substrate (such as a printed wiring board) using an epoxy adhesive.

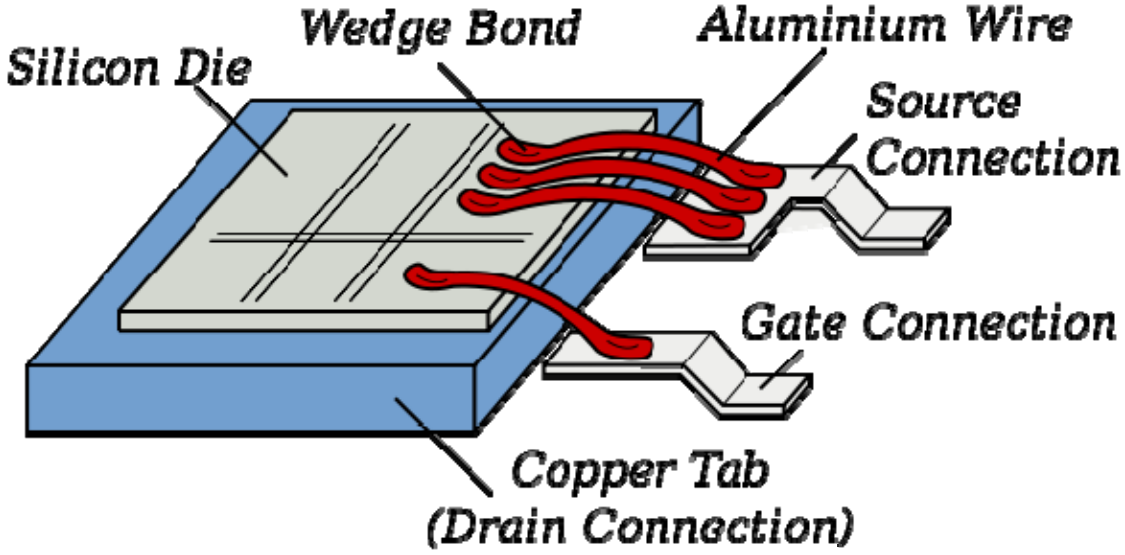
Wire bonding



Gold wire ball-bonded to a gold contact pad



Aluminium wires wedge-bonded to a KSY34 transistor die



The interconnections in a power package are made using thick (250 to 400 μm), wedge-bonded, aluminium wires.

Wire bonding is the primary method of making interconnections between an integrated circuit (IC) and a printed circuit board (PCB) during semiconductor device fabrication.

Although less common, wire bonding can be used to connect an IC to other electronics or to connect from one PCB to another. Wire bonding is generally considered the most cost-effective and flexible interconnect technology, and is used to assemble the vast majority of semiconductor packages.

Bondwires usually consist of one of the following materials:

- Aluminum
- Copper
- Gold

Wire diameters start at 15 µm and can be up to several hundred micrometres for high-powered applications.

Copper wire has become one of the preferred materials for wire bonding interconnects in many semiconductor and microelectronic applications. Copper is used for fine wire ball bonding in sizes up to 0.003 inch (75 micrometres). Copper wire has the ability of being used at smaller diameters providing the same performance as gold without the high material cost.

Copper wire up to 0.010 inch (250 micrometres) can be successfully wedge bonded with the proper set-up parameters. Large diameter copper wire can and does replace aluminum wire where high current carrying capacity is needed or where there are problems with complex geometry. Annealing and process steps used by manufacturers enhance the ability to use large diameter copper wire to wedge bond to silicon without damage occurring to the die.

Copper wire does pose some challenges in that it is harder than both gold and aluminum, so bonding parameters must be kept under tight control. The formation of oxides is inherent with this material, so storage and shelf life are issues that must be considered. Special packaging is required in order to protect copper wire and achieve a longer shelf life.

Pure gold wire doped with controlled amounts of beryllium and other elements is normally used for ball bonding. This process brings together the two materials that are to be bonded using heat, pressure and ultrasonic energy referred to as Thermosonic Bonding. The most common approach in Thermosonic Bonding is to ball-bond to the chip, then stitch-bond to the substrate. Very tight controls during processing enhance looping characteristics and eliminate sagging.

Junction size, bond strength and conductivity requirements typically determine the most suitable wire size for a specific wire bonding application. Typical manufacturers make gold wire in diameters from 0.0005 inch (12.5 micrometres) and larger. Production tolerance on gold wire diameter is +/-3%.

Alloyed aluminum wires are generally preferred to pure aluminum wire except in high-current devices because of greater drawing ease to fine sizes and higher pull-test strengths in finished devices. Pure aluminum and 0.5% magnesium-aluminum are most commonly used in sizes larger than 0.004 inch.

All aluminum systems in semiconductor fabrication eliminate the "purple plague" (brittle gold-aluminum intermetallic compound) sometimes associated with pure gold bonding wire. Aluminum is particularly suitable for ultrasonic bonding.

In order to assure that high quality bonds can be obtained at high production speeds, special controls are used in the manufacture of 1% silicon-aluminum wire. One of the most important characteristics of high grade bonding wire of this type is homogeneity of the alloy system. Homogeneity is given special attention during the manufacturing process. Microscopic checks of the alloy structure of finished lots of 1% silicon-aluminum wire are performed routinely. Processing also is carried out under conditions which yield the ultimate in surface cleanliness and smooth finish and permits entirely snag-free de-reeling.

There are two main classes of wire bonding:

- Ball bonding
- Wedge bonding

Ball bonding usually is restricted to gold and copper wire and usually requires heat. Wedge bonding can use either gold or aluminum wire, with only the gold wire requiring heat.

In either type of wire bonding, the wire is attached at both ends using some combination of heat, pressure, and ultrasonic energy to make a weld.

Thermosonic bonding

Thermosonic bonding is widely used to permanently interconnect metallized silicon integrated circuits (also known as the “chip”) and other components into computers as well as into a myriad of other electronic equipment.

Thermosonic bonding is generally used to form solid-state bonds well below the melting point of the mating metals such as gold wires to gold metallized pads deposited on silicon integrated circuits. It is formed by using a combination of heat, ultrasonic energy and pressure which is generally applied by a bonding tool. Since relatively low bonding parameters are required to form acceptable bonds, the integrity of devices, such as integrated circuits, are further assured throughout their intended lifetimes.

History

Three methods of solid state wire bonding were sequentially developed to improve the permanent interconnections between the silicon integrated circuits (ICs) to the outside circuitry.

Thermocompression bonding: In the mid 1950s, solid state wire bonds were made using heat and pressure and was referred to as Thermocompression bonding . The process was generally limited to bonding pre-cleaned oxide-free gold-to-gold since the desired shearing action at the bonding interface was limited. Orson L. Anderson, a coauthor of the thermocompression bonding paper , later described the importance of interfacial shear in forming reliable solid-state bonds .

Ultrasonic bonding: In the early 1960s commercial ultrasonic wire bonders were introduced which used vibratory energy and pressure to form solid-state bonds without a provision to add heat . The vibratory action enhanced the shearing action at the bonding interface which enhanced the reliability of forming gold-to-gold solid state bonds and also extended the range of mating metals to aluminum and copper.

THERMOSONIC BONDING: In the mid 1960s, **Alexander Coucoulas** reported the first Thermosonic wire bonds using a combination of heat, ultrasonic vibrations and pressure. He used a commercial ultrasonic wire bonder to investigate the attachment of aluminum wires to tantalum thin films deposited on glass substrates. He observed that the ultrasonic energy and pressures levels needed to sufficiently deform the wire and form the required contact areas significantly increased the incidences of cracks in the glass substrates. A means of heating the bond region was then added to the ultrasonic bonder. The bond region was then heated during the ultrasonic bonding cycle which virtually eliminated the glass failure mode since the required contact area was achieved with lower ultrasonic energy and pressure levels. The enhanced wire deformation during the ultrasonic bonding cycle was attributed to the transition from cold working (or strain hardening of the wire) to hot working where its softness was largely maintained due to the onset of recrystallization. Christian Hagar and George Harman stated that in 1970 Alexander Coucoulas, reported additional work in forming thermosonic-type bonds which he initially called *hot work ultrasonic bonding*. In this case, copper wires were bonded to palladium thin films deposited on aluminum oxide substrates.

Summary

In the 1980s, commercial Thermosonic Bonders began to emerge for wire bonding to metallized silicon integrated circuits, flip chip-IC bonding as well for attaching many other components included in today's electronic packages.

At present, the majority of connections to integrated circuits are made using Thermosonic Bonding . Thermosonic bonding is the process of choice because it employs lower bonding temperatures, forces and dwell times than thermocompression bonding as well as lower vibratory energy levels than Ultrasonic bonding to form the required bond area and, therefore, eliminates damaging the relative fragile silicon integrated circuit “chip”.

Such potential failure modes could be costly whether they occur during the manufacturing stage or later, during an operational failure in the field.

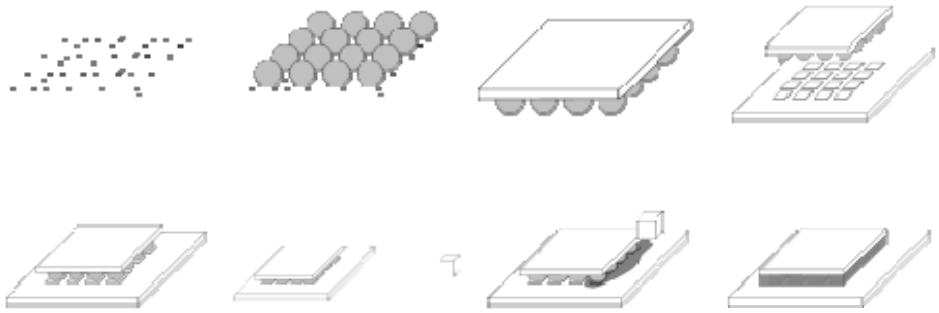
In summary, the use of the **Thermosonic Bonding** process, which was introduced by **Alexander Coucoulas** in the mid 1960's, has ensured the operational integrity of interconnected silicon integrated circuits which are used in computers and other solid-state electronic packages throughout the world.

Flip chip

Flip chip, also known as **Controlled Collapse Chip Connection** or its acronym, **C4**, is a method for interconnecting semiconductor devices, such as IC chips and Microelectromechanical systems (MEMS), to external circuitry with solder bumps that have been deposited onto the chip pads. The solder bumps are deposited on the chip pads on the top side of the wafer during the final wafer processing step. In order to mount the chip to external circuitry (e.g., a circuit board or another chip or wafer), it is flipped over so that its top side faces down, and aligned so that its pads align with matching pads on the external circuit, and then the solder is flowed to complete the interconnect. This is in contrast to wire bonding, in which the chip is mounted upright and wires are used to interconnect the chip pads to external circuitry.

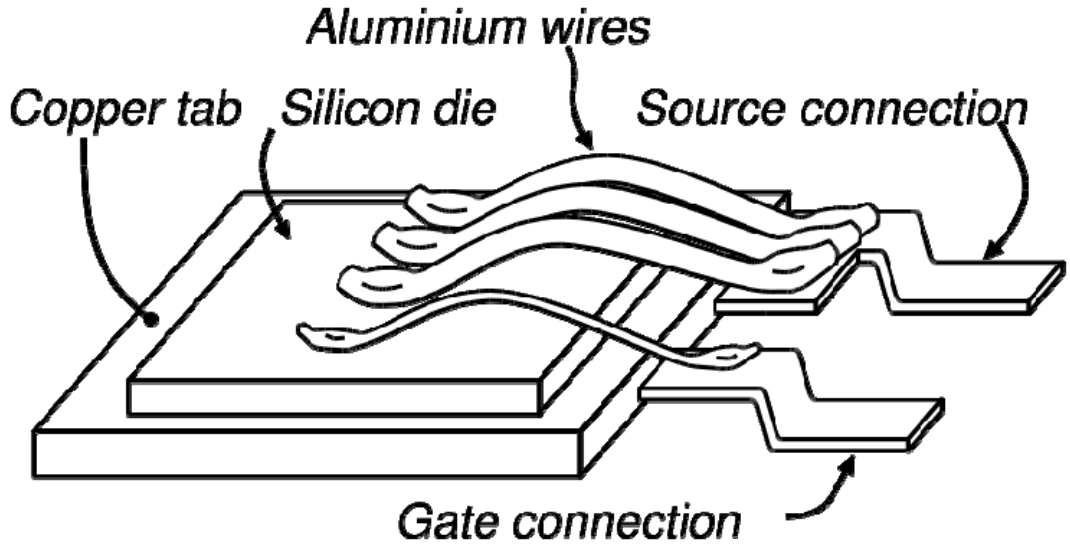
Process steps

- Integrated circuits are created on the wafer
- Pads are metalized on the surface of the chips
- Solder dots are deposited on each of the pads
- Chips are cut
- Chips are flipped and positioned so that the solder balls are facing the connectors on the external circuitry
- Solder balls are then remelted (typically using hot air reflow)
- Mounted chip is “underfilled” using an electrically-insulating adhesive



Comparison of mounting technologies

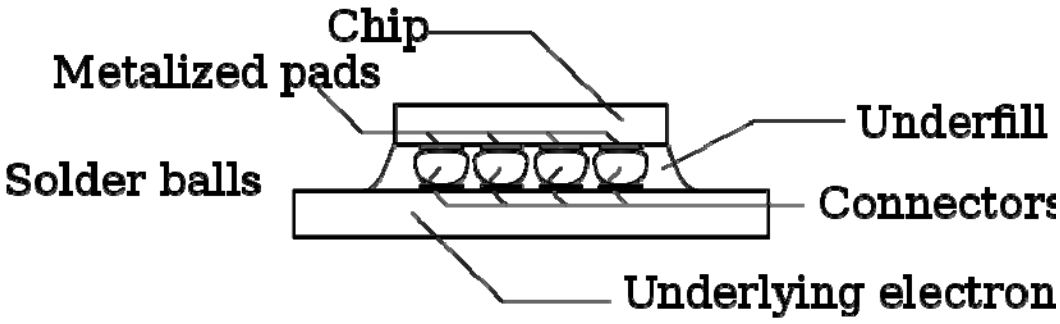
Wire bonding/Thermosonic bonding



The interconnections in a power package are made using thick aluminium wires (250 to 400 μm) wedge-bonded

In typical semiconductor fabration systems chips are built up in large numbers on a single large wafer of semiconductor material, typically silicon. The individual chips are patterned with small pads of metal near their edges that serve as the connections to an eventual mechanical carrier. The chips are then cut out of the wafer and attached to their carriers, typically via wire bonding such as Thermosonic Bonding. These wires eventually lead to pins on the outside of the carriers, which are attached to the rest of the circuitry making up the electronic system.

Flip chip



Side-view schematic of a typical flip chip mounting

Processing a flip chip is similar to conventional IC fabrication, with a few additional steps. Near the end of the manufacturing process, the attachment pads are metalized to make them more receptive to solder. This typically consists of several treatments. A small dot of solder is then deposited on each metalized pad. The chips are then cut out of the wafer as normal.

Recently, high-speed mounting methods evolved through a cooperation between Reel Service Ltd. and Siemens AG in the development of a high speed mounting tape known as 'MicroTape.' . By adding a tape-and-reel process into the assembly methodology, placement at high speed, typically 20,000 placements per hour are achievable using standard PCB assembly equipment.

To attach the flip chip into a circuit, the chip is inverted to bring the solder dots down onto connectors on the underlying electronics or circuit board. The solder is then re-melted to produce an electrical connection, typically using an ultrasonic or alternatively reflow solder process. This also leaves a small space between the chip's circuitry and the underlying mounting. In most cases an electrically-insulating adhesive is then "underfilled" to provide a stronger mechanical connection, provide a heat bridge, and to ensure the solder joints are not stressed due to differential heating of the chip and the rest of the system.

Advantages

The resulting completed flip chip assembly is much smaller than a traditional carrier-based system; the chip sits directly on the circuit board, and is much smaller than the carrier both in area and height. The short wires greatly reduce inductance, allowing higher-speed signals, and also carry heat better.

Disadvantages

Flip chips have several disadvantages. The lack of a carrier means they are not suitable for easy replacement, or manual installation. They also require very flat surfaces to mount to, which is not always easy to arrange, or sometimes difficult to maintain as the boards heat and cool. Also, the short connections are very stiff, so the thermal expansion of the chip must be matched to the supporting board or the connections can crack.

History

The process was originally introduced commercially by IBM in the 1960s for ICs being used in the mainframe systems. DEC followed IBM's lead, but was unable to achieve the quality they demanded, and eventually gave up on the concept. It was pursued once again in the mid-90s for the Alpha product line, but then abandoned due to the fragmentation of the company and subsequent sale to Compaq. In the 1970s it was taken up by Delco Electronics, and has since become very common in automotive applications.

Alternatives

Since the flip chip's introduction a number of alternatives to the solder bumps have been introduced, including gold balls or molded studs, electrically conductive polymer and the "plated bump" process that *removes* an insulating plating by chemical means. Flip chips have recently gained popularity among manufacturers of cell phones, pagers and other small electronics where the size savings are valuable.

Integrated circuit encapsulation

Integrated circuit encapsulation (**IC encapsulation**, **encapsulation**) is the design and manufacturing of protective packages for integrated circuits.

It is often the last stage of IC packaging (semiconductor package assembly) in semiconductor device fabrication. The integrated circuit die is being encapsulated with ceramic, plastic, or epoxy to prevent physical damage or corrosion.

Sometimes the term "encapsulation" is used synonymously to "packaging".

Plating

Plating is a surface covering in which a metal is deposited on a conductive surface. Plating has been done for hundreds of years, but it is also critical for modern technology. Plating is used to decorate objects, for corrosion inhibition, to improve solderability, to harden, to improve wearability, to reduce friction, to improve paint adhesion, to alter conductivity, for radiation shielding, and for other purposes. Jewelry typically uses plating to give a silver or gold finish. Thin-film deposition has plated objects as small as an atom, therefore plating finds uses in nanotechnology.

There are several plating methods, and many variations. In one method, a solid surface is covered with a metal sheet, and then heat and pressure are applied to fuse them (a version of this is Sheffield plate). Other plating techniques include vapor deposition under vacuum and sputter deposition. Recently, plating often refers to using liquids. Metallizing refers to coating metal on non-metallic objects.

Electroplating

In electroplating, an ionic metal is supplied with electrons to form a non-ionic coating on a substrate. A common system involves a chemical solution with the ionic form of the metal, an anode (positively charged) which may consist of the metal being plated (a soluble anode) or an insoluble anode (usually carbon, platinum, titanium, lead, or steel), and finally, a cathode (negatively charged) where electrons are supplied to produce a film of non-ionic metal.

Electroless plating

Electroless plating, also known as chemical or auto-catalytic plating, is a non-galvanic type of plating method that involves several simultaneous reactions in an aqueous solution, which occur without the use of external electrical power. The reaction is accomplished when hydrogen is released by a reducing agent, normally sodium hypophosphite (Note: the hydrogen leaves as a hydride ion), and oxidized thus producing a negative charge on the surface of the part. The most common electroless plating method is electroless nickel plating, although silver, gold and copper layers can also be applied in this manner, as in the technique of Angel gilding.

Specific cases

Gold plating

Gold plating is a method of depositing a thin layer of gold on the surface of glass or metal, most often copper or silver.

Gold plating is often used in electronics, to provide a corrosion-resistant electrically conductive layer on copper, typically in electrical connectors and printed circuit boards. With direct gold-on-copper plating, the copper atoms have the tendency to diffuse through the gold layer, causing tarnishing of its surface and formation of an oxide/sulfide layer. A layer of a suitable barrier metal, usually nickel, has therefore to be deposited on the copper substrate, forming a copper-nickel-gold sandwich.

Metals and glass may also be coated with gold for ornamental purposes, using a number of different processes usually referred to as *gilding*.

Silver plating

A silver-plated alto saxophone

For less demanding applications in electronics, silver is often used as a cheaper replacement for gold. (Although silver is a better conductor than gold it does oxidize and so gold is better for contacts. However, variable capacitors are considered of the highest quality when they have silver plated plates. In this application there is no make and break contact so gold would not offer any advantage over silver).

Care should be used for parts exposed to high humidity environments. When the silver layer is porous or contains cracks, the underlying copper undergoes rapid galvanic corrosion, flaking off the plating and exposing the copper itself; a process known as red plague.

Historically, silver plate was used to provide a cheaper version of items that might otherwise be made of silver, including cutlery and candlesticks. The earliest kind was Old Sheffield Plate, but in the 19th century new methods of production (including electroplating) were introduced.

Another method that can be used to apply a thin layer of silver to several objects, such as glass, is the Tollen's Test method, which usually is prepared as follows. Using this method the final reaction can occur by placing Tollen's Reagent in a glass and then adding Glucose/Dextrose and shaking the bottle to perform the reaction.

$AgNO_3 + KOH \rightarrow AgOH + KNO_3$
$AgOH + 2NH_3 \rightarrow [Ag(NH_3)_2]^{1+} + [OH]^{1-}$
$[Ag(NH_3)_2]^{1+} + [OH]^{1-} + \text{Aldehyde(Usually Glucose/Dextrose)} \rightarrow Ag + 2NH_3 + H_2O$

Rhodium plating

Rhodium plating is occasionally used on white gold, silver or copper and its alloys. A barrier layer of nickel is usually deposited on silver first, though in this case it is not to prevent migration of silver through rhodium, but to prevent contamination of the rhodium bath with silver and copper, which slightly dissolve in the sulfuric acid usually present in the bath composition.

Chrome plating

Chrome plating is a finishing treatment utilizing the electrolytic deposition of chromium. The most common form of chrome plating is the thin, decorative *bright chrome*, which is typically a 10-µm layer over an underlying nickel plate. When plating on iron or steel, an underlying plating of copper allows the nickel to adhere. The pores (tiny holes) in the nickel and chromium layers also promote corrosion resistance. Bright chrome imparts a mirror-like finish to items such as metal furniture frames and automotive trim. Thicker deposits, up to 1000 µm, are called *hard chrome* and are used in industrial equipment to reduce friction and wear.

The traditional solution used for industrial hard chrome plating is made up of about 250 g/l of CrO_3 and about 2.5 g/l of SO_4^-. In solution, the chrome exists as chromic acid, known as hexavalent chromium. A high current is used, in part to stabilize a thin layer of chromium(+2) at the surface of the plated work. Acid chrome has poor throwing power, fine details or holes are further away and receive less current resulting in poor plating.

Zinc plating

Zinc coatings prevent oxidation of the protected metal by forming a barrier and by acting as a sacrificial anode if this barrier is damaged. Zinc oxide is a fine white dust that (in contrast to iron oxide) does not cause a breakdown of the substrate's surface integrity as it is formed. Indeed the zinc oxide, if undisturbed, can act as a barrier to further oxidation,

in a way similar to the protection afforded to aluminum and stainless steels by their oxide layers. The majority of hardware parts are zinc plated, rather than cadmium plated.

Tin plating

The tin-plating process is used extensively to protect both ferrous and nonferrous surfaces. Tin is a useful metal for the food processing industry since it is non-toxic, ductile and corrosion resistant. The excellent ductility of tin allows a tin coated base metal sheet to be formed into a variety of shapes without damage to the surface tin layer. It provides sacrificial protection for copper, nickel and other non-ferrous metals, but not for steel.

Tin is also widely used in the electronics industry because of its ability to protect the base metal from oxidation thus preserving its solderability. In electronic applications, lead may be added to prevent the growth of metallic "whiskers" in compression stressed deposits, which would otherwise cause electrical shorting.

Alloy plating

In some cases, it is desirable to co-deposit two or more metals resulting in an electroplated alloy deposit. Depending on the alloy system, an electroplated alloy may be solid solution strengthened or precipitation hardened by heat treatment to improve the plating's physical and chemical properties. Nickel-Cobalt is a common electroplated alloy.

Composite plating

Metal matrix composite plating can be manufactured when a substrate is plated in a bath containing a suspension of ceramic particles. Careful selection of the size and composition of the particles can fine-tune the deposit for wear resistance, high temperature performance, or mechanical strength. Tungsten carbide, silicon carbide, chromium carbide, and aluminum oxide (alumina) are commonly used in composite electroplating.

Cadmium plating

Cadmium plating is under scrutiny because of the environmental toxicity of the cadmium metal. However, cadmium plating is still widely used in some applications such as aerospace fasteners and it remains in military and aviation specs. Cadmium plating (or "cad plating") offers a long list of technical advantages such as excellent corrosion resistance even at relatively low thickness and in salt atmospheres, softness and malleability, freedom from sticky and/or bulky corrosion products, galvanic compatibility with aluminum, freedom from stick-slip thus allowing reliable torquing of plated threads, can be dyed to many colors and clear, has good lubricity and solderability, and works well either as a final finish or as a paint base. If environmental concerns matter, in most

aspects cadmium plating can be directly replaced with gold plating as it shares most of the material properties - gold is more expensive and cannot serve as a paint base.

Nickel plating

The chemical reaction for nickel plating is:

At cathode: $Ni \rightarrow Ni^{2+} + 2e^-$

At anode: $H_2PO_2 + H_2O \rightarrow H_2PO_3 + 2H^+$

Compared to cadmium plating, nickel plating offers a shinier and harder finish, but lower corrosion resistance, lubricity, and malleability, resulting in a tendency to crack or flake if the piece is further processed.

Electroless nickel plating

Electroless nickel plating, also known as *enickel* and *NiP*, offers many advantages: uniform layer thickness over most complicated surfaces, direct plating of ferrous metals (steel), superior wear and corrosion resistance to electroplated nickel or chrome. Much of the chrome plating done in aerospace industry can be replaced with electroless nickel plating, again environmental costs, costs of hexavalent chromium waste disposal and notorious tendency of uneven current distribution favor electroless nickel plating.

Electroless nickel plating is self-catalyzing process, the resultant nickel layer is NiP compound, with 7-11% phosphorus content. Properties of the resultant layer hardness and wear resistance are greatly altered with bath composition and deposition temperature, which should be regulated with 1 °C precision, typically at 91 °C.

During bath circulation, any particles in it will become also nickel plated, this effect is used to advantage in processes which deposit plating with particles like silicon carbide (SiC) or polytetrafluoroethylene (PTFE). While superior compared to many other plating processes, it is expensive because the process is complex. Moreover, the process is lengthy even for thin layers. When only corrosion resistance or surface treatment is of concern, very strict bath composition and temperature control is not required and the process is used for plating many tons in one bath at once.

Electroless nickel plating layers are known to provide extreme surface adhesion when plated properly. Electroless nickel plating is non-magnetic and amorphous. Electroless nickel plating layers are not easily solderable, nor do they seize with other metals or another electroless nickel plated workpiece under pressure. This effect benefits electroless nickel plated screws made out of malleable materials like titanium. Electrical resistance is higher compared to pure metal plating.