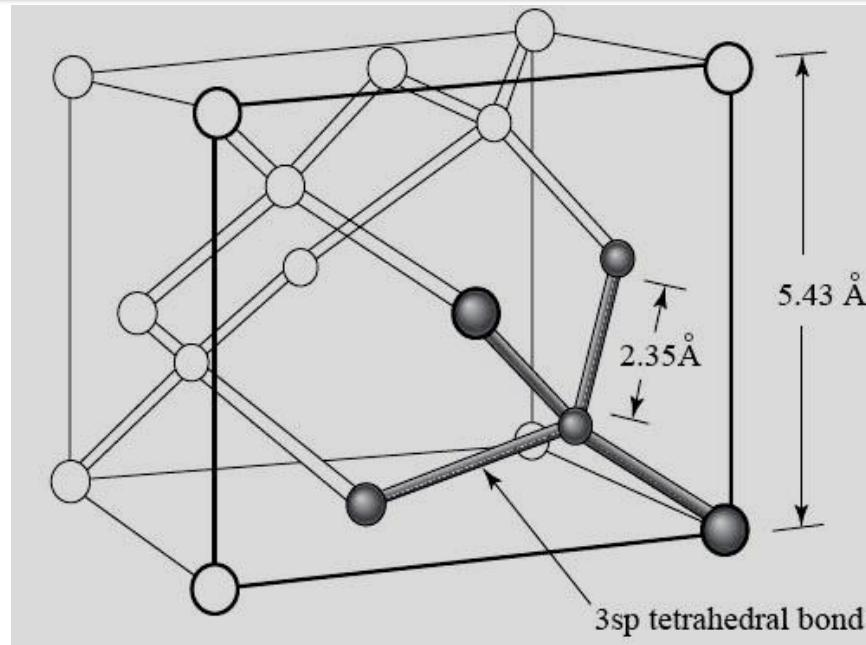


Chapter 1 Introduction and Historical Perspective

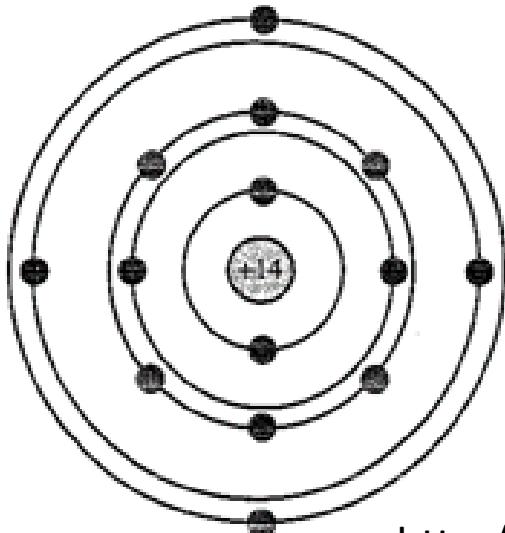
1. Semiconductors.
2. Semiconductor devices, semiconductor technology families.

Read textbook page 13-43

Silicon



“Diamond” structure



$$\text{Si # density: } 8/(5.43\text{\AA})^3 = 5 \times 10^{22}/\text{cm}^3$$

http://en.wikipedia.org/wiki/Diamond_cubic

III ^A	IV ^A	V ^A	VI ^A
5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00
II ^B			
30 Zn 65.39	31 Ga 69.72	32 Ge 72.59	33 As 74.92
48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8
80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0
84 Po 209			

Figure 1-14 Portion of the periodic table relevant to semiconductor materials and doping. Elemental semiconductors are in column IV. Compound semiconductors are combinations of elements from columns III and V (or II and VI).

Carrier concentrations of intrinsic (undoped) Si

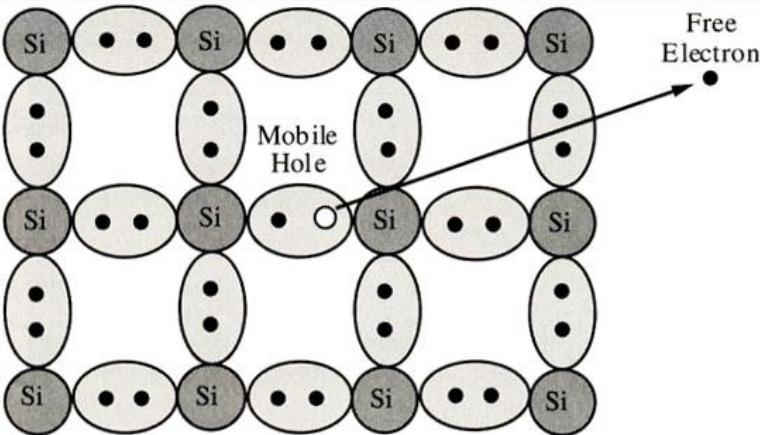
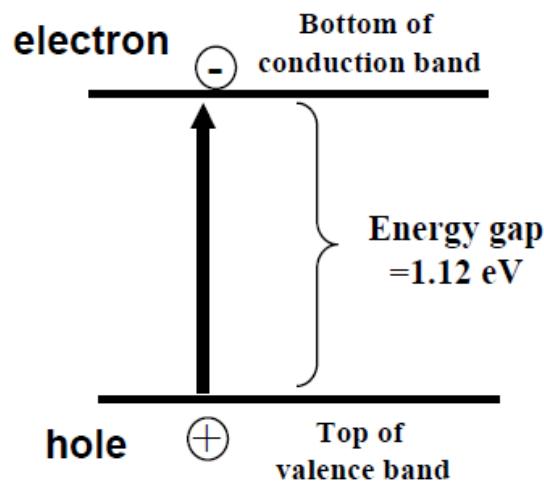
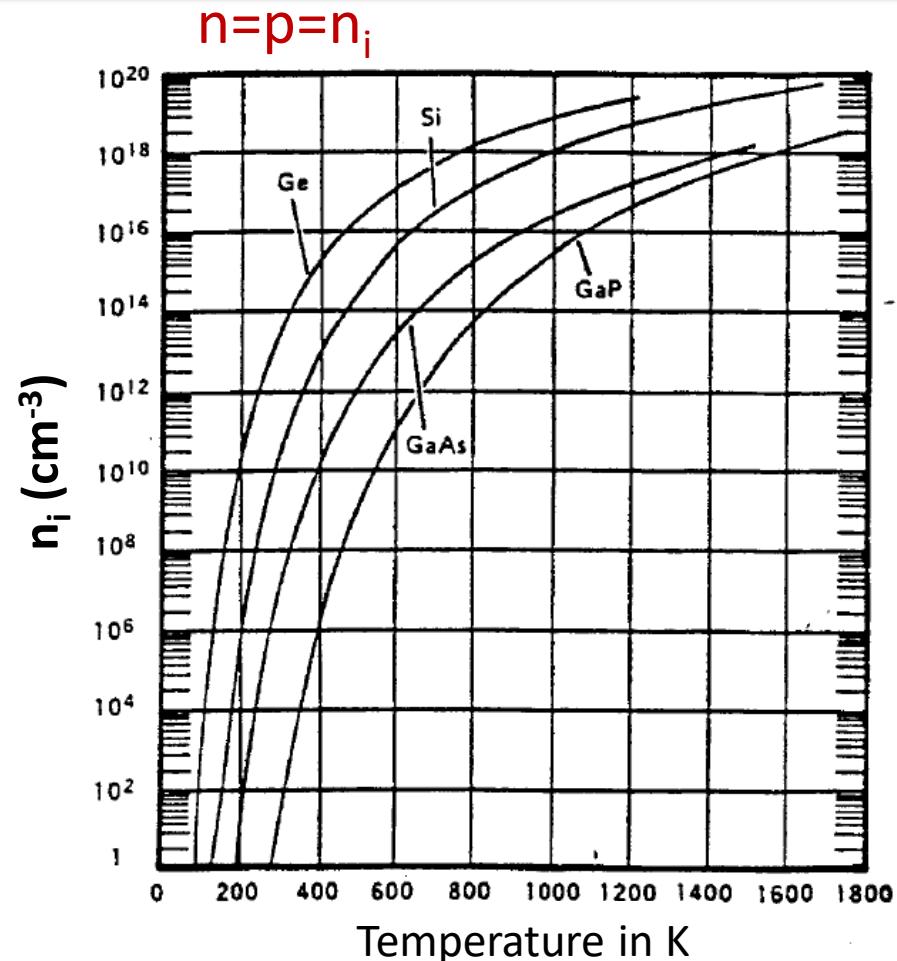


Figure 1-15 Electron (-) and hole (+) pair generation represented by a broken bond in the crystal. Both carriers are mobile and can carry currents in devices.



$1.12\text{eV} \gg kT = 0.026\text{eV}$ for $T=300\text{K}$, so n_i is very low at room temperature.



For Si

$$n_i = 3.9 \times 10^{16} T^{3/2} e^{-\frac{0.605\text{eV}}{kT}} / \text{cm}^3$$

$n_i \approx 1.45 \times 10^{10} \text{ cm}^{-3}$ at room temperature

Doping of silicon

III A	IV A	V A
5 B 10.81	6 C 12.01	7 N 14.01
13 Al 26.98	14 Si 28.09	15 P 30.97
31 Ga 69.72	32 Ge 72.59	33 As 74.92
49 In 114.8	50 Sn 118.7	51 Sb 121.8

By substituting a Si atom with a special impurity atom (Column V for donor, Column III for acceptor), a conduction electron or hole is created.

Semiconductor with both acceptors and donors has 4 kinds of charge carriers

Mobile, contribute to current flow when electric field is applied.

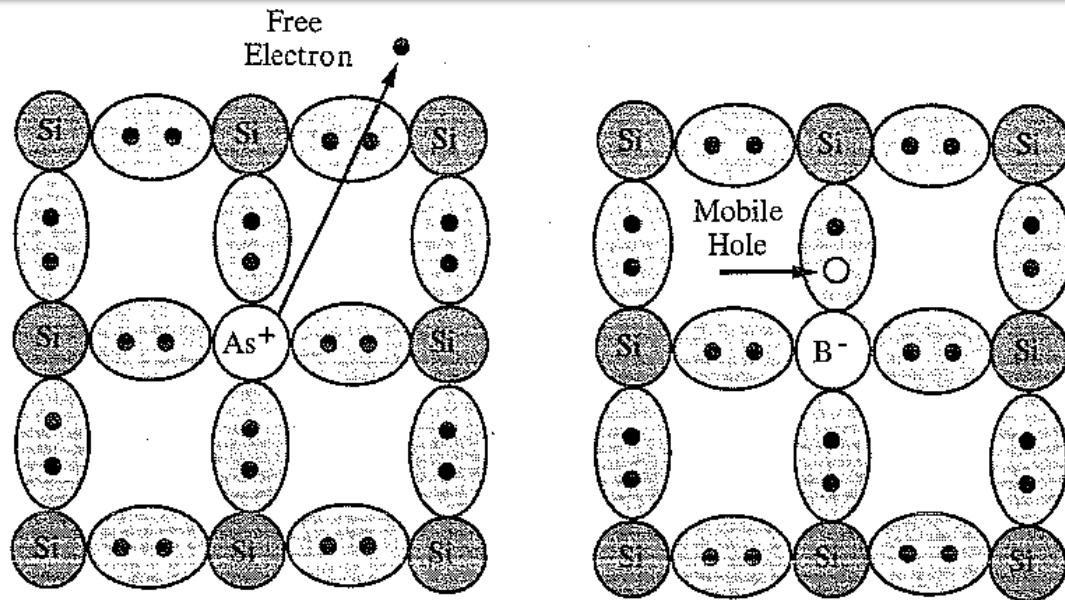


Figure 1-17 Doping of group IV semiconductors using elements from column V (arsenic) or III (boron) of the periodic table.



Hole



Electron



Ionized
Donor

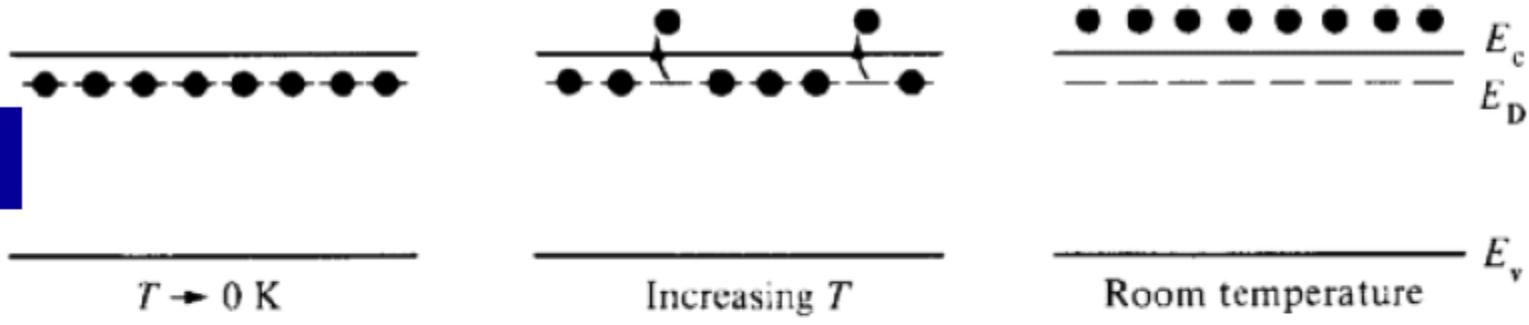


Ionized
Acceptor

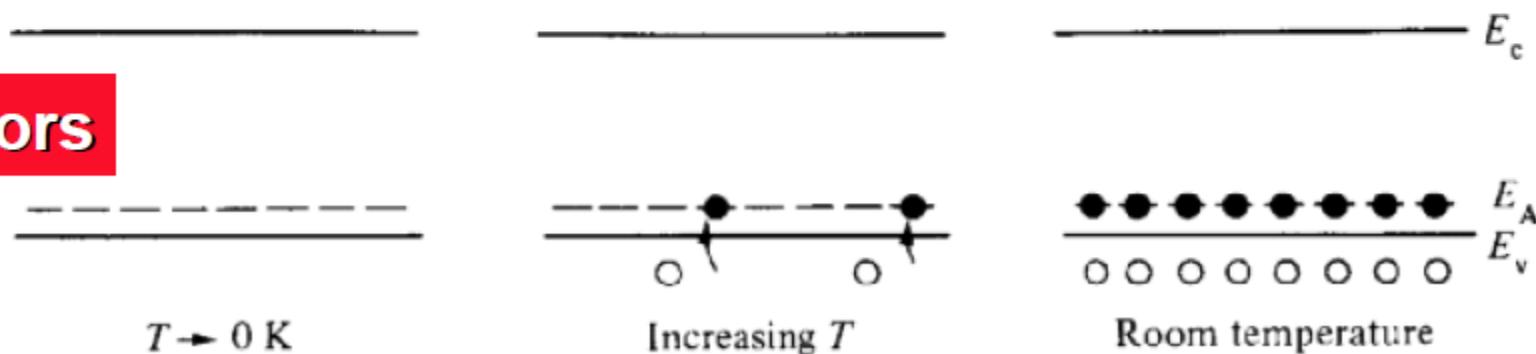
Immobile, DO NOT contribute to current flow with electric field is applied. However, they affect the local electric field

Energy band description of electrons and holes contributed by donors and acceptors

Donors



Acceptors



E_c = bottom of conduction band

E_v = top of valence band

E_D = Donor energy level

E_A = Acceptor energy level

At room temperature, the dopants of interest are essentially fully ionized.

Intrinsic and extrinsic silicon

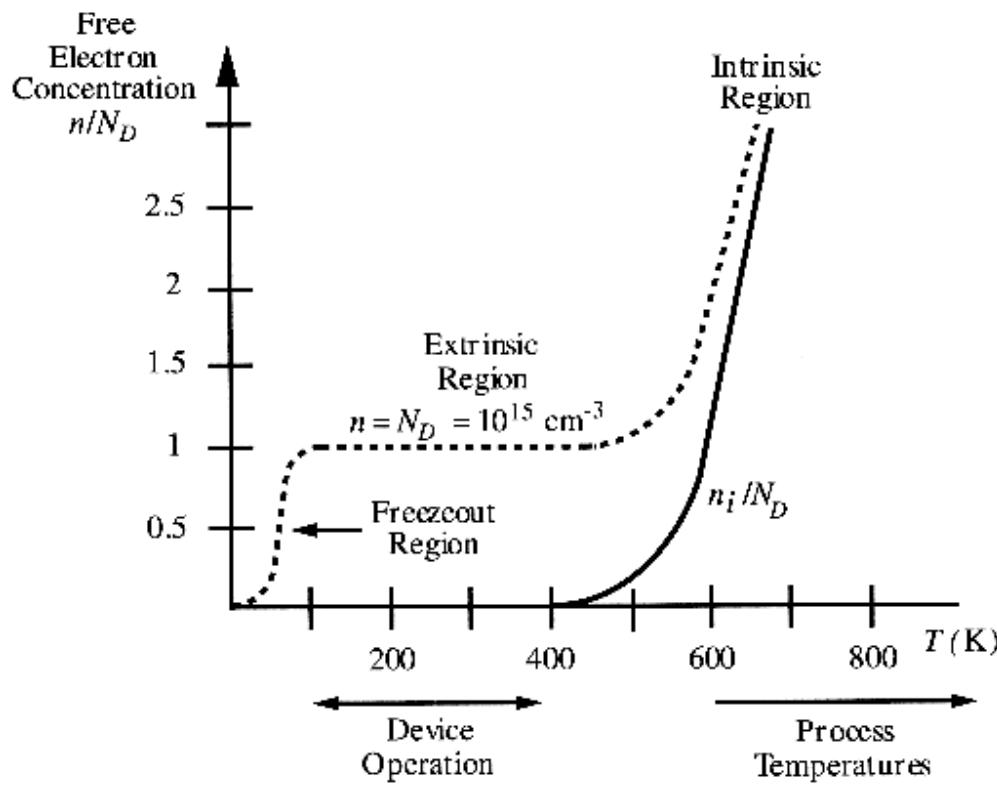


Figure 1-21 Behavior of free carrier concentration versus temperature. Arsenic in silicon is qualitatively illustrated as a specific example ($N_D=10^{15} \text{ cm}^{-3}$). Note that at high temperatures n_i becomes larger than the 10^{15} doping and $n \approx n_i$. Devices are normally operated where $n=N_D$. Fabrication occurs at much higher temperatures where n is often $\approx n_i$.

Intrinsic: un-doped, or doping level lower than n_i .

Extrinsic: carrier density determined/controlled by doping level.

For semiconductor device, it is usually extrinsic at room temperature.

But the semiconductor often becomes intrinsic at device fabrication temperatures (e.g. oxidation is done at $>900^\circ\text{C}$).

Approximate definition of doping levels:

N^- or P^- : N_D or $N_A < 10^{14} \text{ cm}^{-3}$

N^- or P^- : $10^{14} \text{ cm}^{-3} < N_D$ or $N_A < 10^{16} \text{ cm}^{-3}$

N^- or P^- : $10^{16} \text{ cm}^{-3} < N_D$ or $N_A < 10^{18} \text{ cm}^{-3}$

N^+ or P^+ : $10^{18} \text{ cm}^{-3} < N_D$ or $N_A < 10^{20} \text{ cm}^{-3}$

N^{++} or P^{++} : N_D or $N_A > 10^{20} \text{ cm}^{-3}$

Si # density : $5 \times 10^{22} \text{ cm}^{-3}$

Intrinsic Si at RT: $n_i = 1.45 \times 10^{10} \text{ cm}^{-3}$

Electron and hole concentrations for homogeneous semiconductor

n : electron concentration (cm^{-3})

p : hole concentration (cm^{-3})

N_D : donor concentration (cm^{-3})

N_A : acceptor concentration (cm^{-3})

Charge neutrality: $N_D^+ + p = N_A^- + n$

At thermal equilibrium, $np = n_i^2$

(for intrinsic semiconductor $n=p=n_i$, so $np=n_i^2$.

This same relation also holds for extrinsic case)

$$n = \frac{N_D - N_A}{2} + \sqrt{\left(\frac{N_D - N_A}{2}\right)^2 + n_i^2}$$

$$p = \frac{N_A - N_D}{2} + \sqrt{\left(\frac{N_A - N_D}{2}\right)^2 + n_i^2}$$

If $N_D \gg N_A$ (so that $N_D - N_A \gg n_i$):

$$n \approx N_D - N_A \quad \text{and} \quad p \approx \frac{n_i^2}{N_D - N_A}$$

$n \gg p \rightarrow$ material is “n-type”

If $N_A \gg N_D$ (so that $N_A - N_D \gg n_i$):

$$p \approx N_A - N_D \quad \text{and} \quad n \approx \frac{n_i^2}{N_A - N_D}$$

$p \gg n \rightarrow$ material is “p-type”

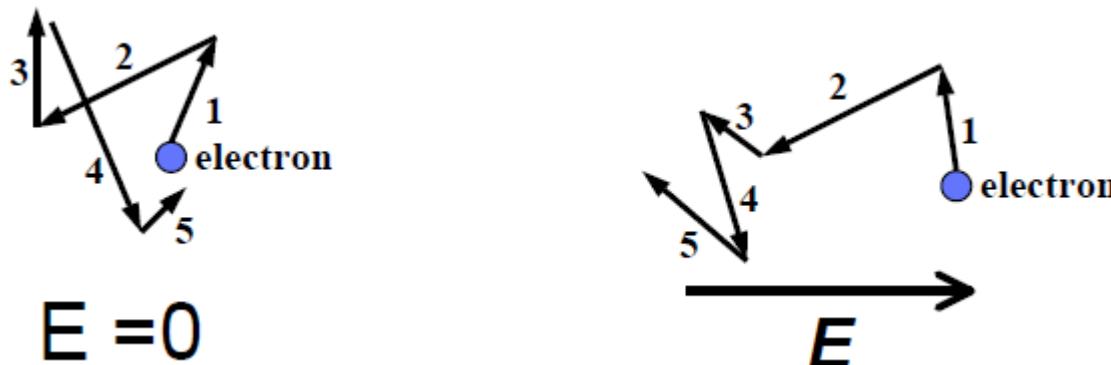
Note: Carrier concentrations depend on NET dopant concentration ($N_D - N_A$)!

Therefore: p-type doping can be realized on n-type substrate if $N_A > N_D$, and vice versa.

Carrier drift and carrier mobility

When an electric field is applied to a semiconductor, mobile carriers will be accelerated by the electrostatic force.

This force superimposes on the random thermal motion of carriers:



E.g. electrons drift in the direction opposite to the E -field.

Average drift velocity = $|v| = \mu E$, μ is carrier mobility.

Electron current density: $J_n = (-q)n v_n = qn\mu_n E$, n is free electron concentration.

Hole current density: $J_p = (+q)p v_p = qp\mu_p E$, p is hole concentration.

Total current density: $J = J_n + J_p = \sigma E$, conductivity $\sigma = (1/\rho) = qn\mu_n + qp\mu_p$

ρ is resistivity. Usually either n or p dominates. E.g. if $n \gg n_i$, then $p = n_i^2/n \ll n_i$.

For Si at RT, with low doping concentration and small fields, maximum values:

$\mu_n = 1500 \text{ cm}^2/\text{V}\cdot\text{sec}$; $\mu_p = 500 \text{ cm}^2/\text{V}\cdot\text{sec} < \mu_n$, so NMOS is faster than PMOS.

Example: dopant compensation

Consider a Si sample doped with $10^{16}/\text{cm}^3$ Boron. What is its electrical resistivity?

Carrier mobility: $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{sec}$.

$$N_A = 10^{16}/\text{cm}^3, N_D = 0 \quad (\textcolor{red}{N_A} \gg \textcolor{blue}{N_D} \rightarrow \text{p-type})$$

$$\rightarrow p \approx 10^{16}/\text{cm}^3 \text{ and } n \approx 10^4/\text{cm}^3$$

$$\begin{aligned} \rho &= \frac{1}{qn\mu_n + qp\mu_p} \cong \frac{1}{qp\mu_p} \\ &= [(1.6 \times 10^{-19})(10^{16})(450)]^{-1} = 1.4 \Omega\text{-cm} \end{aligned}$$

Consider the same Si sample (with $10^{16}/\text{cm}^3$ Boron), doped additionally with $10^{17}/\text{cm}^3$ Arsenic. What is the new resistivity?

Carrier mobility: $\mu_n = 600 \text{ cm}^2/\text{V}\cdot\text{sec}$. (lower μ_n because higher doping reduces mobility)

$$N_A = 10^{16}/\text{cm}^3, N_D = 10^{17}/\text{cm}^3 \quad (\textcolor{blue}{N_D} \gg \textcolor{red}{N_A} \rightarrow \text{n-type})$$

$$\rightarrow n \approx 9 \times 10^{16}/\text{cm}^3 \text{ and } p \approx 1.1 \times 10^3/\text{cm}^3$$

$$\begin{aligned} \rho &= \frac{1}{qn\mu_n + qp\mu_p} \cong \frac{1}{qn\mu_n} \\ &= [(1.6 \times 10^{-19})(9 \times 10^{16})(600)]^{-1} = 0.12 \Omega\text{-cm} \end{aligned}$$

The sample is converted to n-type material by adding more donors than acceptors, and is said to be “compensated”.

Chapter 1 Introduction and Historical Perspective

1. Semiconductors.
2. Semiconductor devices, semiconductor technology families.

p - n junction diode

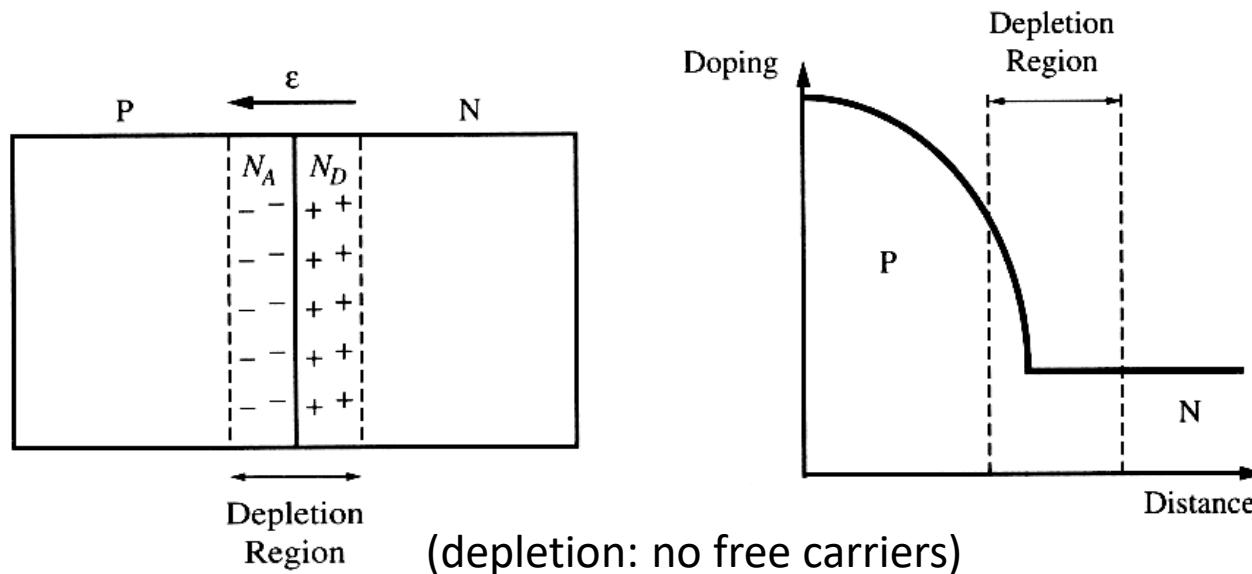


Figure 1-28 Schematic representation of a PN diode on the left. Typical doping profile on the right (here P is doped into the substrate having uniform doping N).

- In equilibrium (no bias), drift current (due to ‘built-in’ electric field ϵ) and diffusion current (due to free carrier concentration gradient) exactly balance, so that no net current flows.
- For forward bias, the applied field partially cancels the built-in field, allowing majority carriers from both sides to diffuse across the junction.
- For reverse bias, the depletion region is widened, only very small leakage current flows.
- The overall I-V relation is simply:

$$I = I_0 \left(\exp \frac{qV}{kT} - 1 \right)$$

MOS transistor

MOS: metal oxide semiconductor.

MOSFET: MOS field effect transistor.

G: gate
S: source
D: drain

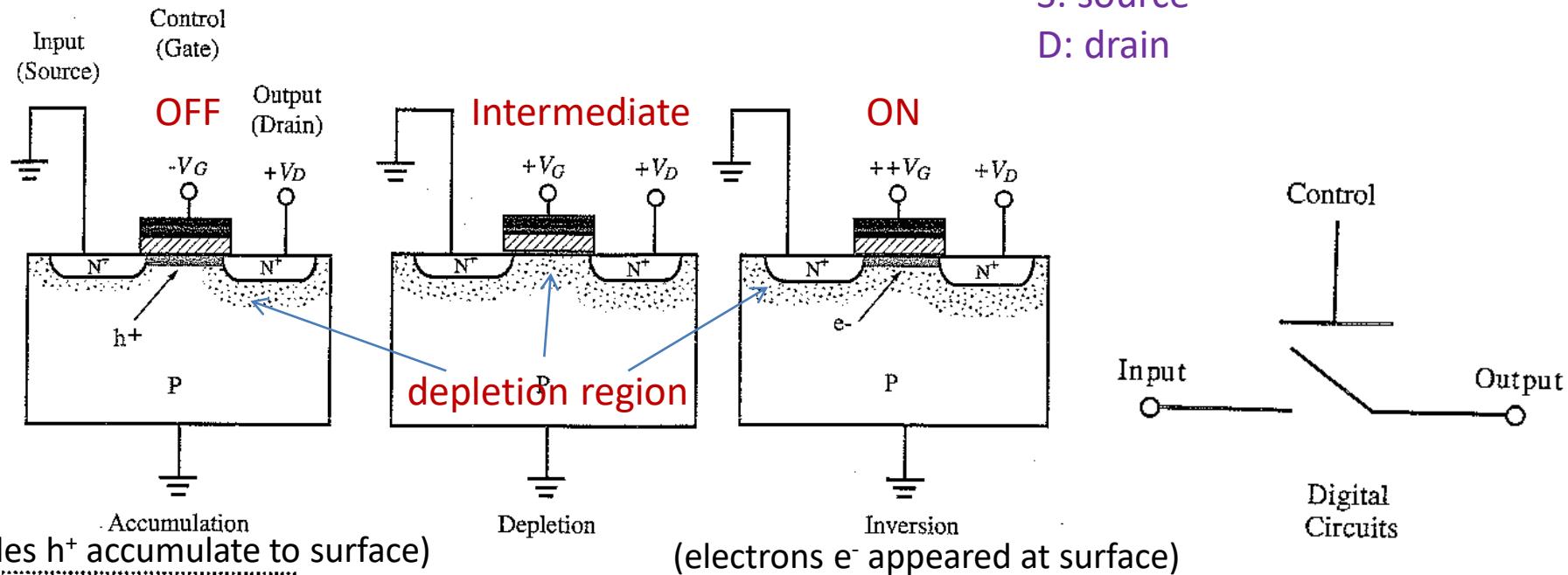


Figure 1-30 Simplified cross section of an MOS transistor biased in its OFF (accumulation), intermediate (depletion) ,and ON (inversion) states. The dotted areas are the depletion regions.

In accumulation, the channel is rich with holes with little free electrons, and the two PN^+ diodes are either zero bias or reverse biased, so there is no/little current between source and drain. The same is true for depletion state where there is no carrier in the channel.

In inversion, the gate voltage is very high which attracts electrons to the very top surface of the channel, so now there is a conduction path of free electrons between source and drain.

Bipolar junction transistor (BJT)

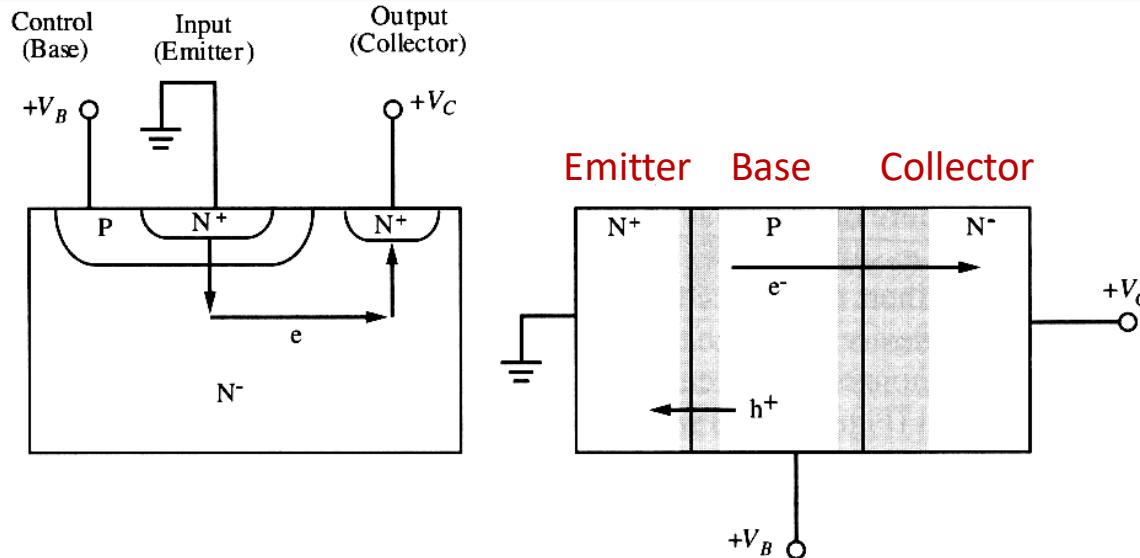


Figure 1-31 Simplified cross section (left) and 1D representation (right) of a bipolar transistor. The shaded areas are the depletion regions. The arrows indicate the path of carrier through the device.

- The key is that the base is very narrow, so it is totally different from two *independent* p-n junctions (one forward, one reverse biased) connected through the base region.
- In operation, the emitter is grounded, a small positive voltage to base, and a large positive voltage applied to the collector.
- A tiny change of V_B leads to a large (exponential) change of I_E that is very close to collector current I_C . (i.e. V_B to control I_C)
- Since most of the current in a BJT flows below the silicon surface, the device is much less sensitive to passivation/protection problems than is the MOS transistor.
- For this reason, BJT was used in the earliest ICs in the 1960s while researchers were trying to understand the stability problems of the Si/SiO₂ interface for MOS transistor.¹³

Semiconductor technology families

1960s, BJT:

BJT: bipolar junction transistor.

Gas phase diffusion for doping.

N⁻ layer grown on P by epitaxy.

Reverse biased p-n junction for device isolation.

6-8 photolithography steps.

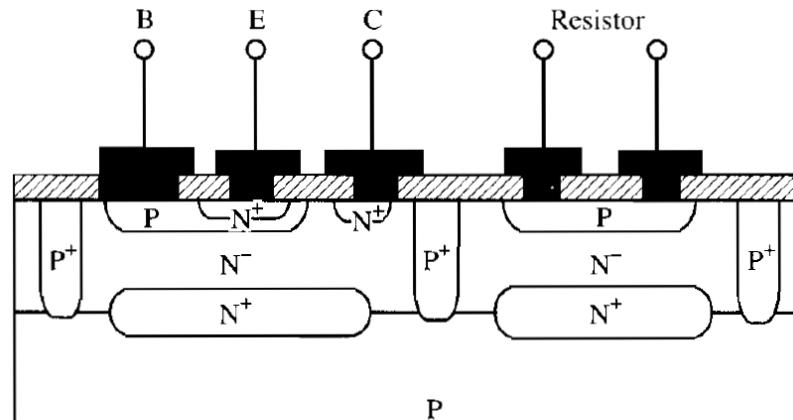


Figure 1-32 Technology typical of the 1960s. Bipolar transistors and resistors were the dominant components.

1970s, E/D NMOS:

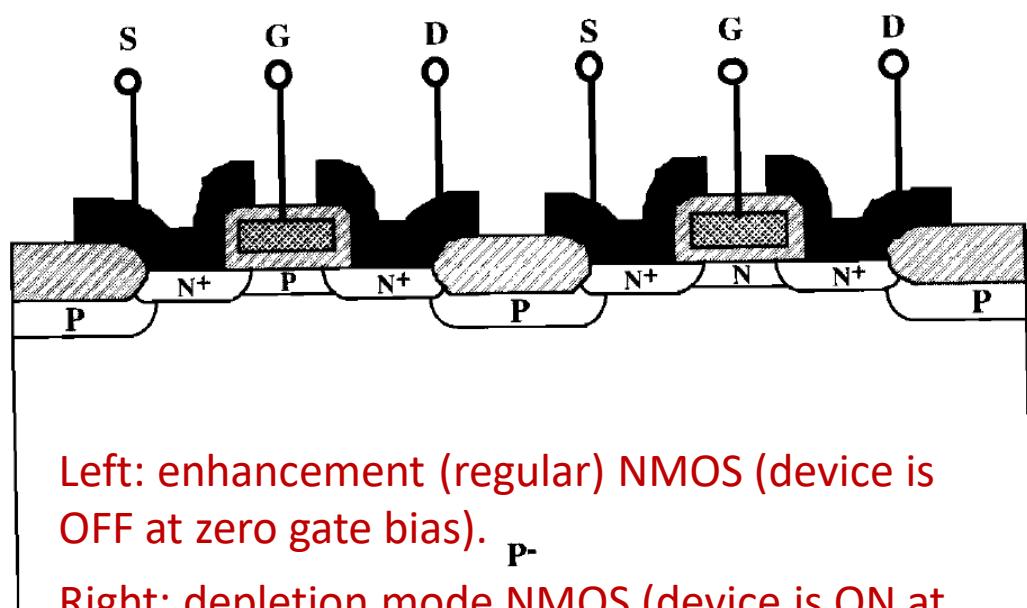
E/D = enhancement/depletion mode

LOCOS (local oxidation) isolation.

NMOS is used since electron mobility is $\sim 3\times$ that of hole mobility.

Depletion NMOS took small area, thus denser circuit.

Again, 6-8 photolithography steps.



Left: enhancement (regular) NMOS (device is OFF at zero gate bias).

P-

Right: depletion mode NMOS (device is ON at zero gate bias).

Semiconductor technology families

1980s, CMOS:

CMOS: complementary (equal number of NMOS and PMOS) MOS.

Low power consumption, low heating.

E.g. the CMOS inverter consumes no DC current in either state (no DC power).

Higher level integration.

12-14 photolithography steps.

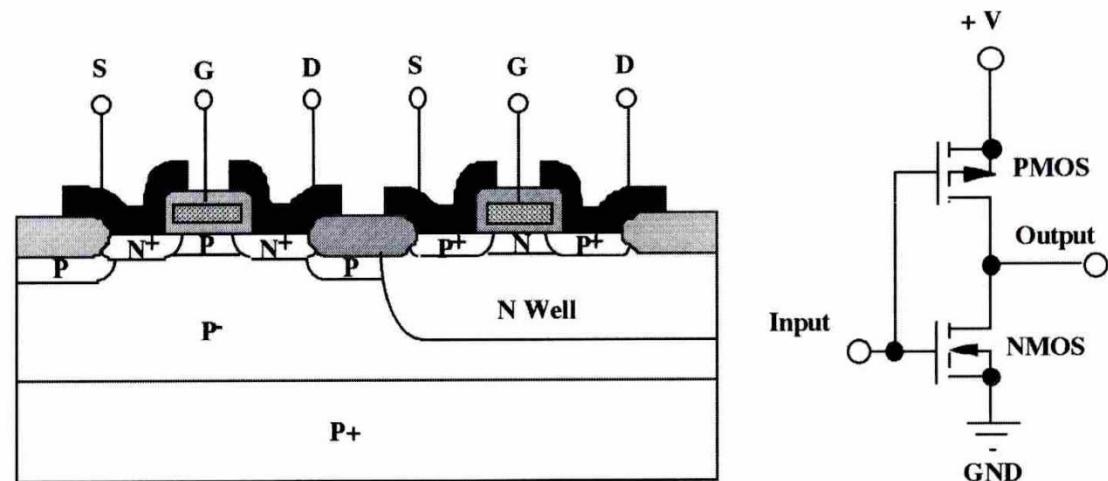


Figure 1-34 Technology typical of the 1980s. CMOS circuits with both NMOS and PMOS devices were dominant.

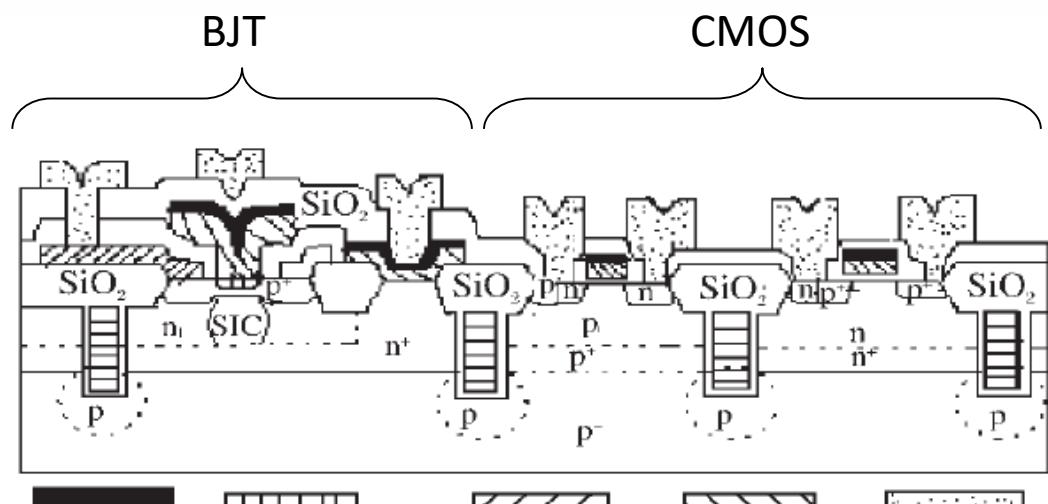
1990s, BiCMOS:

Bipolar and CMOS.

CMOS for highly integrated internal circuit.

BJT for driving circuit.

>20 photolithography steps.



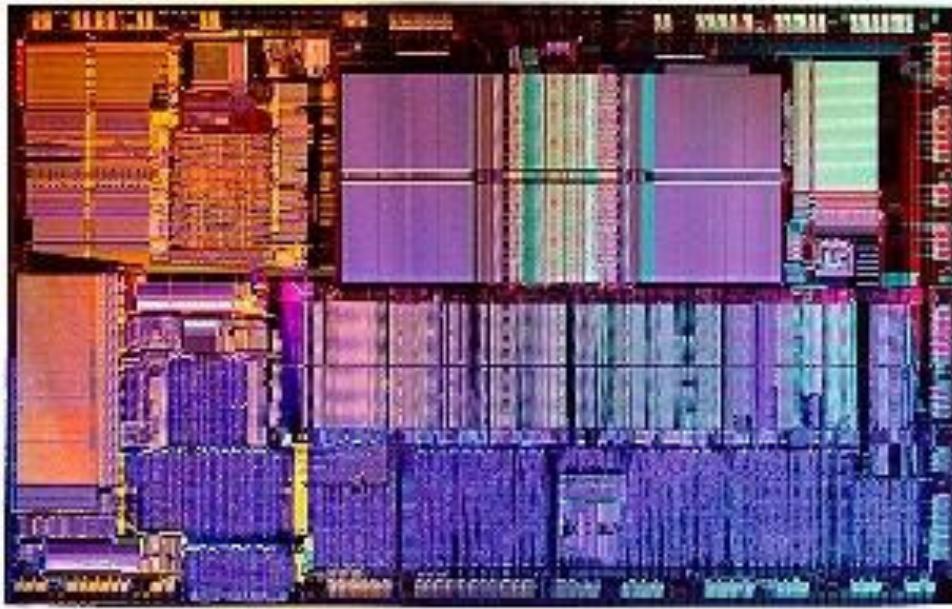
BJT: bipolar junction transistor
poly = poly-crystalline Si.

TiSi Un-doped poly p⁺ poly n⁺ poly metal 15

Chapter 1 Introduction and Historical Perspective: semiconductor and electronic device

1. Introduction.
2. Growth of IC – Moore's law.
3. Some history in IC industry.

Introduction



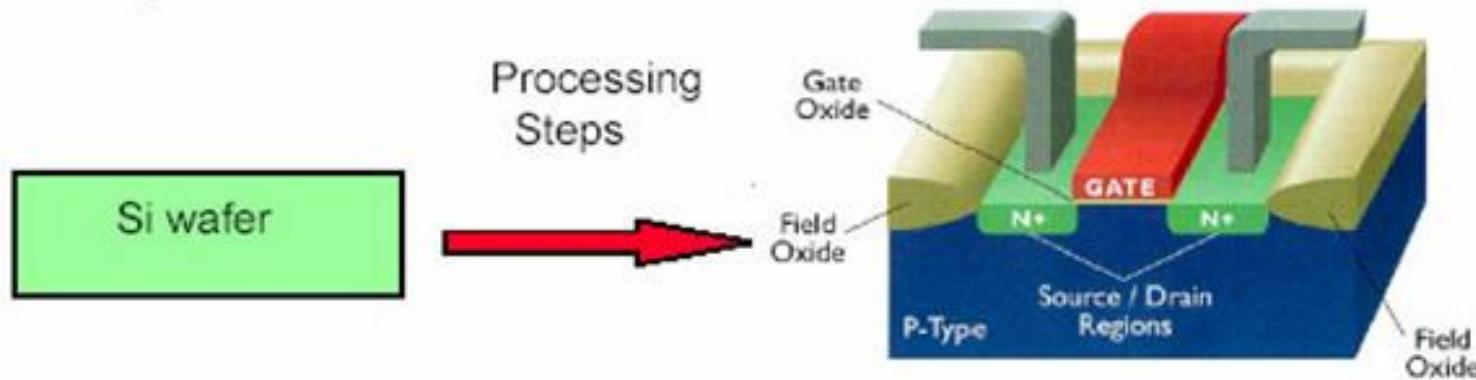
State of art ICs
manufactured in the 1990s.

- This course is basically about silicon chip fabrication, the technologies used to manufacture ICs (CPU, memory – DRAM, flash...).
- However, the same technology is also widely used for applications other than ICs, such as large area displays (LCD), hard disk drive, semiconductor lasers, MEMS (microelectromechanical systems), lab-on-a-chip, solar cell....
- For nano-application, microfabrication is the basis for nanofabrication; with the major differences is that photolithography is used for microfabrication whereas nano-lithography (electron beam lithography...) is used for nanofabrication.
- Therefore, you will find this course very useful even though most of you will not work in the IC industry after graduation.

Basic fabrication components

A sequence of additive and subtractive steps with lateral patterning.

Example: MOSFET



Three components for micro- and nano-fabrication:

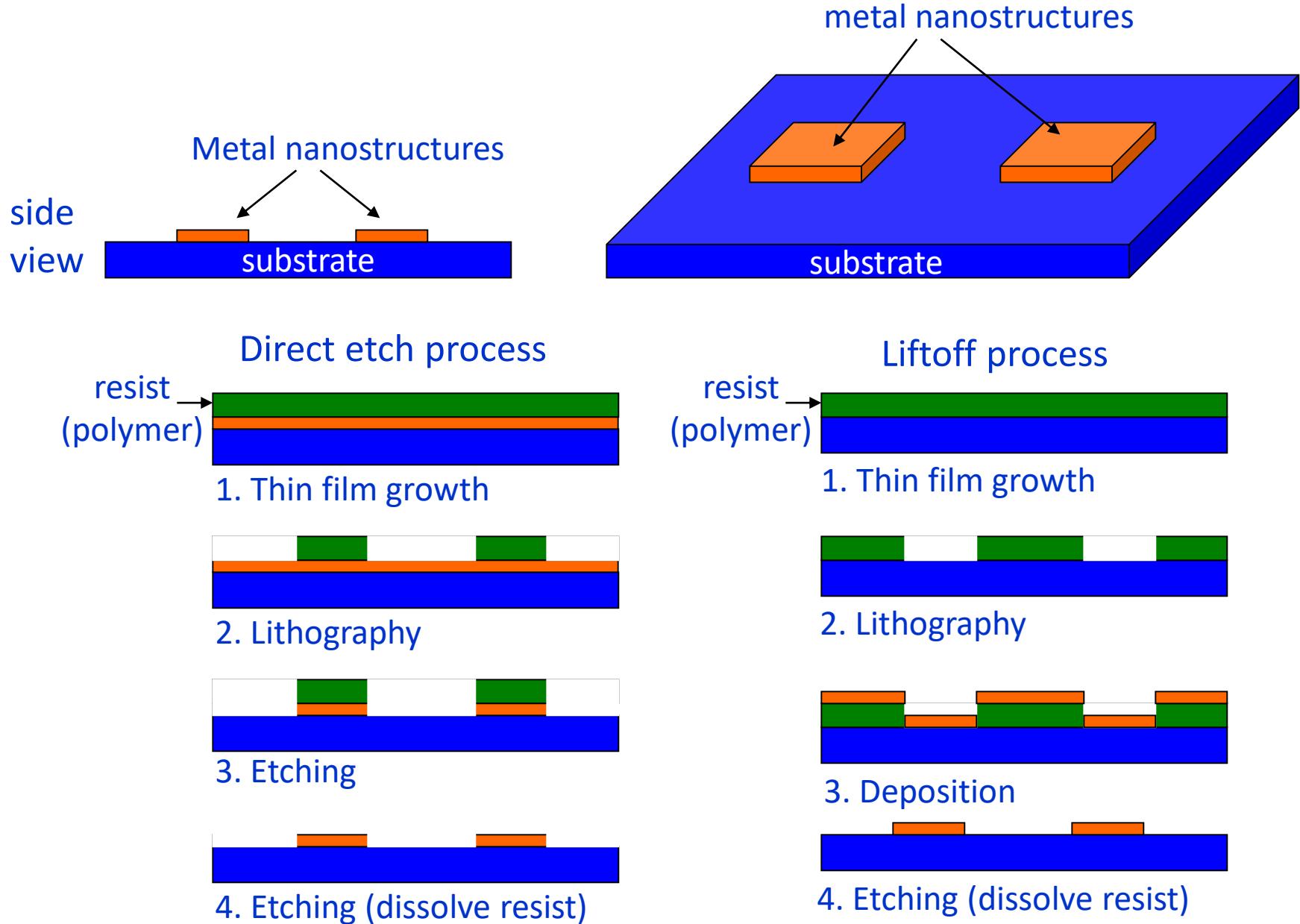
Lithography (lateral patterning): generate pattern in a material *called resist*
photolithography, electron-beam lithography, nanoimprint lithography...

Thin film deposition (additive): spin coating, chemical vapor deposition, molecular beam
epitaxy, sputtering, evaporation, electroplating...

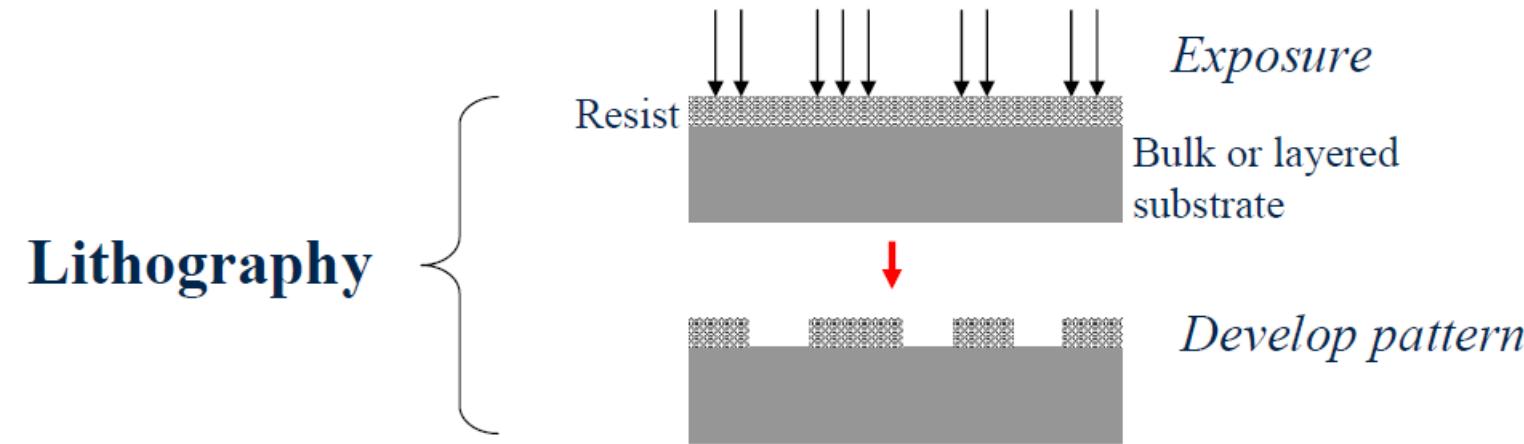
Etching (subtractive): reactive ion etching, ion beam etching, wet chemical etching,
polishing...

Other techniques such as doping (ion implantation) is also important for semiconductor
device.

One fabrication example

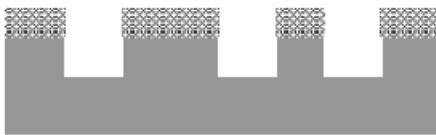


Summary of general fabrication process



Pattern Transfer

Etch



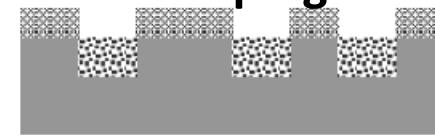
Direct etch

Deposition



Liftoff

Diffusion or doping



Diffusion or doping

Growth



Electroplating...

Chapter 1 Introduction and Historical Perspective: Integrated Circuit (IC)

1. Introduction.
2. Growth of IC – Moore's law.
3. Some history in IC industry.

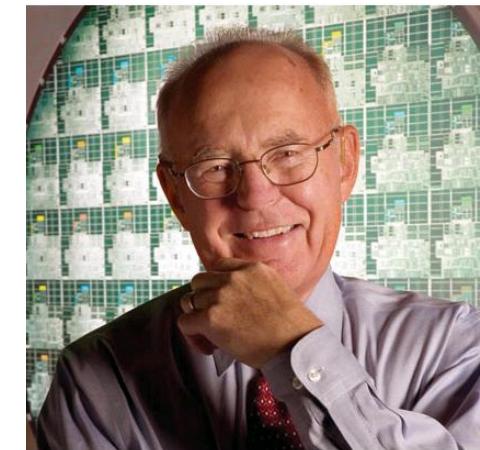
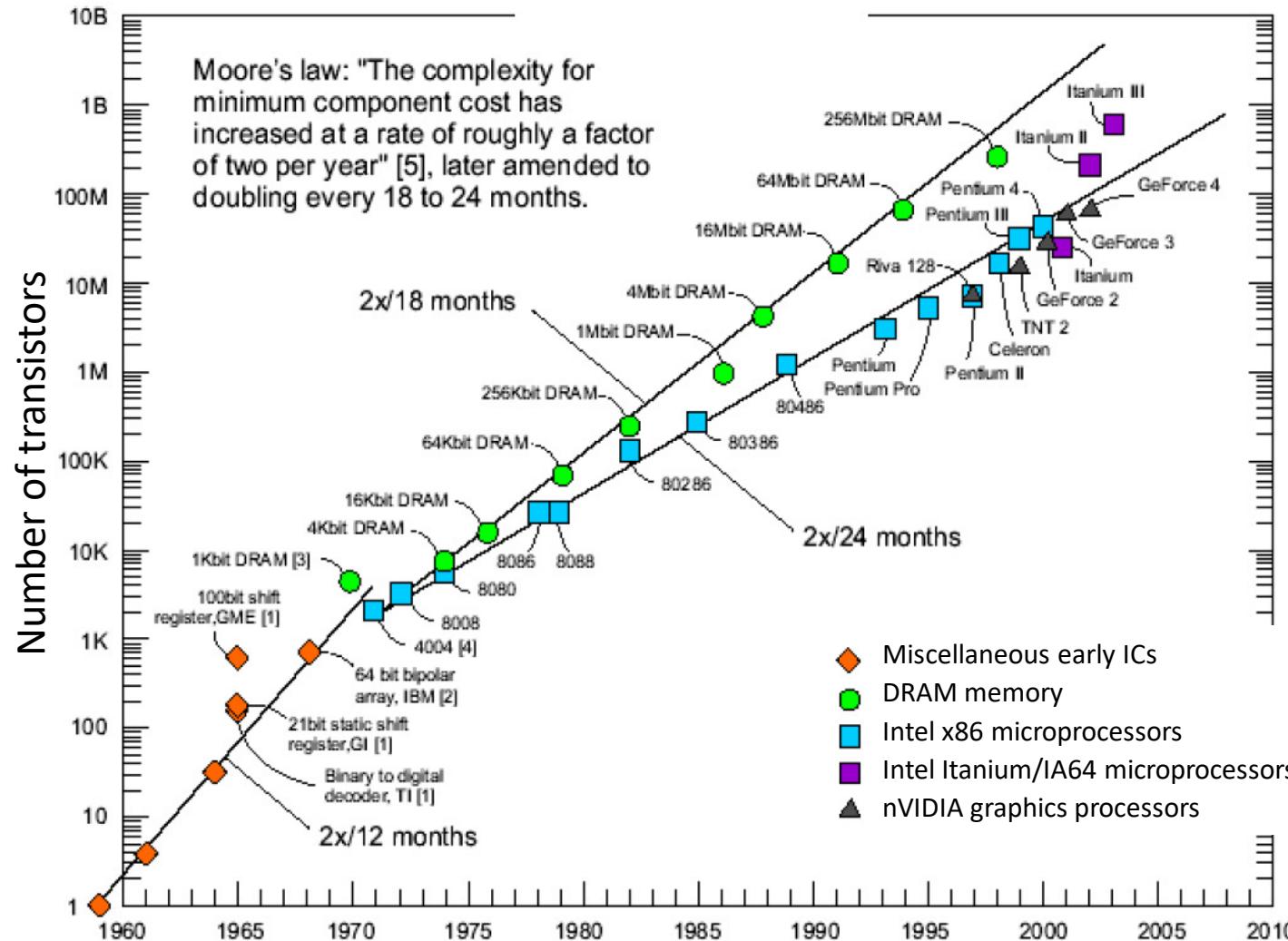
Device scaling down over time in IC industry

Moore's law: doubling of the number of transistors on a chip roughly every two years.

This is realized by:

Making transistor smaller - smallest lateral feature size decreases by 13% each year.

Making chip bigger – chip/wafer size increases 16%/year.



Gordon Moore:
born 3 January 1929,
co-founder and
Chairman Emeritus of
Intel Corporation;
author of Moore's Law
published in 1965.

Device scaling down over time in IC industry

Assumes CMOS technology dominates over entire roadmap.

2 year cycle moving to 3 years (scaling + innovation now required).

SIA-NTRS:	2 year cycle				3 year cycle				
Year of Production	1998	2000	2002	2004	2007	2010	2013	2016	2018
Technology Node (half pitch)	250 nm	180 nm	130 nm	90 nm	65 nm	45 nm	32 nm	22 nm	18 nm
MPU Printed Gate Length (MPU - microprocessor unit)		100 nm	70 nm	53 nm	35 nm	25 nm	18 nm	13 nm	10 nm
DRAM Bits/Chip (Sampling)	256M	512M	1G	4G	16G	32G	64G	128G	128G
MPU Transistors/Chip (x10 ⁶)				550	1100	2200	4400	8800	14,000
Min Supply Voltage (volts)	1.8-2.5	1.5-1.8	1.2-1.5	0.9-1.2	0.8-1.1	0.7-1.0	0.6-0.9	0.5-0.8	0.5-0.7

Each node half pitch $\times(1/\sqrt{2})$, area per transistor $\times(1/2)$, i.e. for same chip size, # of transistors doubled

Scaling down supply voltage because otherwise, as transistors get smaller, the electric fields (voltage/feature size) in these devices will increase to unacceptable levels.

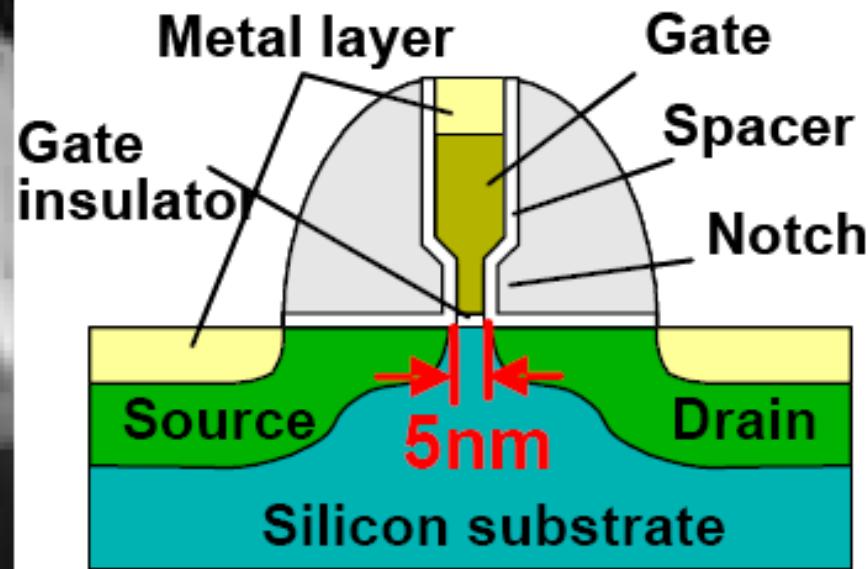
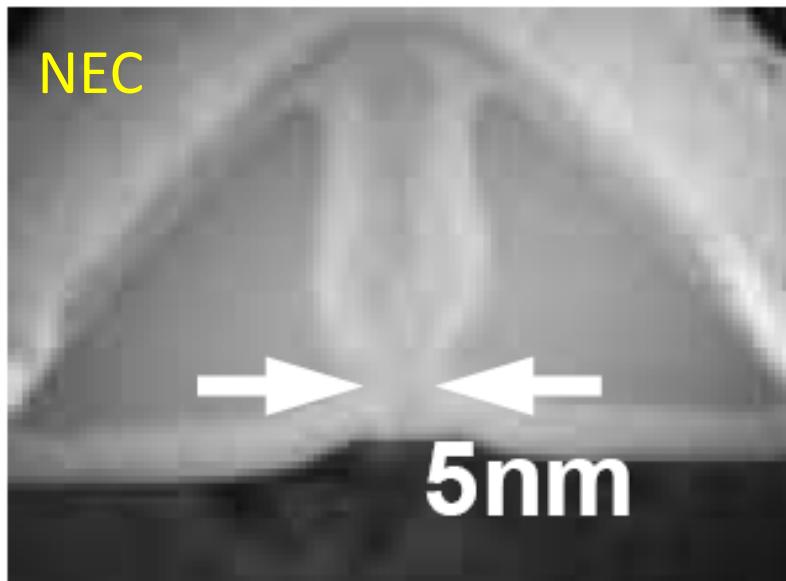
In addition, the number of levels of interconnection and photo-mask also increases.

SIA: Semiconductor Industry Association

NTRS: National Technology Roadmap for Semiconductors.

ITRS: International Technology Roadmap for Semiconductors, <http://www.itrs.net>

Fabrication of truly tiny transistor



Photograph and figure of 5-nm-gate transistor.

However, the key is how to fabricate them with high yield and low cost.

More importantly, even though it can be fabricated, it may not function the way we want (quantum effect, leak current...).

Chapter 1 Introduction and Historical Perspective

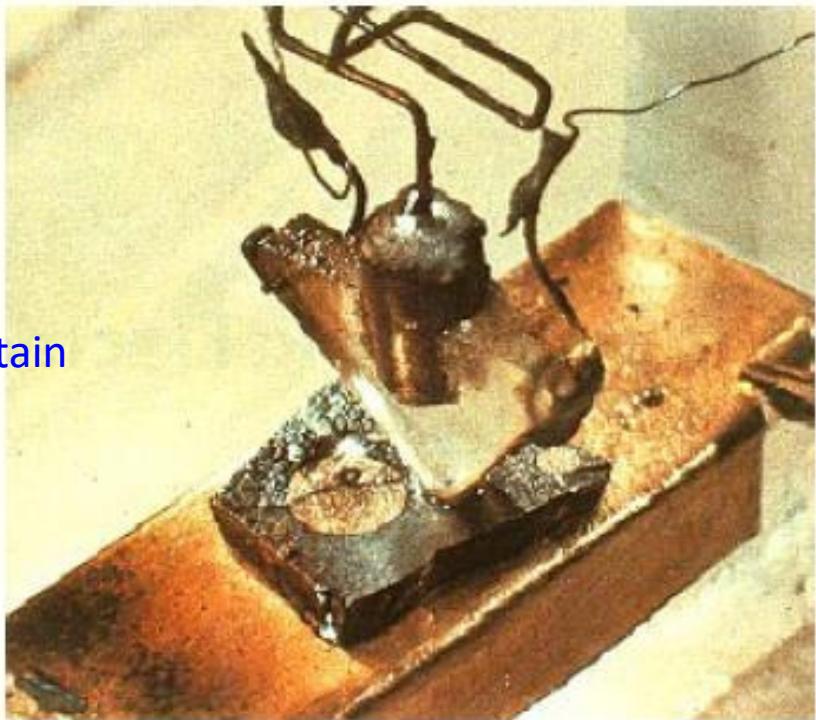
1. Introduction.
2. Growth of IC – Moore's law.
3. Some history in IC industry.

IC fabrication technology: brief history

- 1940s - setting the stage - the initial inventions that made integrated circuits possible.
- In 1945, Bell Labs established a group to develop a semiconductor replacement for the vacuum tube. The group led by William Shockley, included, John Bardeen, Walter Brattain and others.
- In 1947 Bardeen and Brattain and Shockley succeeded in creating an amplifying circuit utilizing a point-contact "transfer resistance" device that later became known as a transistor.
- In 1951 Shockley developed the junction transistor, a more practical form of the transistor.
- By 1954 the transistor was an essential component of the telephone system and the transistor first appeared in hearing aids followed by radios.

Some history in IC industry: first transistor

1st transistor in 1947 by Bell Lab, it is a point contact transistor.



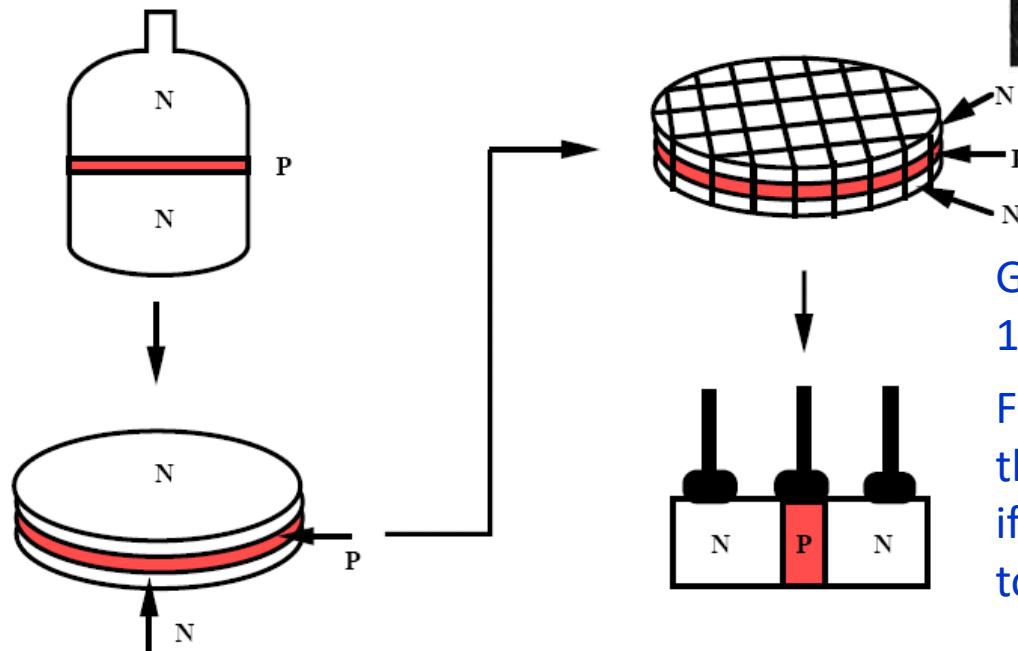
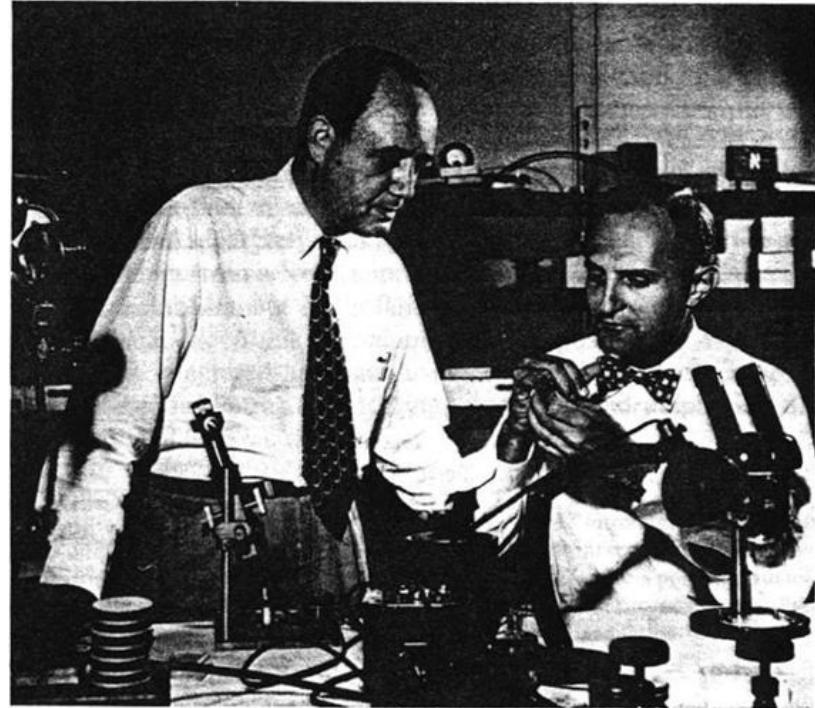
(Reprinted with permission of Lucent Technologies).

Bipolar transistor in
polycrystalline Germanium,
1956 Nobel Prize in physics.

Base is n-type Ge.
Emitter and collector are two metal wires, which
are very thin and pushed onto the Ge base.
Distance between the two metal wires: 200-250 μ m.

First junction transistor

Gordon Teal and Morgan Sparks made the first *junction* transistor, the construction of which eliminated many of the reliability problems of the *point contact* transistors.



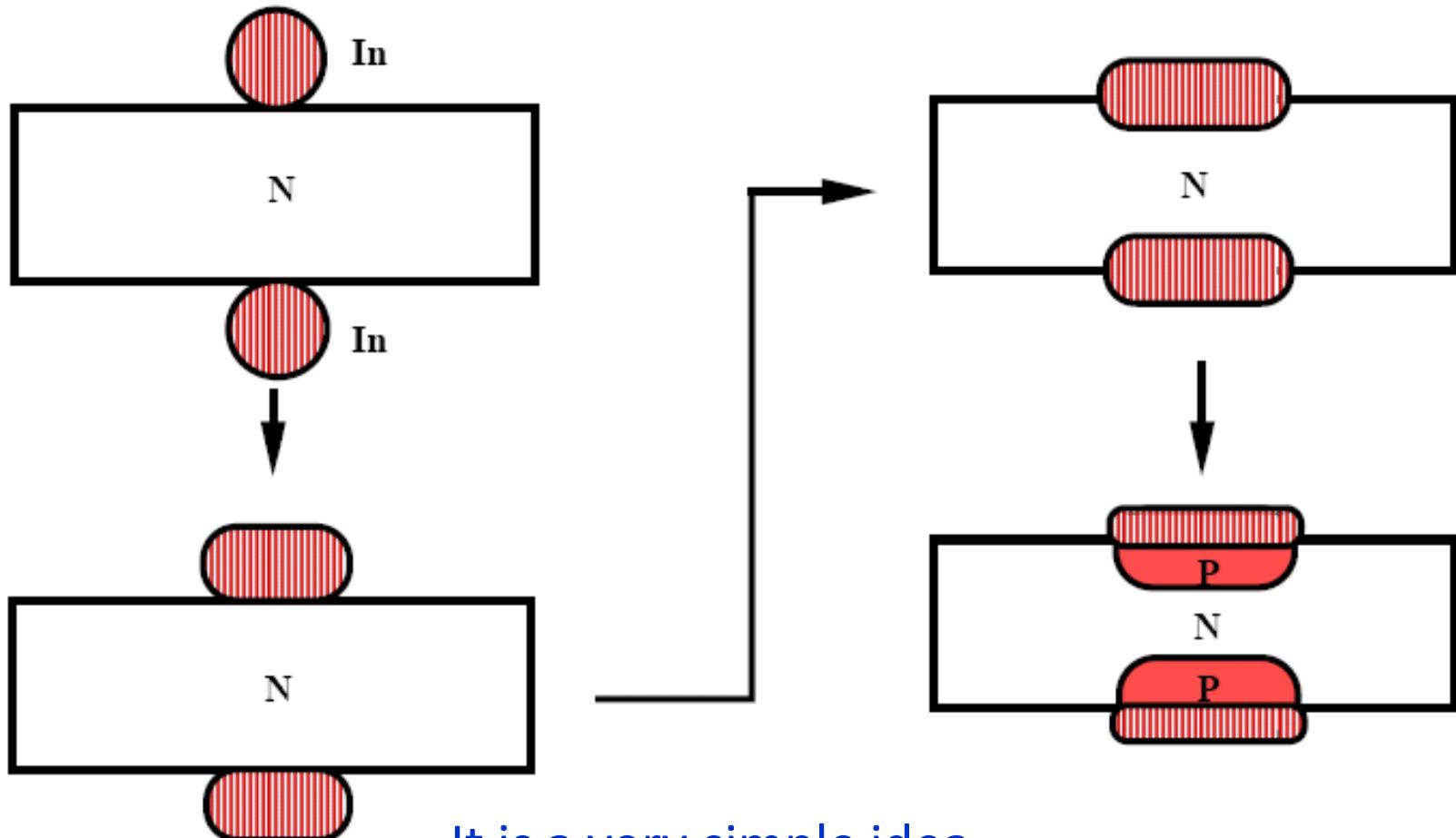
Grown junction transistor technology of the 1950s, in single crystalline Si or Ge.

For Si device, Al wire is used to connect to the middle P base region (it doesn't matter if Al is also contacted to the N-regions due to the high contact resistance with N).

Alloy junction technology of the 1950s

Indium melts at 157°C, and it is a P-type dopant.

(In is in the same group as B, the most popular P-type dopant. B, Al, Ga, In...)



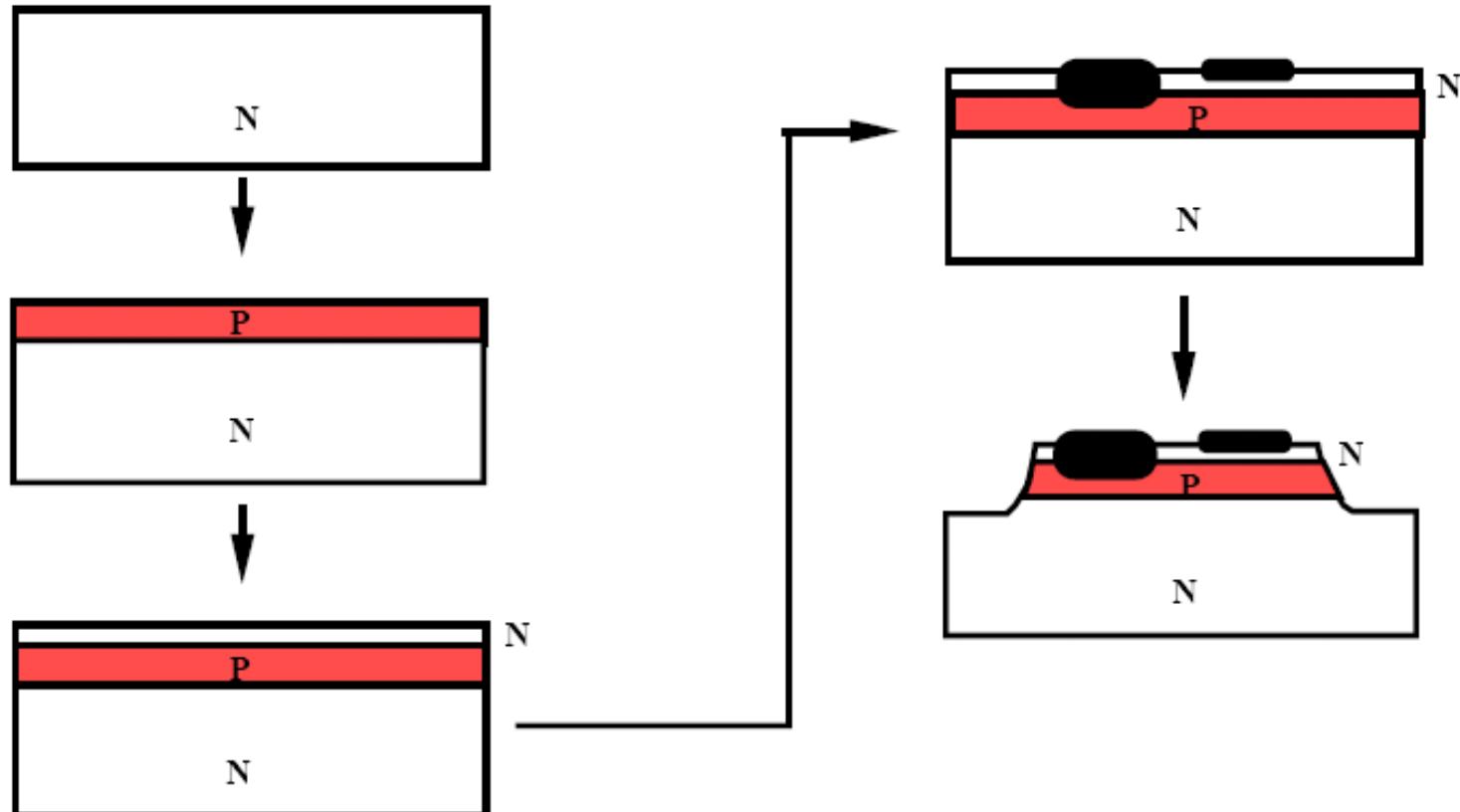
It is a very simple idea.

Doubled diffused mesa transistor technology of the late 1950s

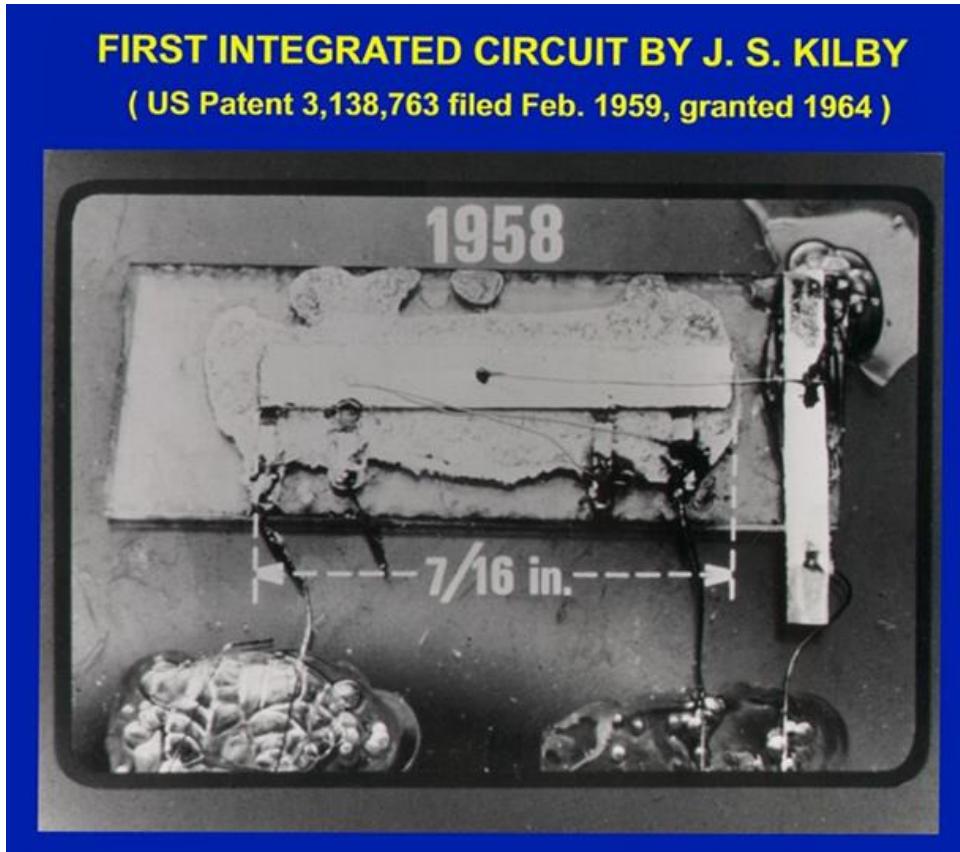
Gas phase diffusion (e.g. PH_3 gas to dope with P) to dope the silicon.

Many devices could be produced from a single substrate.

But exposed junctions were present on the wafer surface or at wafer edges.



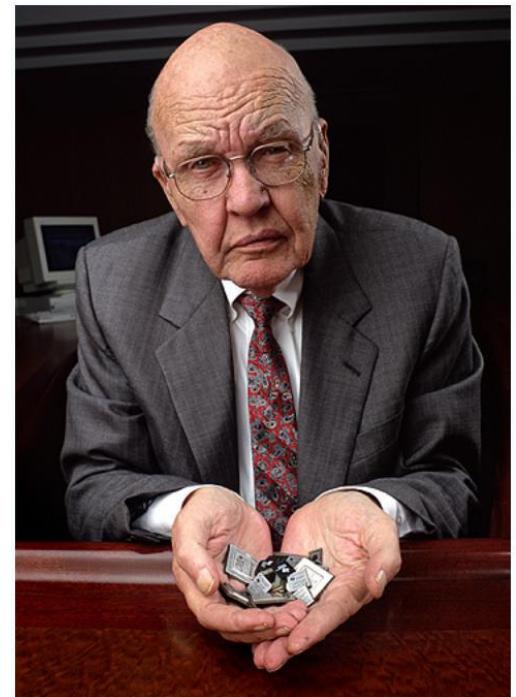
Integrated Circuit (IC) invented by Kilby from TI



A simple oscillator IC with five integrated components (resistors, capacitors, distributed capacitors and transistors) on Ge substrate.

To read (if interested) “Turning Potential into Realities: The Invention of the Integrated Circuit (Nobel Lecture)”

IC: integrate multiple components on the same chip and to interconnect them to form a circuit.

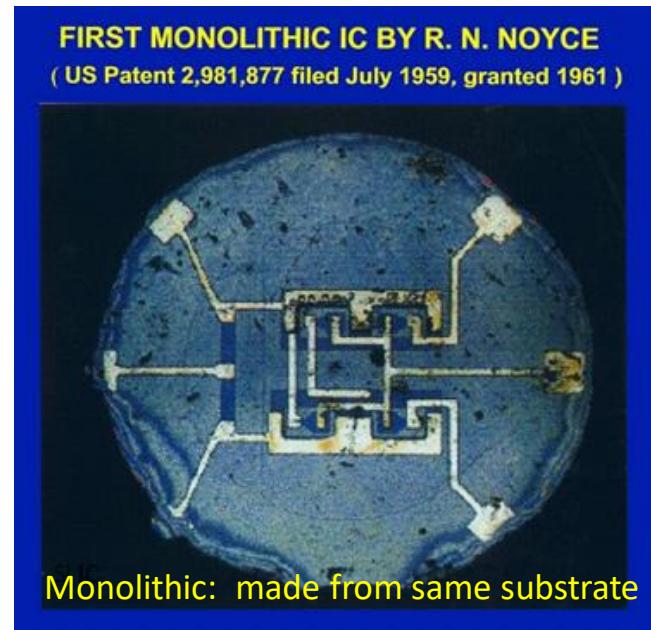


Jack Kilby, Nobel Prize in Physics in 2000

TI: Texas Instrument

Planar process invented in the late 1950s

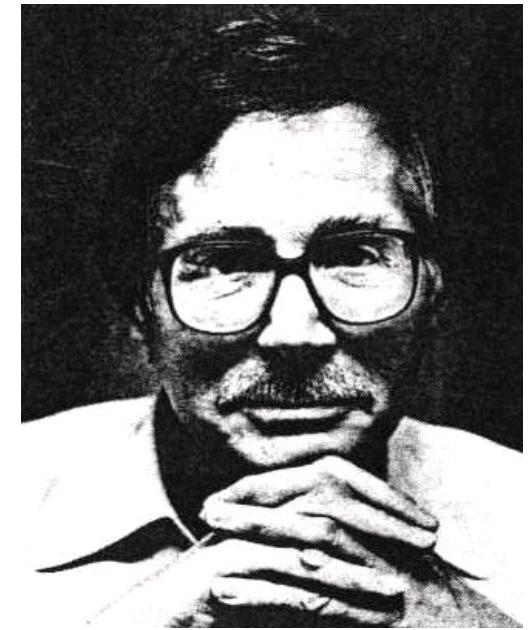
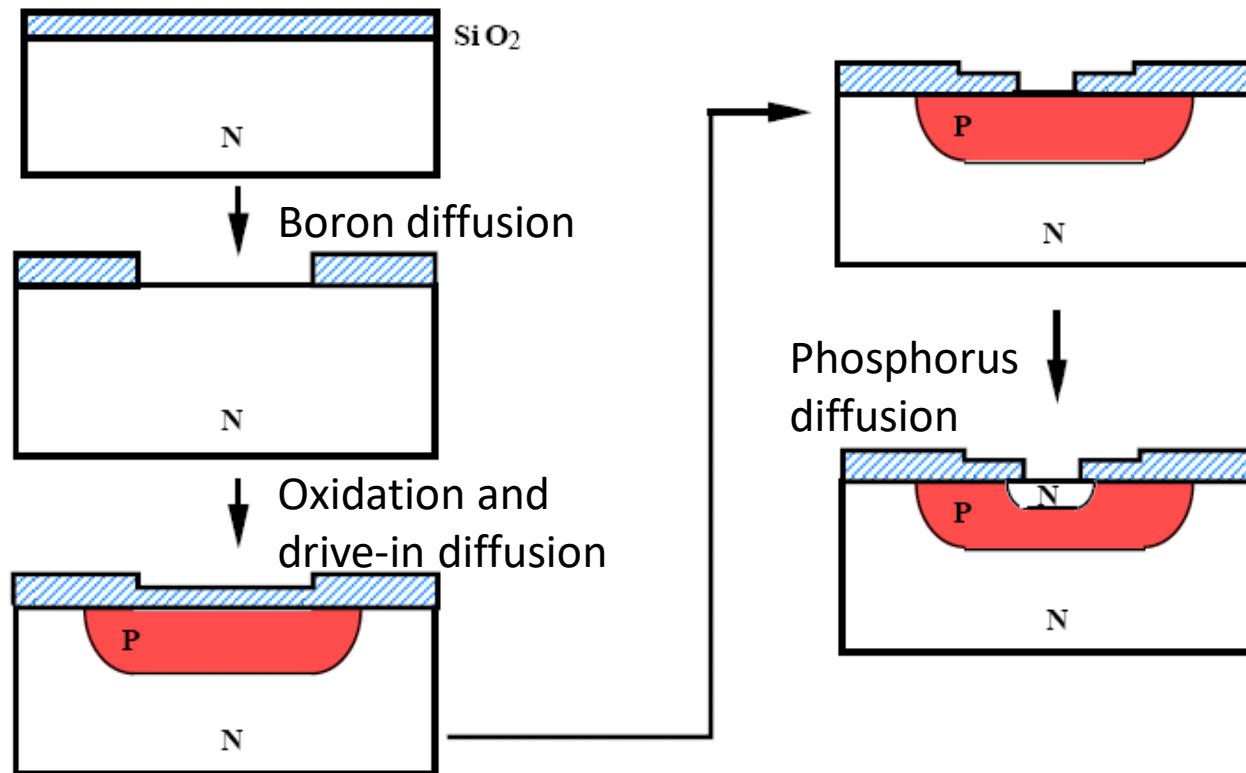
- Kilby's invention had a serious drawback, the individual circuit elements were connected together with gold wires making the circuit difficult to scale up to any complexity.
- By late 1958 Jean Hoerni at Fairchild had developed a structure with N and P junctions formed in silicon. Over the junctions a thin layer of silicon dioxide was used as an insulator and holes were etched open in the silicon dioxide to connect to the junctions.
- In 1959, Robert Noyce also of Fairchild had the idea to evaporate a thin metal layer over the circuits created by Hoerni's process.
- The metal layer connected down to the junctions through the holes in the silicon dioxide and was then etched into a pattern to interconnect the circuit.
- Planar technology set the stage for complex integrated circuits and is the process used today.



Planar technology

Planar process invented in the late 1950s

- Gas phase diffusion masked by SiO_2 .
- SiO_2 patterned by photolithography.
- Since only Si has this perfect oxide that can block the diffusion, technology was shifted from Ge to Si.
- Junction is under SiO_2 surface (no longer exposed to surface/edges), it is thus passivated/protected.



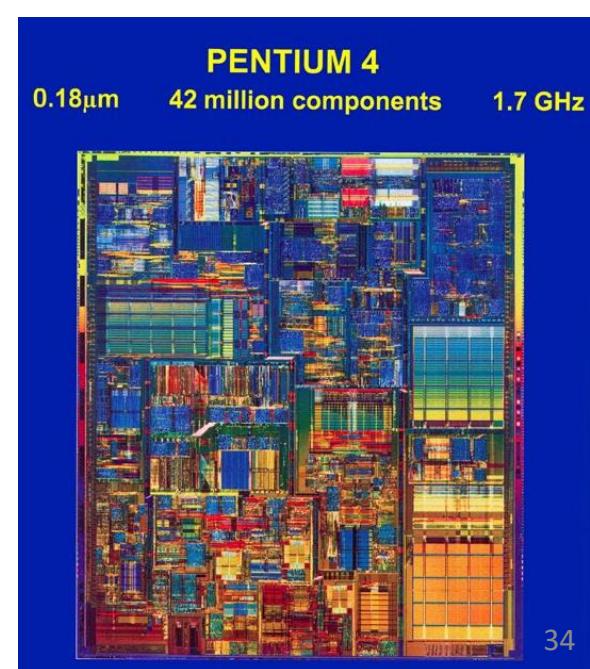
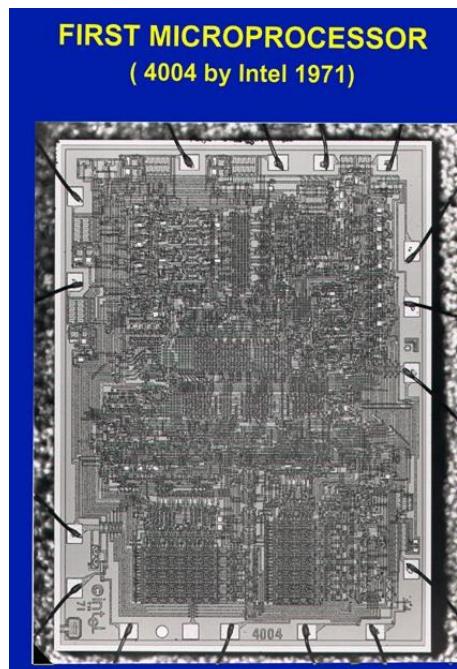
Jean Hoerni from Fairchild,
inventor of planar process

Breakthroughs in IC history (summary)

- Bardeen, Brattain, Shockley, First Ge-based bipolar transistor invented 1947, Bell Labs. Nobel prize in 1956.
- Atalla, First Si-based MOSFET invented 1958, Bell Labs.
- Kilby (TI) & Noyce (Fairchild), Invention of integrated circuits 1959, Nobel prize in 2000.
- Planar technology, Jean Hoerni, Fairchild, 1959
- First CMOS circuit invented 1963, Fairchild
- “Moore’s law” coined 1965, Fairchild
- Dennard, scaling rule presented 1974, IBM
- First Si technology roadmap published 1994, USA

Intel was founded in 1968, by Gordon Moore and Robert Noyce, both from Fairchild.

For Fairchild's history, go to
http://en.wikipedia.org/wiki/Fairchild_Semiconductor

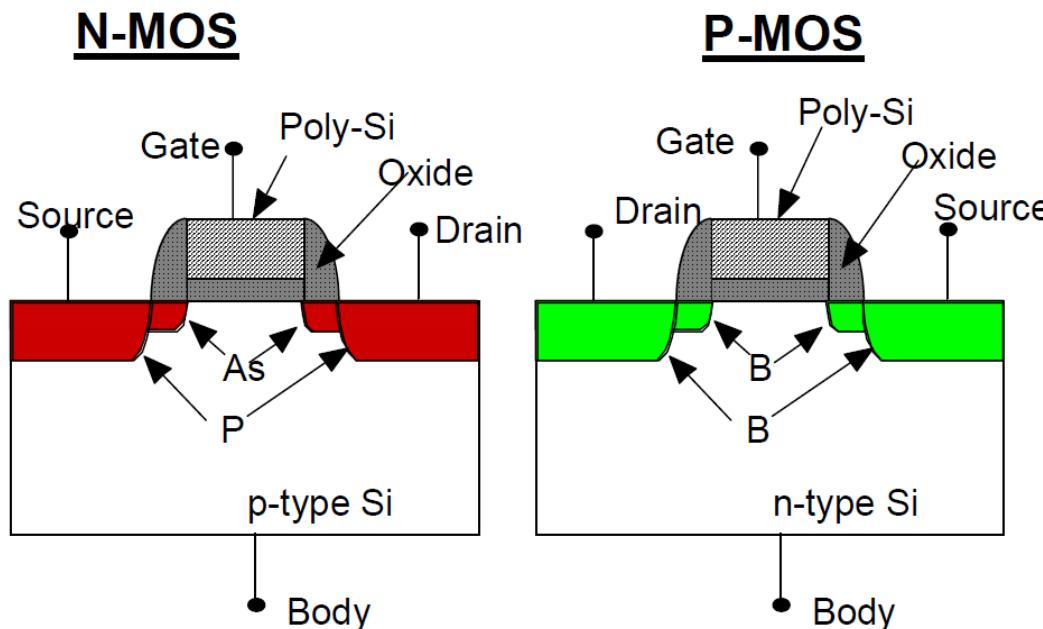


Chapter 2 Modern CMOS technology

1. Introduction.
2. CMOS process flow.

CMOS: complementary metal–oxide–semiconductor

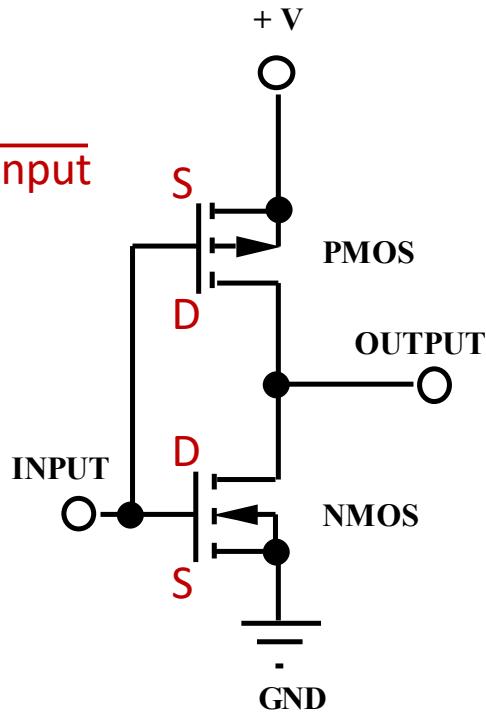
- In the simplest CMOS technologies, we need to realize simply NMOS and PMOS transistors for circuits like those illustrated below.
- Typical CMOS technologies in manufacturing add additional steps to implement multiple device V_{TH} , thin film transistors (TFT) in SRAMs, capacitors for DRAMs etc.
- CMOS described here requires 16 masks (through metal level 2) and >100 process steps.
- There are many possible variations on the process flow (e.g. LOCOS device isolation vs. shallow trench isolation).



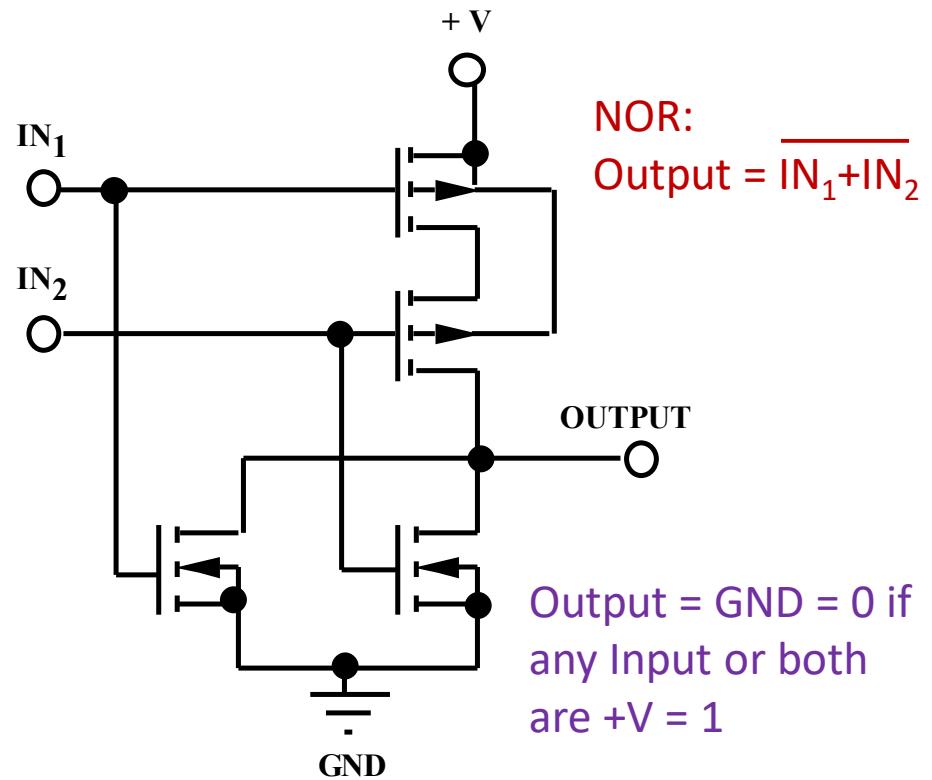
n-MOS & p-MOS require different channel background doping and source/drain region doping.
In CMOS, the gate is no longer “metal”, it is heavily doped poly-crystalline Si with low resistance.

CMOS is required by logic circuits

Inverter:
Output = $\overline{\text{Input}}$



NOR:
Output = $\overline{\text{IN}_1 + \text{IN}_2}$



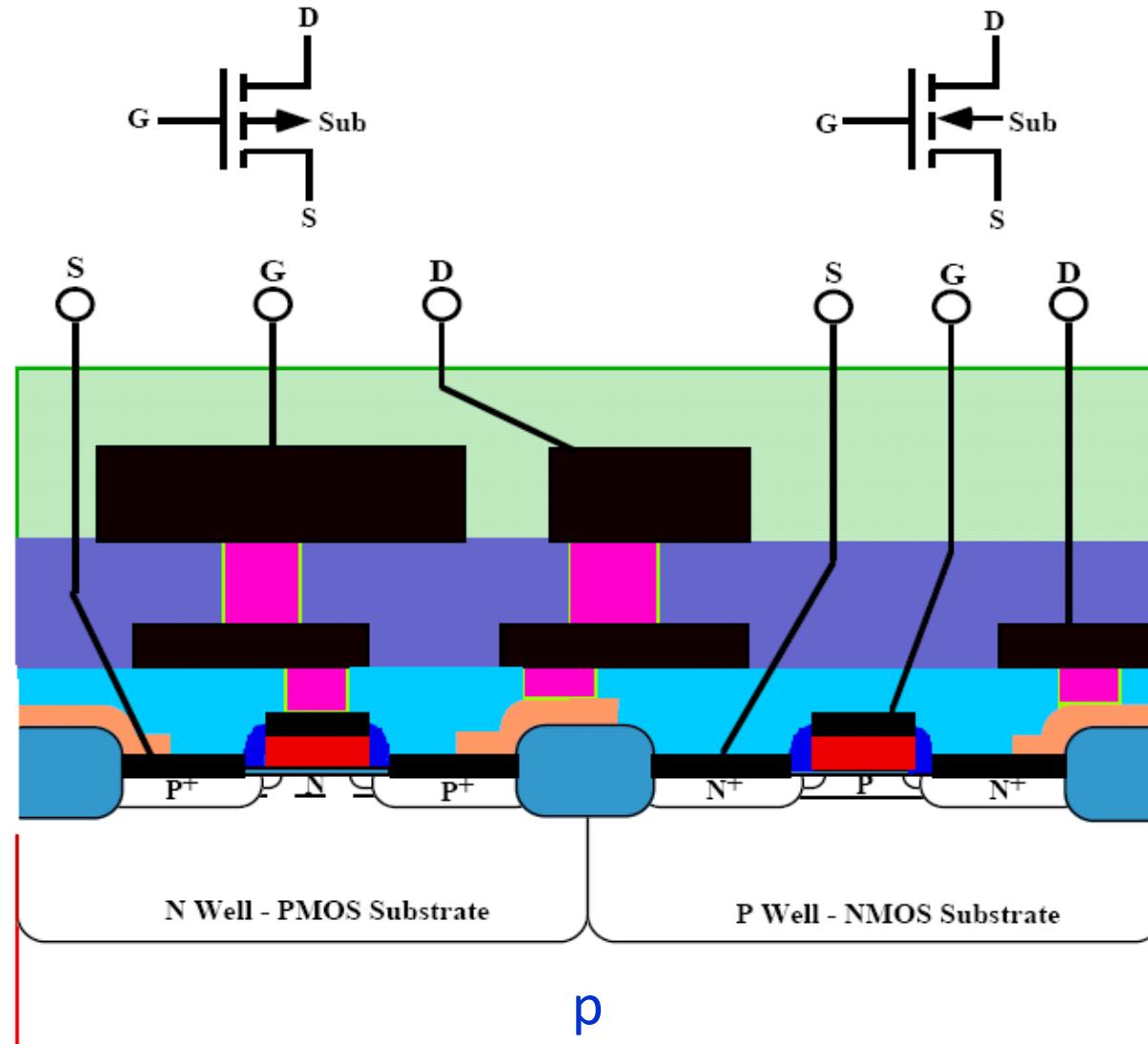
Output = GND = 0 if
any Input or both
are +V = 1

CMOS (n-MOS & p-MOS) reduces static power dissipation.

Because (e.g. for the inverter) there is no current flow from +V to GND since one of the MOS is always off.

The same inverter logic can also be realized by replacing the top PMOS with a resistor R (ON NMOS << R << OFF NMOS), but current flows when NMOS is on.

Cross-section of the CMOS IC



This is what we are going to fabricate in this chapter.

Fabrication “toolkit”

- Insulating Layers
 - Oxidation, nitridation
 - Deposition (LPCVD, PECVD, APCVD)
- Selective doping of silicon
 - Diffusion (in-situ doping)
 - Ion implantation
 - Epitaxy (in-situ doping)
- Material deposition (silicon, metals, insulators)
 - LPCVD
 - PECVD
 - Sputter deposition
- Patterning of Layers
 - Lithography (UV, deep UV, e-beam & x-ray)
- Etching of (deposited) material
 - Dry etches—plasma, RIE, sputter etch, DRIE
 - Wet etches—etch in liquids, CMP etc

LPCVD: low pressure chemical vapor deposition.

PECVD: plasma enhanced CVD.

APCVD: atmospheric pressure CVD

RIE: reactive ion etching

DRIE: deep RIE.

CMP: chemical mechanical polishing

Chapter 2 Modern CMOS technology

1. Introduction.
2. CMOS process flow.

Choosing the substrate and active region formation



Nitride has high tensile stress, oxide has compressive stress.
The two stress can balance/compensate each other to reduce
stress in Si that may cause defects in Si.

LPCVD nitride: $3\text{SiH}_4 + 4\text{NH}_3 \rightarrow \text{Si}_3\text{N}_4 + 12\text{H}_2$, 800°C.
LPCVD: low pressure chemical vapor deposition

Si, (100), P Type, 5-50 Ωcm

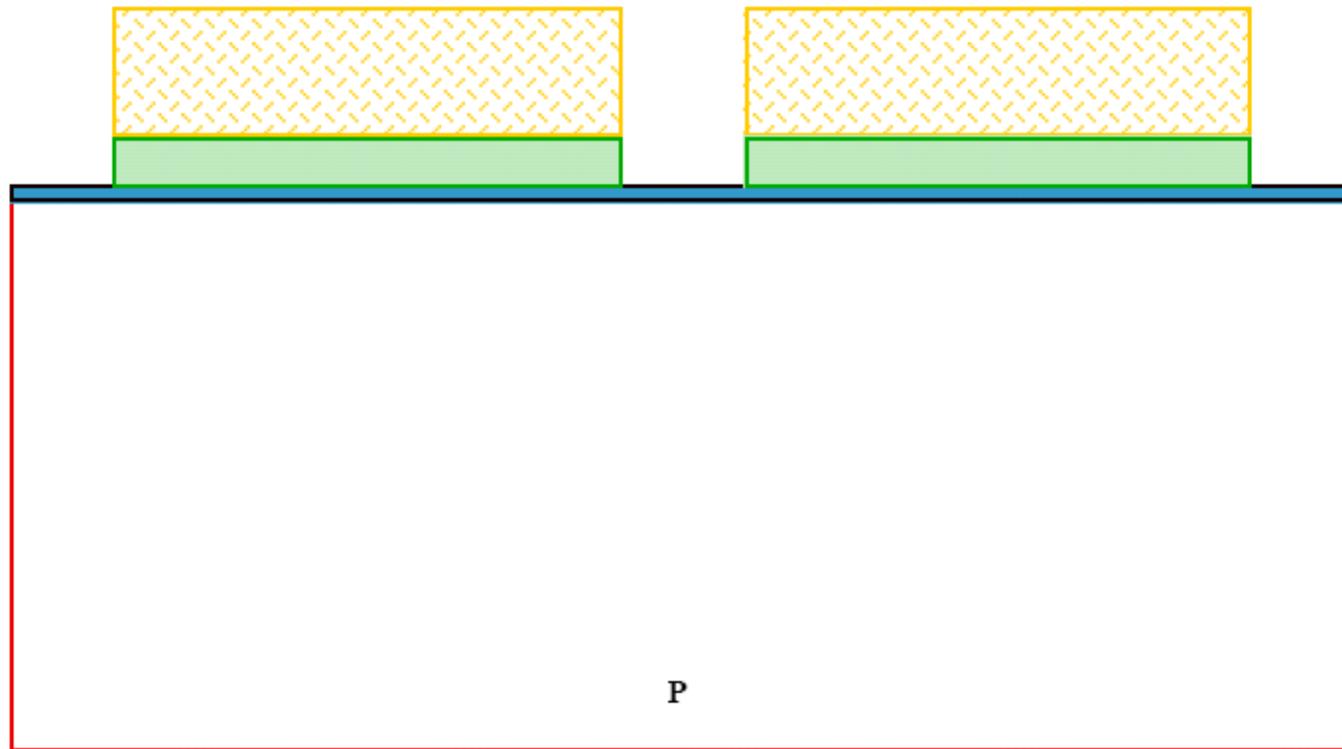
Substrate selection: moderately high resistivity (lightly doped, 10^{15}cm^{-3}), (100) orientation substrate (better Si/SiO₂ interface than other orientations), P type.

Start from low doping, then dope P-well and N-well by ion implantation that is much better controlled than substrate doping (done during crystal growth).

Wafer cleaning, thermal oxidation (≈ 40 nm, using O₂, or H₂O generated from H₂ and O₂ reaction, cleaner than H₂O vapor from boiling water), Si₃N₄ LPCVD (≈ 80 nm), photoresist spinning and baking ($\approx 0.5 - 1.0 \mu\text{m}$).

Active region formation

Photolithography, nitride etching



Mask #1 patterns the active areas. The nitride is dry etched.

Dry etch = plasma etch, reactive species are generated in a plasma (like arc discharge). E.g F is generated in CF_4 plasma. Atomic F is extremely reactive.



LOCOS isolation

LOCOS: LOCal Oxidation of Silicon

Remove resist, thermal oxidation



Si_3N_4 is very dense material and prevents/blocks H_2O or O_2 from diffusion to the Si surface, thus no oxidation under nitride.

P

Remove photoresist.

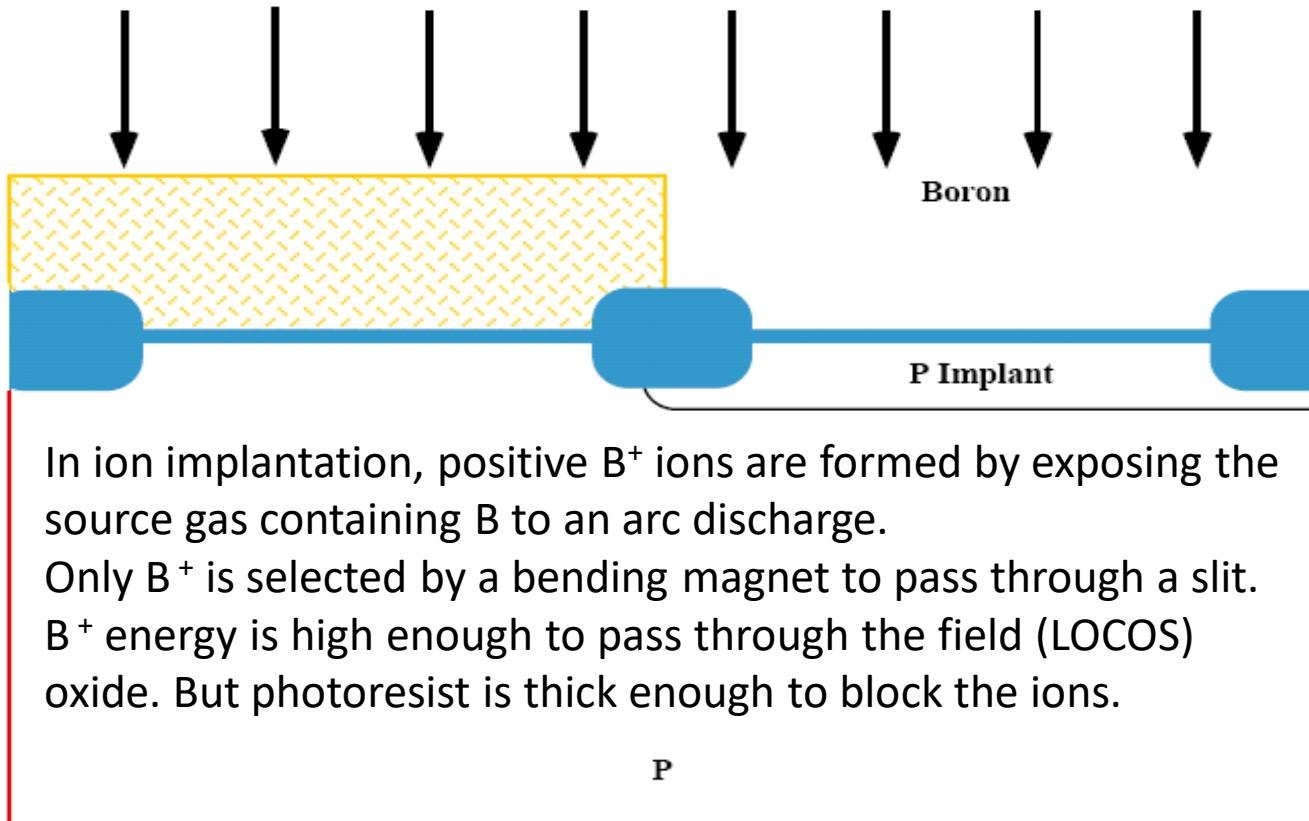
Field oxide is grown using a LOCOS process.

Typically 90min @ 1000°C in H_2O grows $\text{SiO}_2 \approx 0.5 \mu\text{m}$.

Field oxide is partially recessed into the surface (oxidation consume some of the silicon)
Field oxides forms a lateral extension under the nitride layer – bird's beak region
Bird's beak region limits device scaling and device density in VLSI circuits!

P-well formation

Wet etch away Si_3N_4 , spin photoresist, lithography, B^+ implantation.



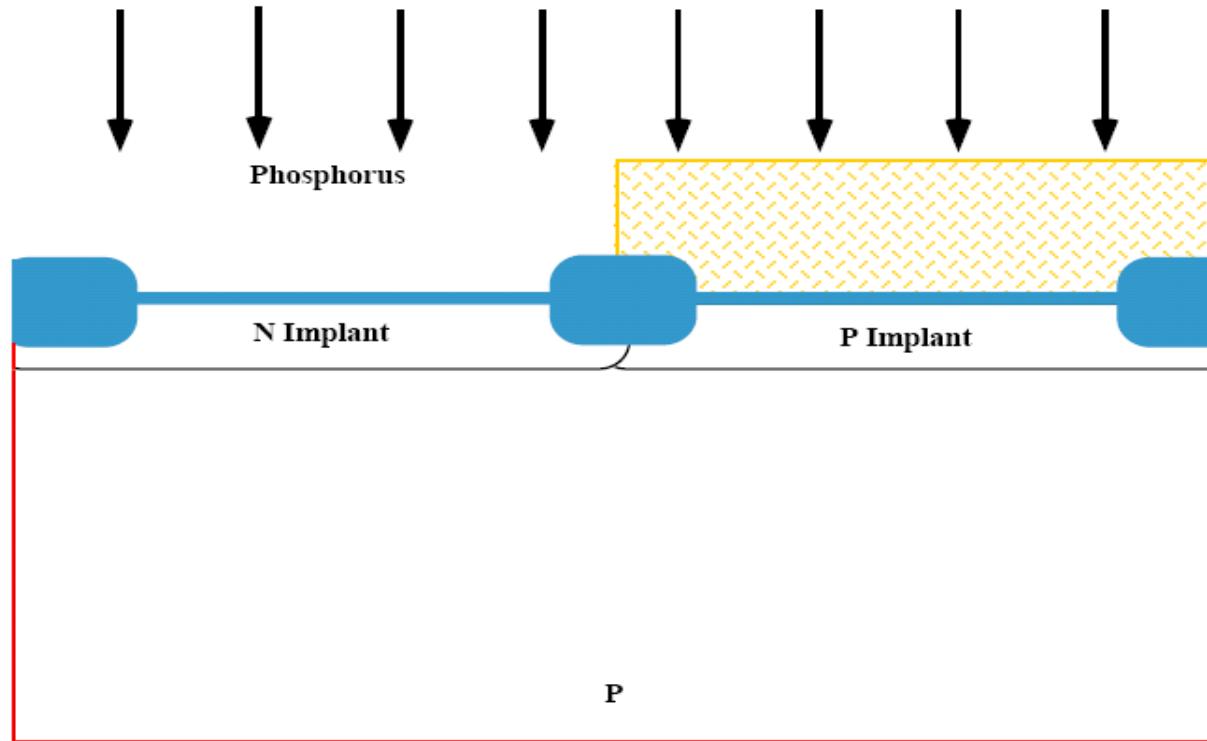
In ion implantation, positive B^+ ions are formed by exposing the source gas containing B to an arc discharge.

Only B^+ is selected by a bending magnet to pass through a slit. B^+ energy is high enough to pass through the field (LOCOS) oxide. But photoresist is thick enough to block the ions.

Mask #2 blocks a B^+ implant to form the wells for the NMOS devices. Typically dose 10^{13}cm^{-2} @ 150-200 KeV (very high energy). (Implant dose is in cm^{-2} , doping concentration is in cm^{-3})

N-well formation

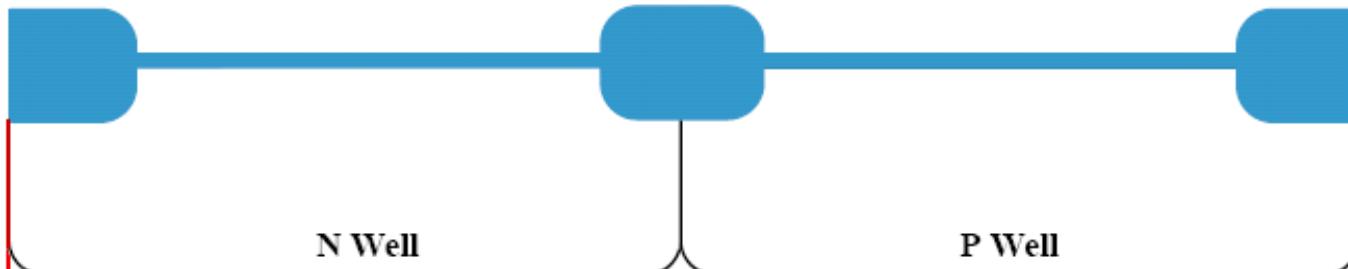
Strip photoresist, spin resist and photolithography, ion implantation



Mask #3 blocks a P^+ implant to form the wells for the PMOS devices.
Typically 10^{13} cm^{-2} @ 300-400 KeV.
(P is heavier than B, so higher energy needed)

N- and P- well formation

Remove resist and anneal



Ion energy is ~100keV, much higher than energy needed to break 4 Si bonds (total 12eV), so ion implantation induces many damages.

B and P have similar diffusion coefficient, so similar final well depth.

P

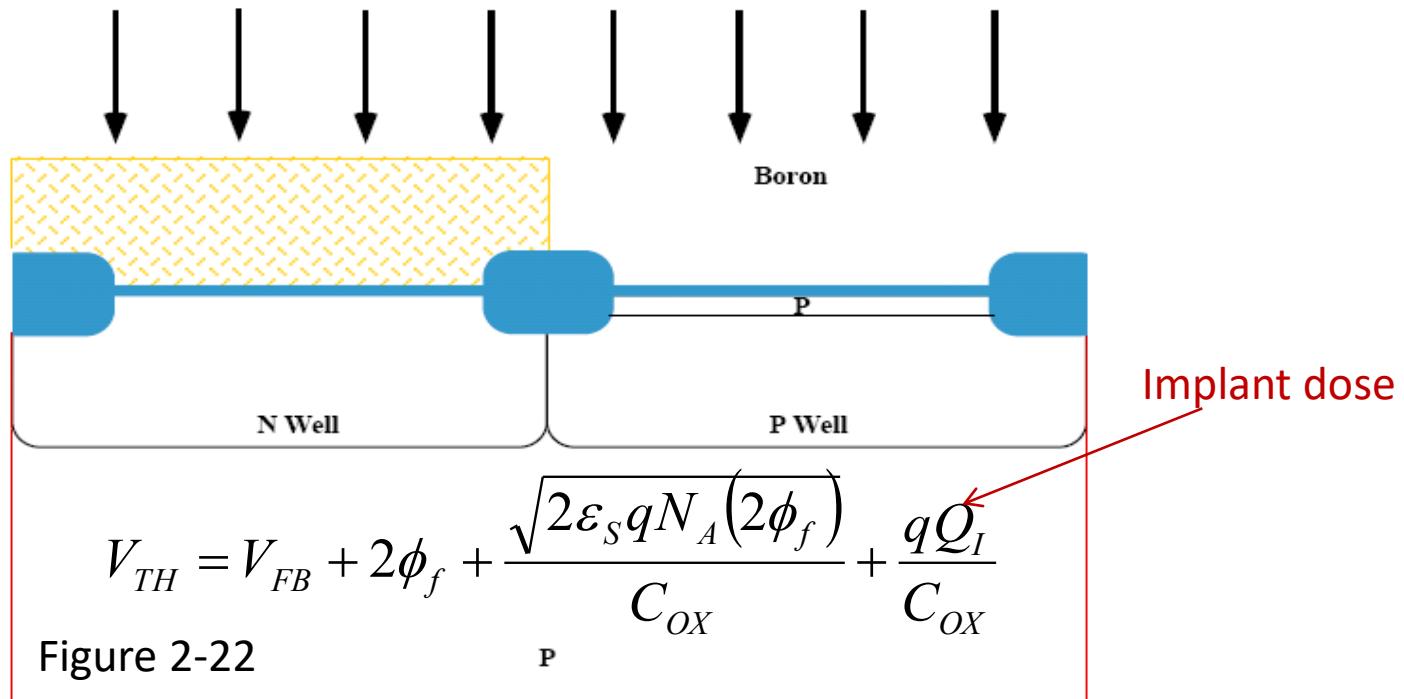
A high temperature drive-in produces the “final” well depths and repairs implant damage.

Typically 4-6 hours @ 1000°C - 1100°C or equivalent Dt .

(here D is diffusion coefficient, t is time)

Threshold voltage (V_{TH}) adjustment

Spin photoresist, photolithography, B^+ ion implantation



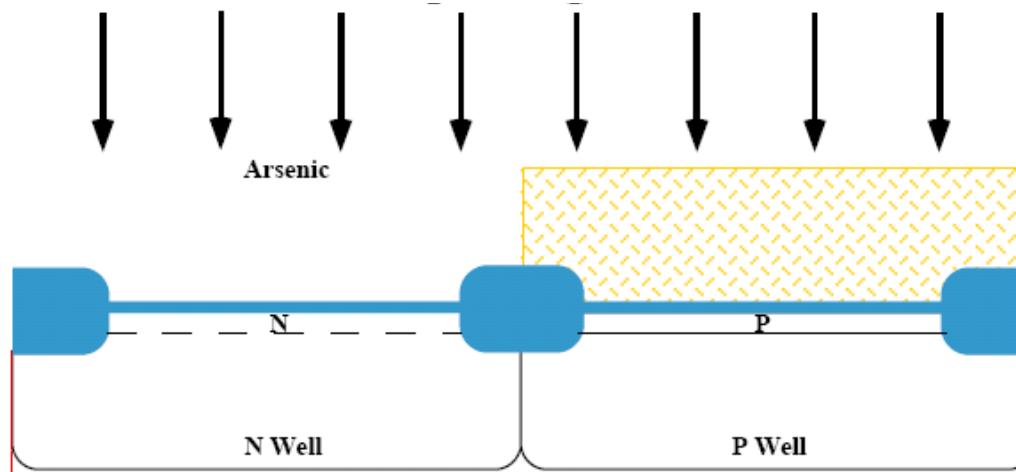
Mask #4 is used to mask the PMOS devices.

A V_{TH} adjust implant is done on the NMOS devices.

Typically $1-5 \times 10^{12} \text{ cm}^{-2}$ B^+ implant @ 50 - 75 KeV.

Threshold voltage (V_{TH}) adjustment

Remove resist, then spin photoresist, photolithography, As^+ ion implantation



Again, adjust V_{TH} by controlling implant dose Q_i .

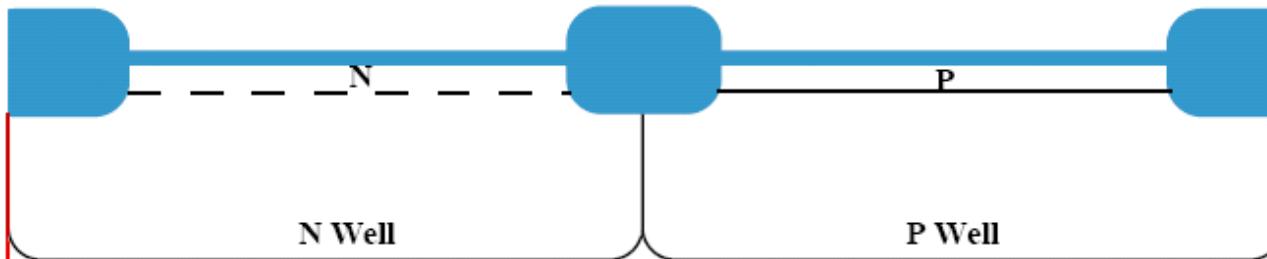
Mask #5 is used to mask the NMOS devices.

A V_{TH} adjust implant is done on the PMOS devices.

Typically $1-5 \times 10^{12} \text{ cm}^{-2}$ As^+ implant @ 75 - 100 KeV.

Gate oxide growth

Remove resist, etch oxide, re-grow thermal oxide



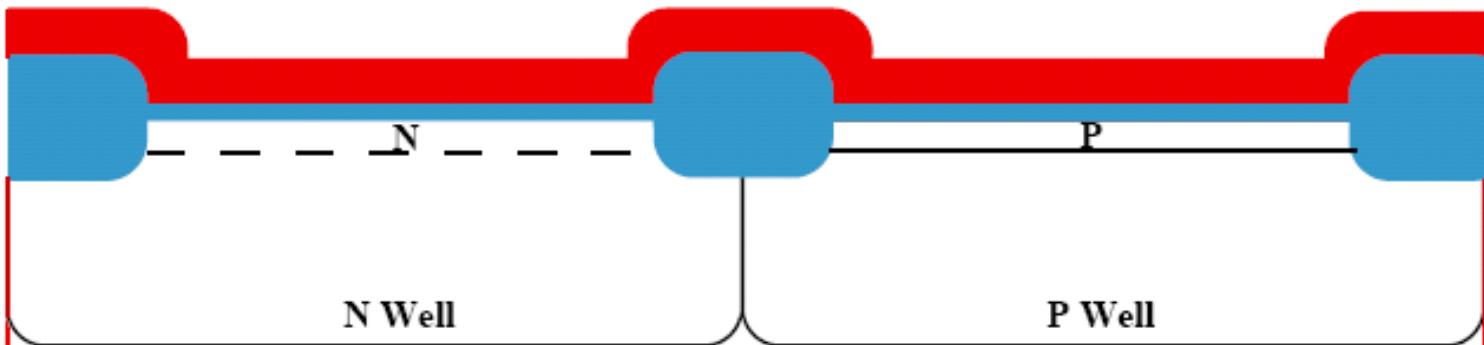
The 'old' oxide (to compensate stress of Si_3N_4) is too thick, and may be damaged during the several implantation steps.

P

Figure 2-24

The thin oxide over the active regions is stripped and a new gate oxide grown, typically 3 - 5nm, which could be grown in 0.5 - 1 hrs @ 800°C in O_2 .

Poly-crystalline silicon deposition



MOSFET: metal oxide semiconductor field effect transistor

But actually metal is no longer used, instead, low resistance heavily doped poly-Si is used.



P

Poly-silicon is deposited by LPCVD ($\approx 0.5 \mu\text{m}$).

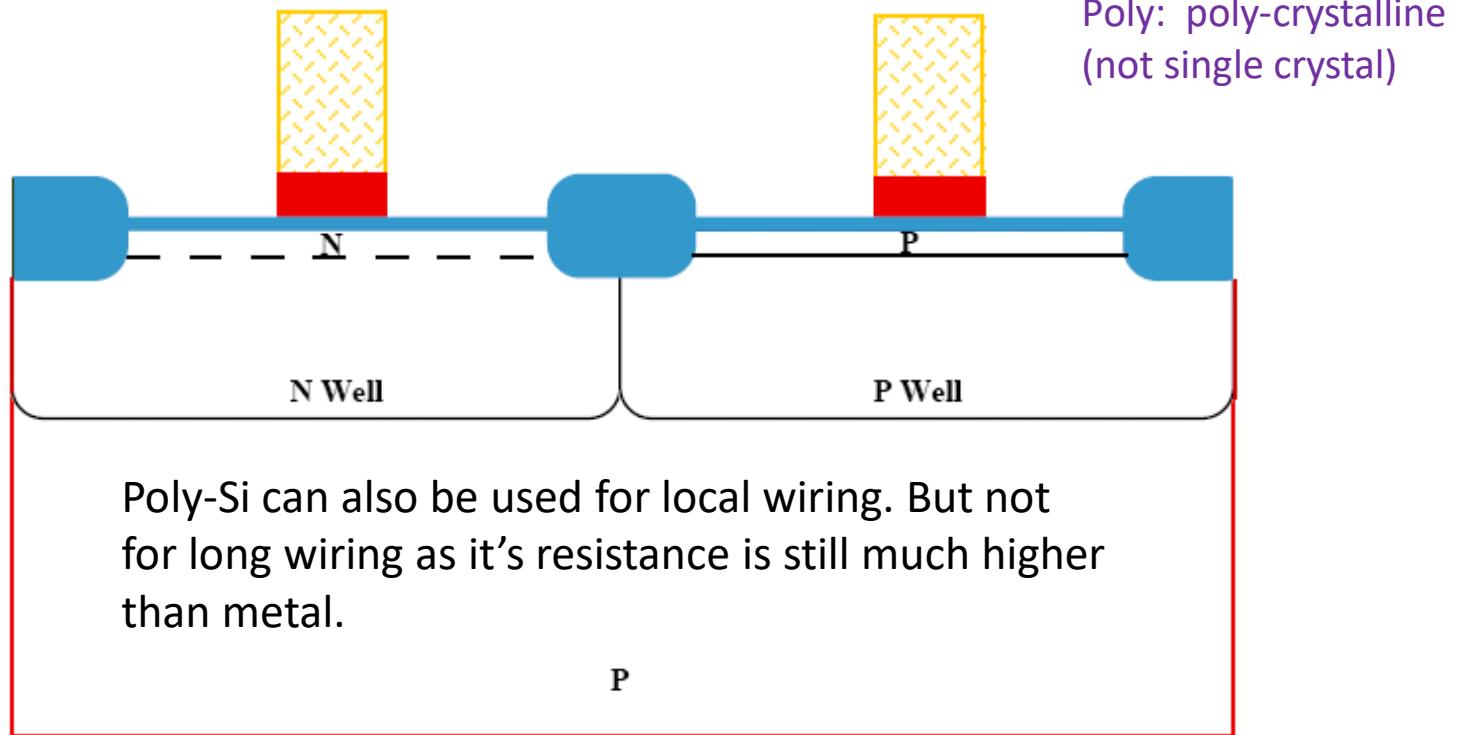
An unmasked P^+ or As^+ implant dopes the poly (typically $5 \times 10^{15} \text{ cm}^{-2}$, high doping to reduce gate resistance).

Both P and As have high solubility in Si, good for heavy doping.

When heated, they will diffuse quickly through grain boundary (now that *poly*) to achieve uniform doping.

Gate formation

Spin resist, photolithography, selective anisotropic etch of poly-Si.



Poly-Si can also be used for local wiring. But not for long wiring as it's resistance is still much higher than metal.

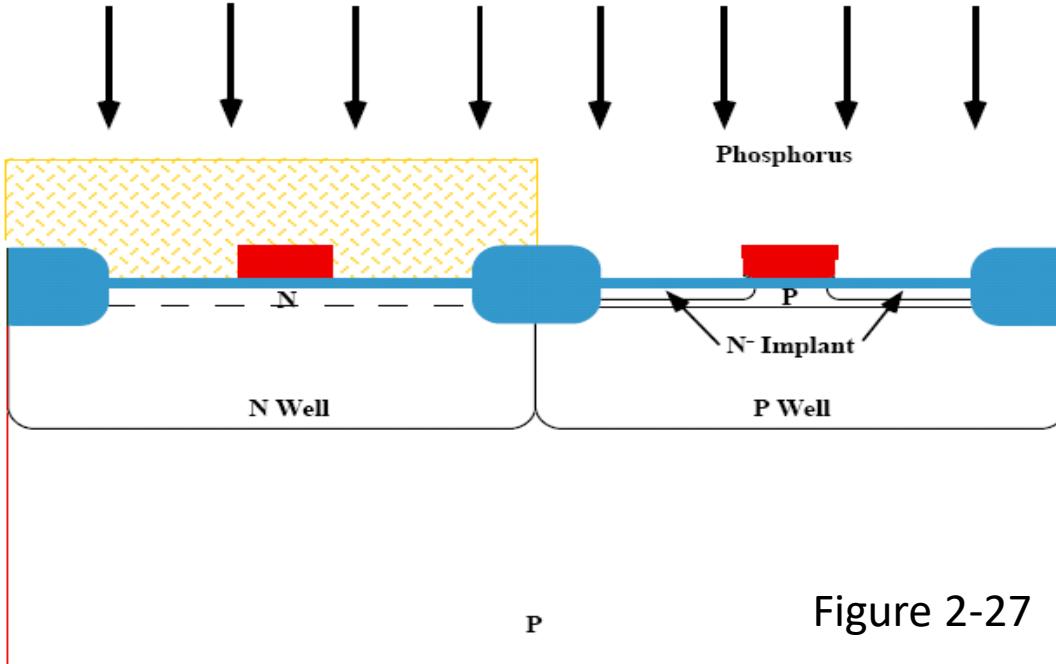
Mask #6 is used to protect the MOS gates.

The poly-Si is plasma etched using an anisotropic etch.

The photolithography in this step is the most demanding since it requires the finest resolution to create the narrow MOS channels.

Tip or extension/LDD (lightly doped drain) formation

Strip resist, spin resist, photolithography, ion implantation



When channel length shrinks more than drive voltage, electric field in the channel may become very high, creating “hot electrons” that may create additional hole-electron pair or inject into gate oxide.

LDD (graded doping) allows drain voltage to be dropped over larger distance, thus reducing peak electric field/hot electron effect.

Mask #7 protects the PMOS devices.

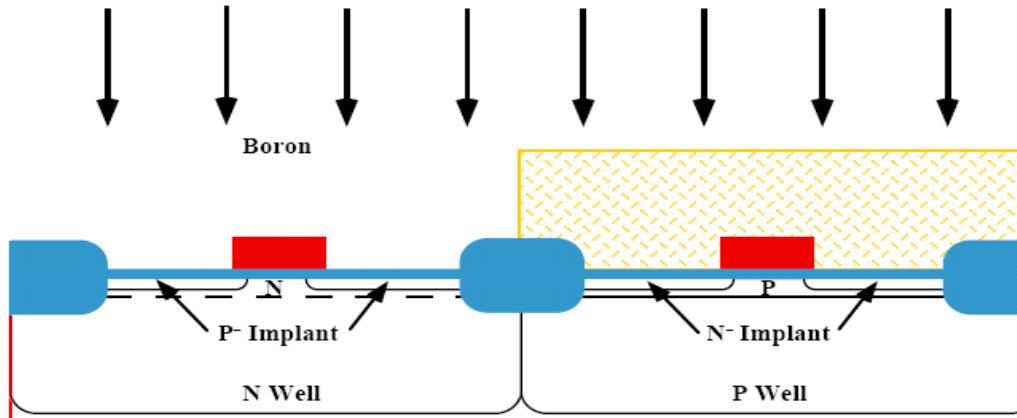
A P⁺ implant forms the LDD regions in the NMOS devices.

Typically $5 \times 10^{13} \text{ cm}^{-2}$ @ 50 KeV.

The polysilicon gate acts like a barrier for this implant to protect the channel region. This is thus a self-aligned process.

Tip or extension/LDD (lightly doped drain) formation

Remove resist, spin resist, photolithography, B⁺ ion implantation



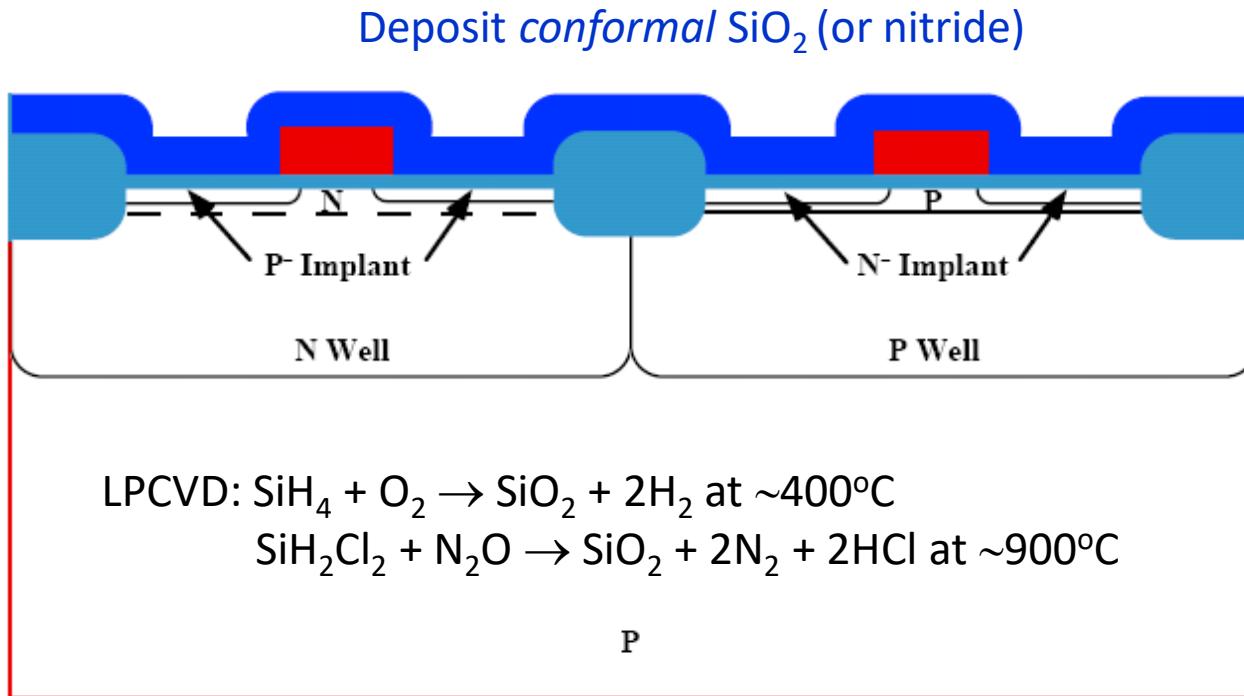
As LDD is shallower than N⁺ and P⁺ source/drain doping, it also reduces “short channel effect” due to the shallower channel.

Mask #8 protects the NMOS devices.

A B⁺ implant forms the LDD regions in the PMOS devices.

Typically $5 \times 10^{13} \text{ cm}^{-2}$ @ 50 KeV.

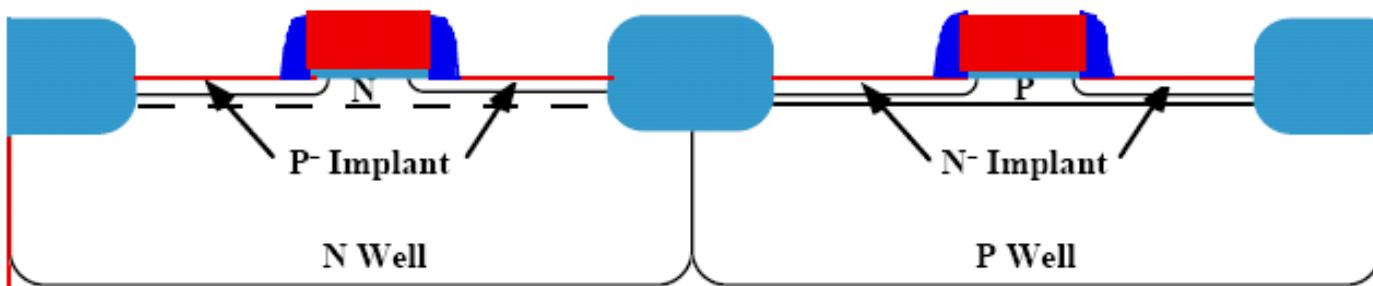
Sidewall spacer formation



Conformal layer of SiO_2 is deposited using LPCVD (typically 0.5 μm).

Sidewall spacer formation

Selective anisotropic dry etching



This works because deposition is conformal (*isotropic*), whereas etching is *anisotropic* (faster etching along vertical direction, little along horizontal direction).

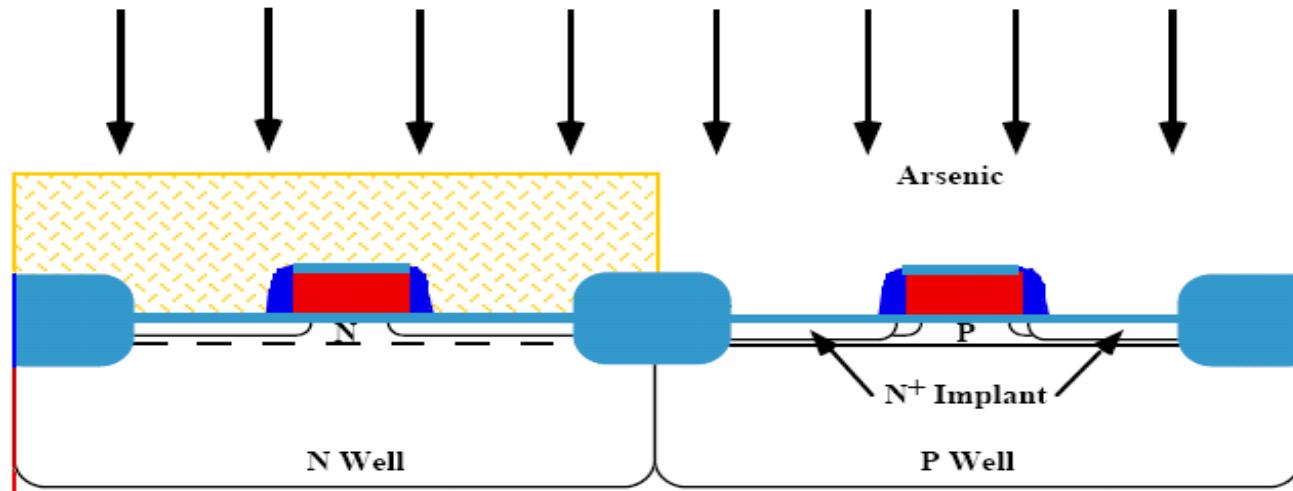
P

Anisotropic etching leaves “sidewall spacers” along the edges of the poly gates.

Timed etch of oxide/nitride using very directional etch (RIE).
Just enough time to remove oxide from the source, drain and gate regions.

Source/drain formation for NMOS

Thermal oxidation, spin resist, photolithography, As⁺ ion implantation



Grow a thin oxide to reduce “channeling effect” during ion implantation, as well as protect the Si surface from contaminants.

As atom has low diffusion coefficient, good for shallow junction.

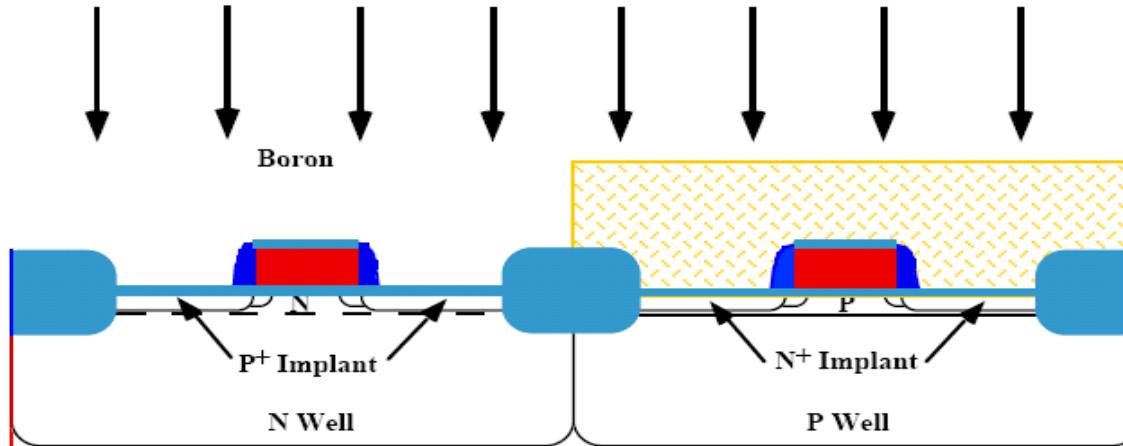
Mask #9 protects the PMOS devices.

An As⁺ implant forms the NMOS source and drain regions.

Typically $2-4 \times 10^{15} \text{ cm}^{-2}$ @ 75 KeV.

Source/drain formation for PMOS

Strip resist, spin resist, photolithography, B⁺ ion implantation



P

Figure 2-32

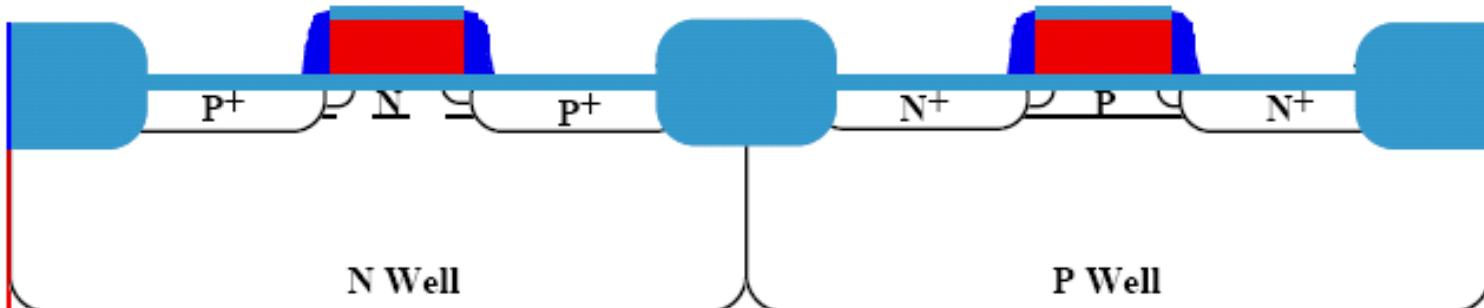
Mask #10 protects the NMOS devices.

A B⁺ implant forms the PMOS source and drain regions.

Typically $1-3 \times 10^{15} \text{ cm}^{-2}$ @ 50 KeV.

Drive-in anneal

Remove resist and anneal (diffusion, damage repair and dopant activation)



Anneal is always needed after ion implantation to repair Si lattice damage caused by energetic ion, and active dopant (bring it to crystalline sites).

P

A final high temperature anneal drives-in the junctions and repairs implant damage.

Typically 30 min @ 900°C or 1 min RTA @ 1000°C.

(RTA: rapid thermal annealing)

Contact and local interconnect formation

Etch away oxide, deposit Ti

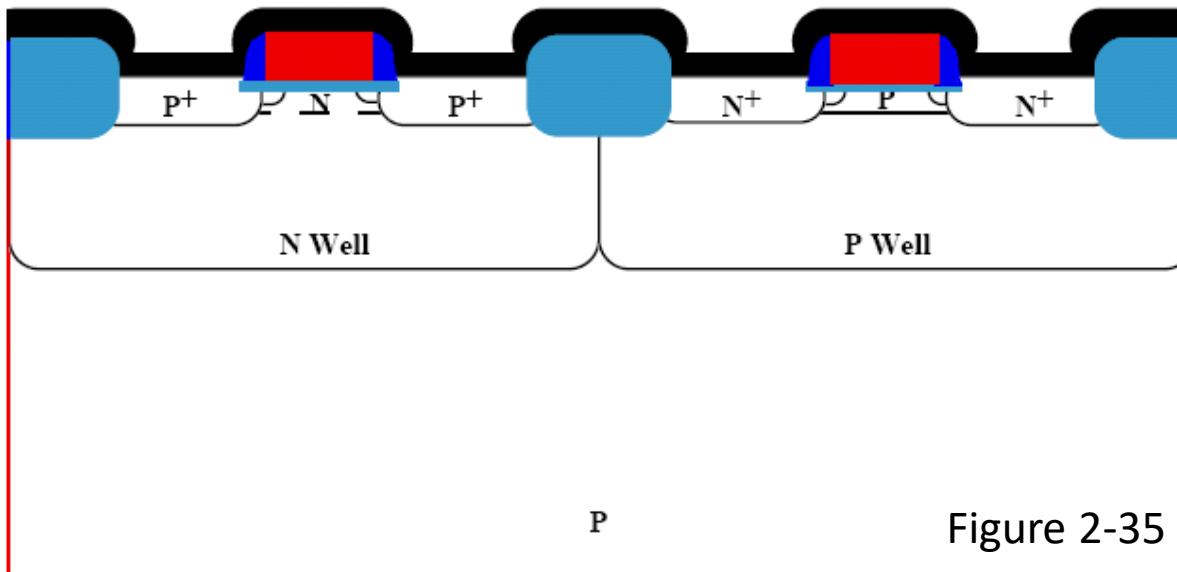
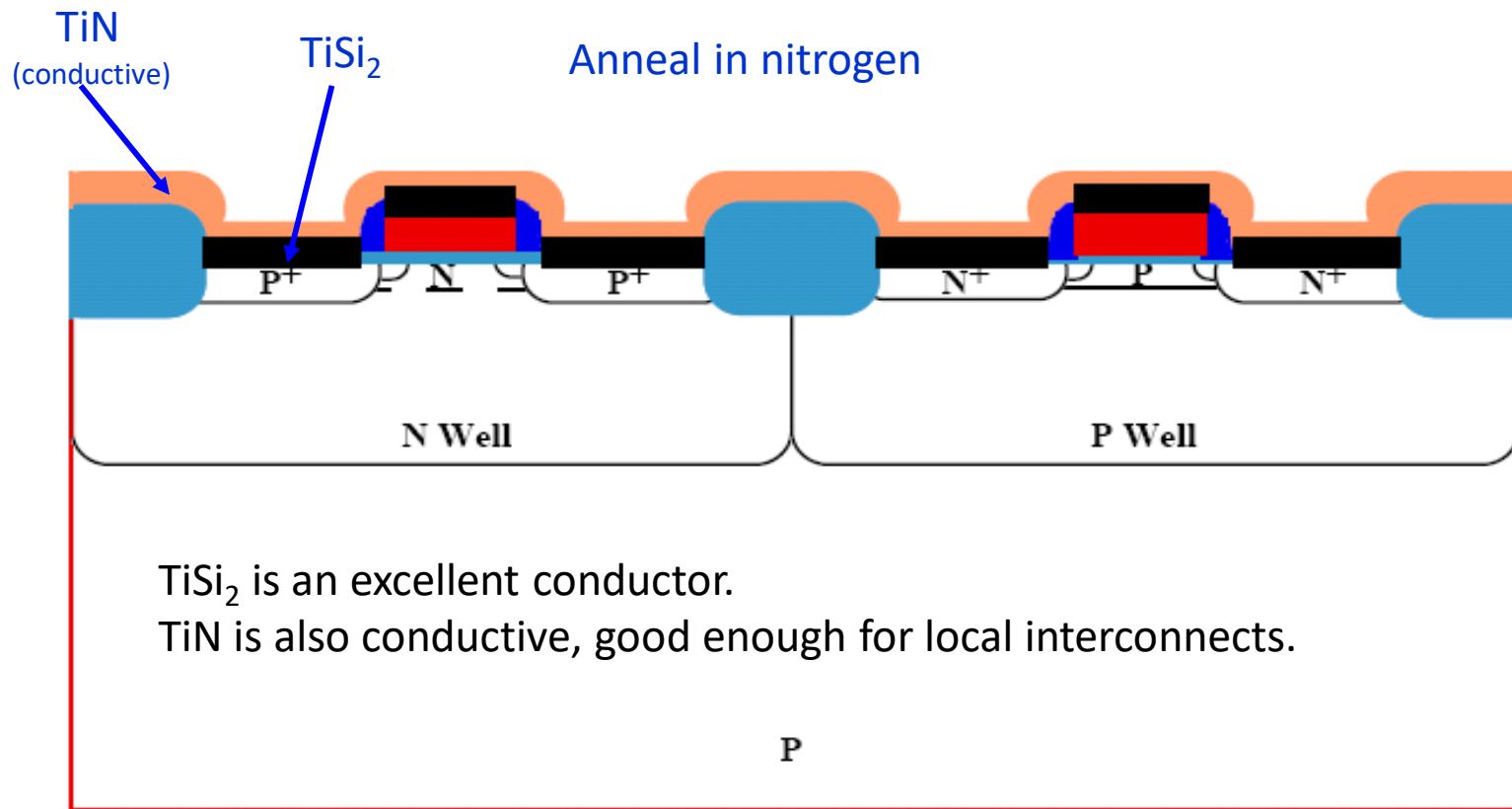


Figure 2-35

An unmasked oxide etch allows contacts to Si and poly regions.
Ti is deposited by sputtering (typically 100nm).

Contact and local interconnect formation



The Ti is reacted in an N₂ ambient, forming TiSi₂ and TiN.

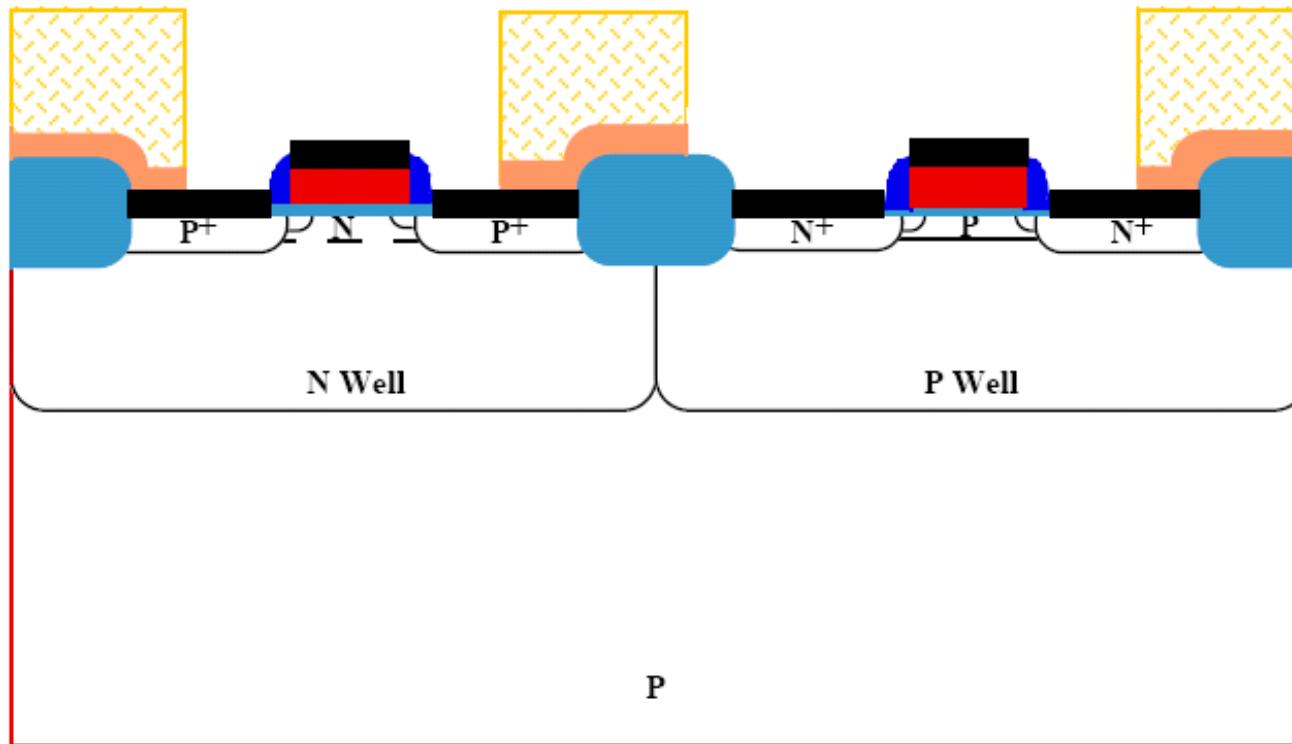
Typically RTP 1 min @ 600 - 700°C. (RTP: rapid thermal processing)

This process is called self-aligned silicide (salicide), since TiSi₂ is formed only on the Si surface at source, drain and gate. TiN is formed at SiO₂ surface.

Salicide reduces gate resistance, source/drain contact resistance.

Contact and local interconnect formation

Spin resist, photolithography, TiN selective etching



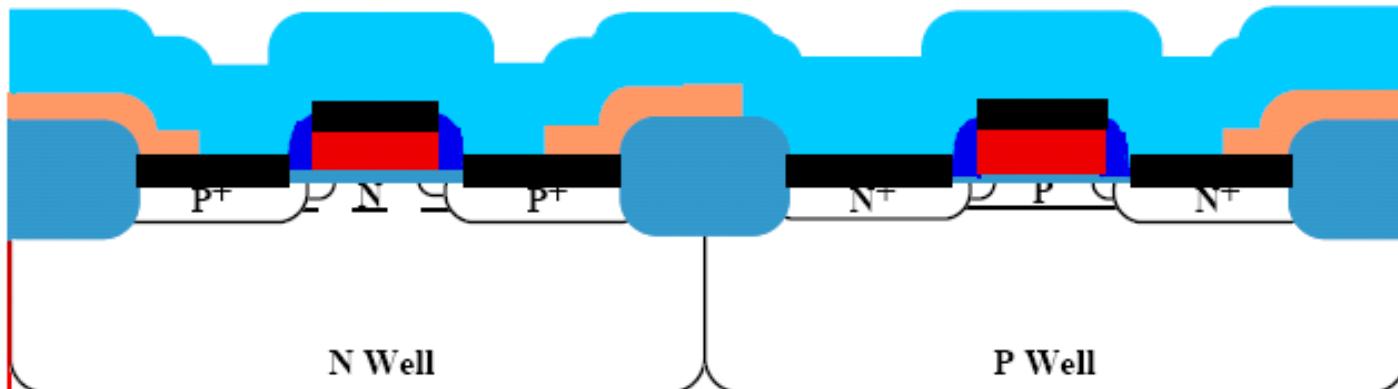
Mask #11 is used to etch the TiN, forming local interconnects.

TiN etched by $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (1:1:5)

Then an anneal at $\sim 800^\circ\text{C}$ in Ar for 1 minute to reduce the resistivity of TiN and TiSi_2 to their final values.

Multi-level metal formation

Remove resist, deposit SiO₂

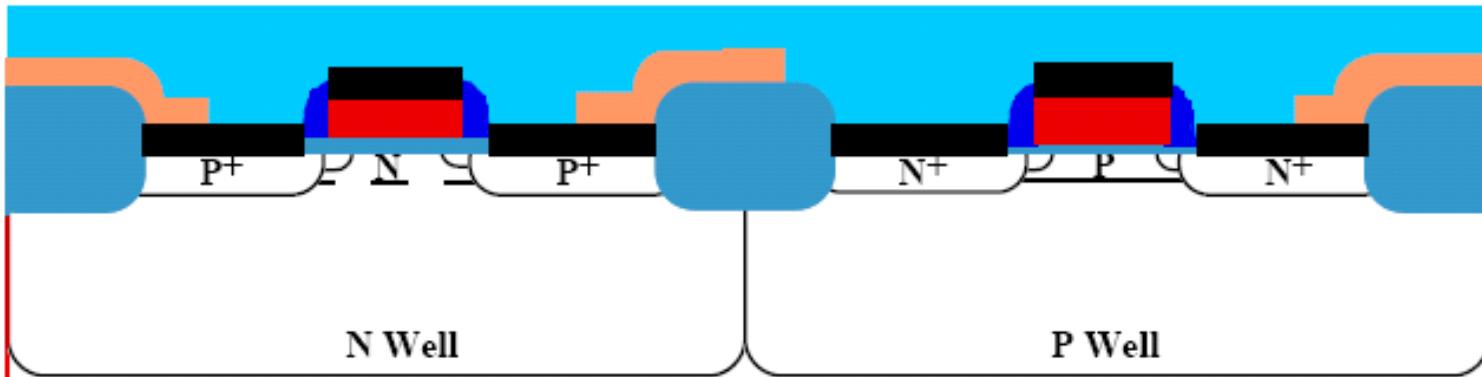


The SiO₂ layer is often doped with P (PSG – phosphosilicate glass) that protects the device against mobile ions like Na⁺. B may also be added (BPSG – borophosphosilicate glass) to reduce the flowing temperature of the glass (flow to smooth out the surface, good for planarization).

A conformal layer of SiO₂ is deposited by LPCVD (typically 1 μm).

Surface planarization

Chemical mechanical polishing (CMP)



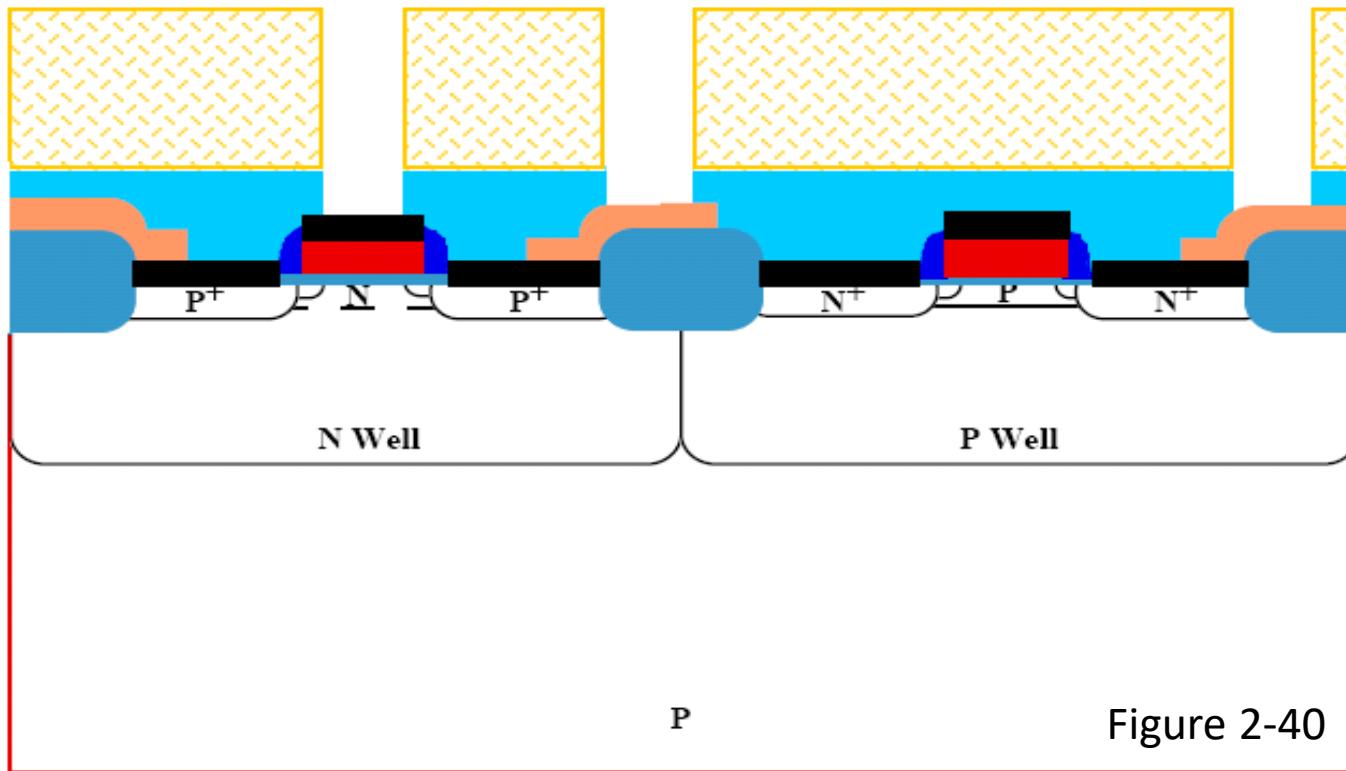
In CMP, besides mechanical polishing (by nanoparticles in the slurry), chemical reaction (e.g. by adjusting pH) is also important. The total polishing rate is much higher than mechanical polishing rate and chemical reaction rate alone.

P

Besides CMP, planarization can also be done by spinning resist and etching back, using a recipe where etching rates for resist and glass are the same.

Multi-level metal formation

Spin resist, photolithography, oxide etching

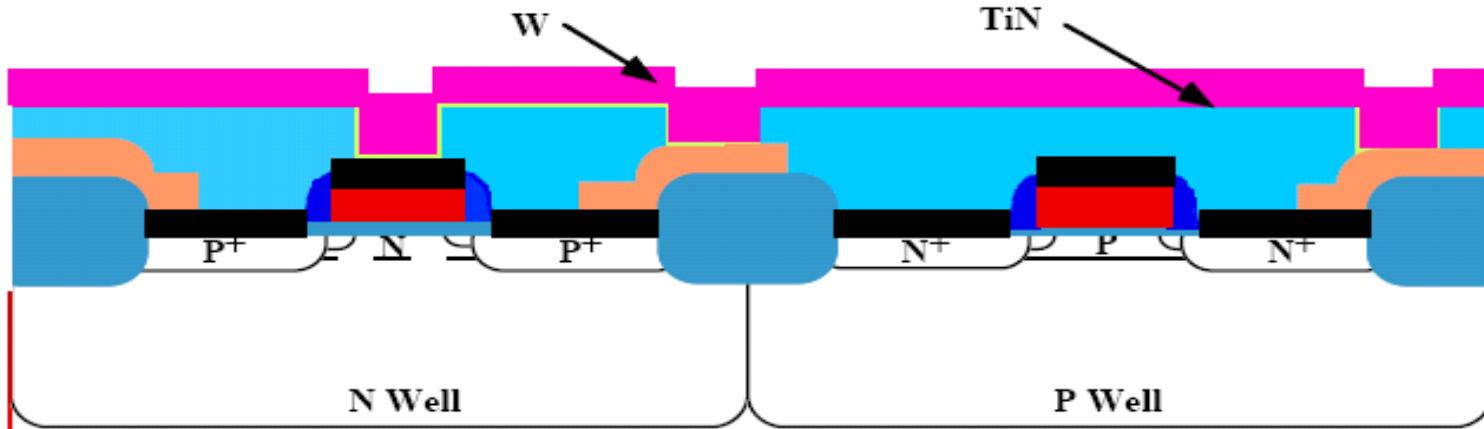


Mask #12 is used to define the contact holes.

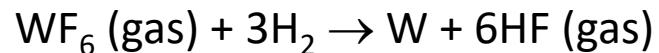
The SiO₂ is plasma etched (reactive ion etching).

W stud (via) formation

Remove resist, deposit TiN diffusion barrier/adhesion layer and W



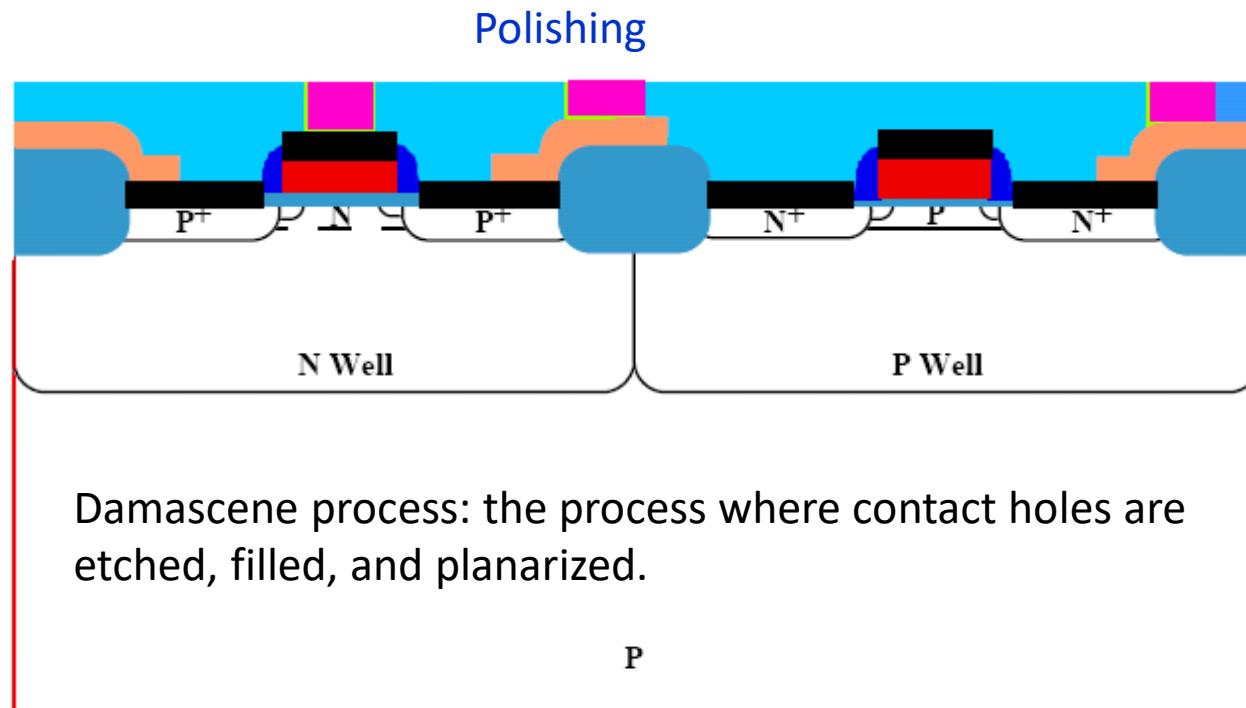
W CVD (chemical vapor deposition):



P

A thin TiN barrier layer is deposited by sputtering (typically a few tens of nm), followed by W CVD deposition.

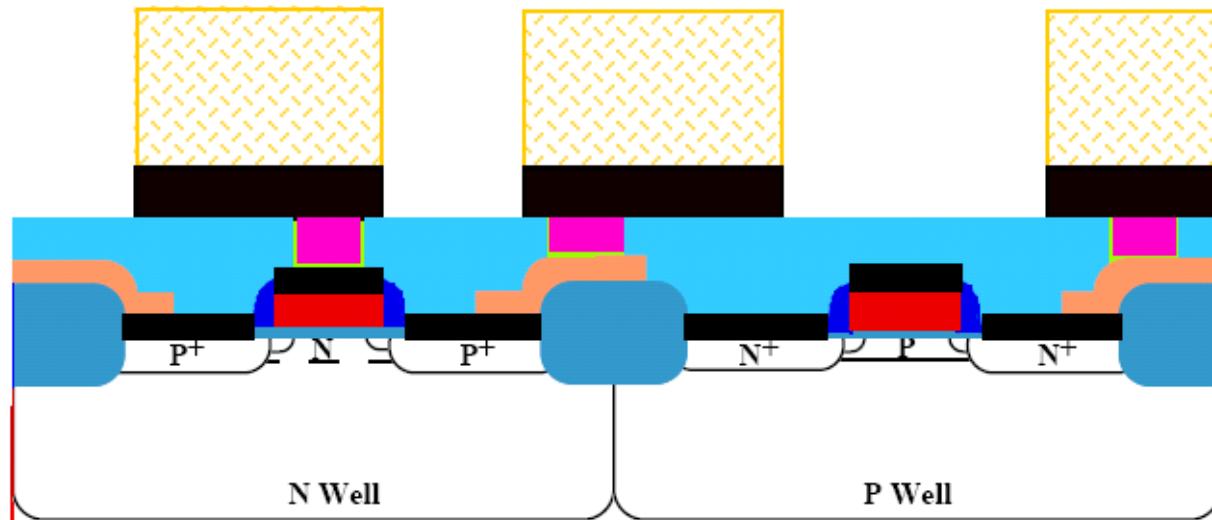
W stud (via) formation



CMP is used to planarize the wafer surface, completing the “damascene process”.

Multi-level metal formation

Deposit Al, spin resist, photolithography, selectively etch Al



Usually small percentage of Si and Cu is added to Al.

Add Si because Si is soluble in Al up to a few percent, and if not added now, Al will take/corrode Si from device region.

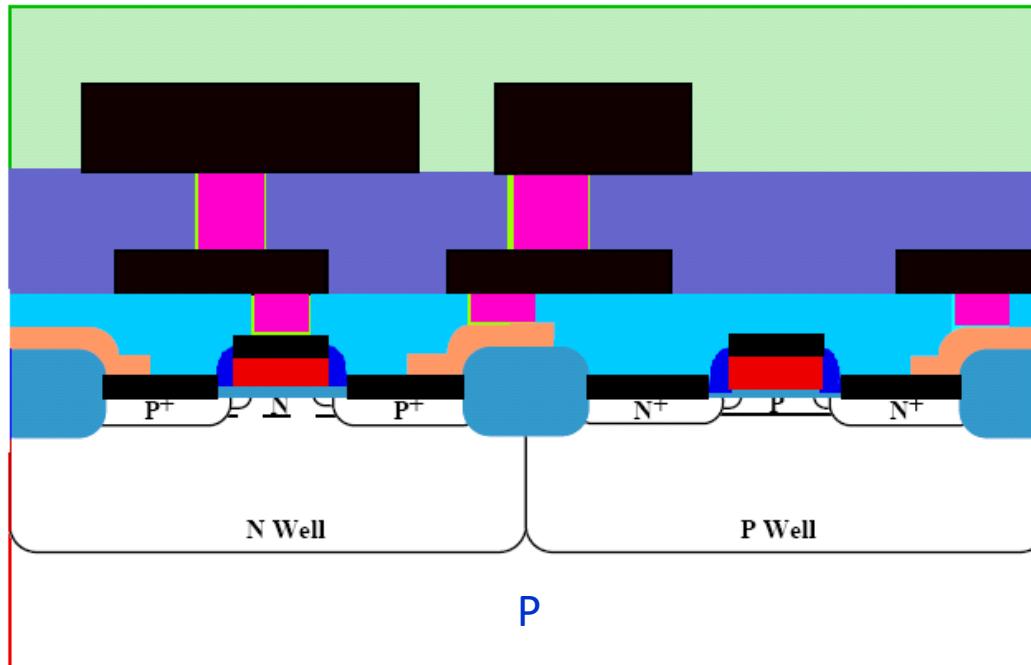
Add Cu to prevent electromigration in Al thin films (Al atoms move around, leaving behind voids)

Al is deposited on the wafer by sputtering.

Mask #13 is used to pattern the Al and plasma etching is used to etch it.
(Al is one of the few metals that can be etched by plasma)

Multi-level metal formation

Strip resist, deposit insulator layer, W stud and Al wire formation, Si_3N_4 passivation layer deposition.



Inter-metal dielectric and second level metal are deposited and defined in the same way as level #1.

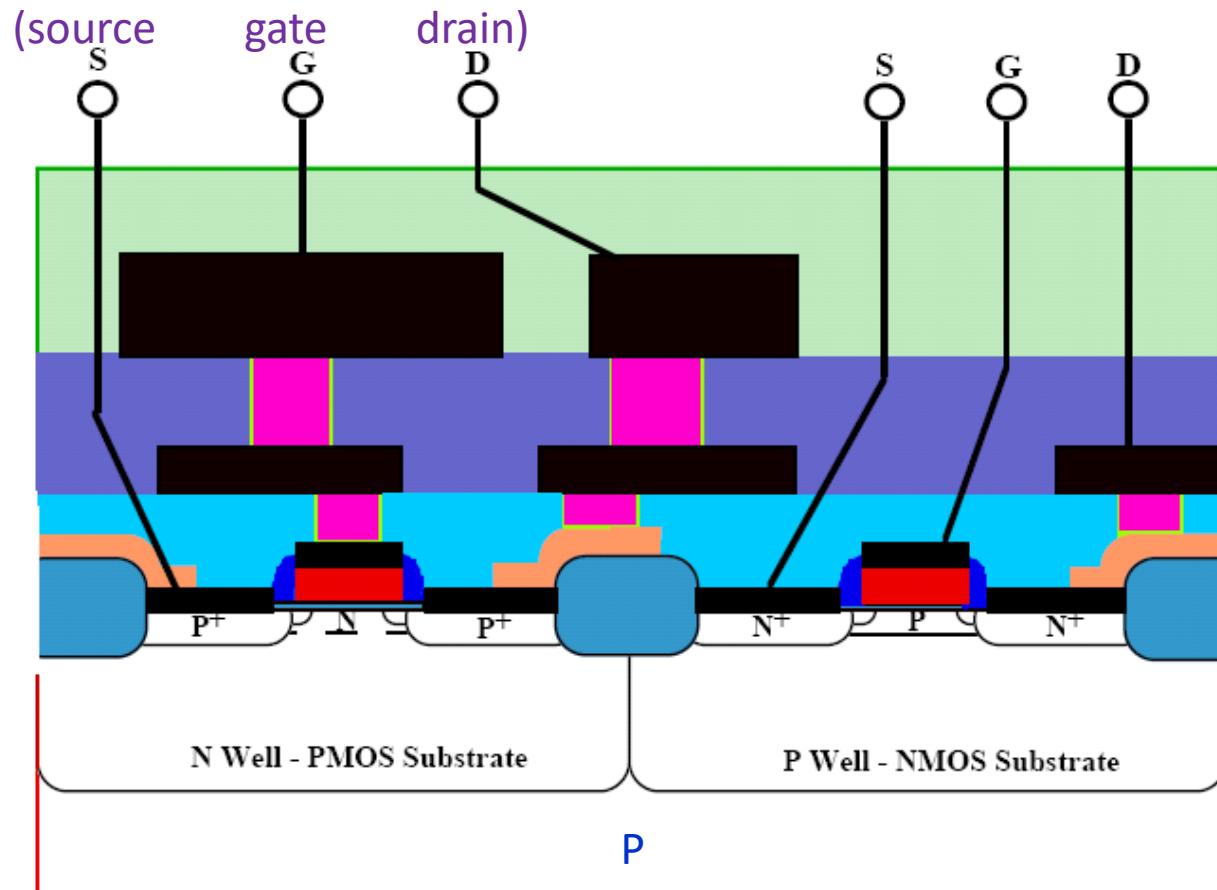
Mask #14 is used to define contact via-holes.

Mask #15 is used to define metal 2.

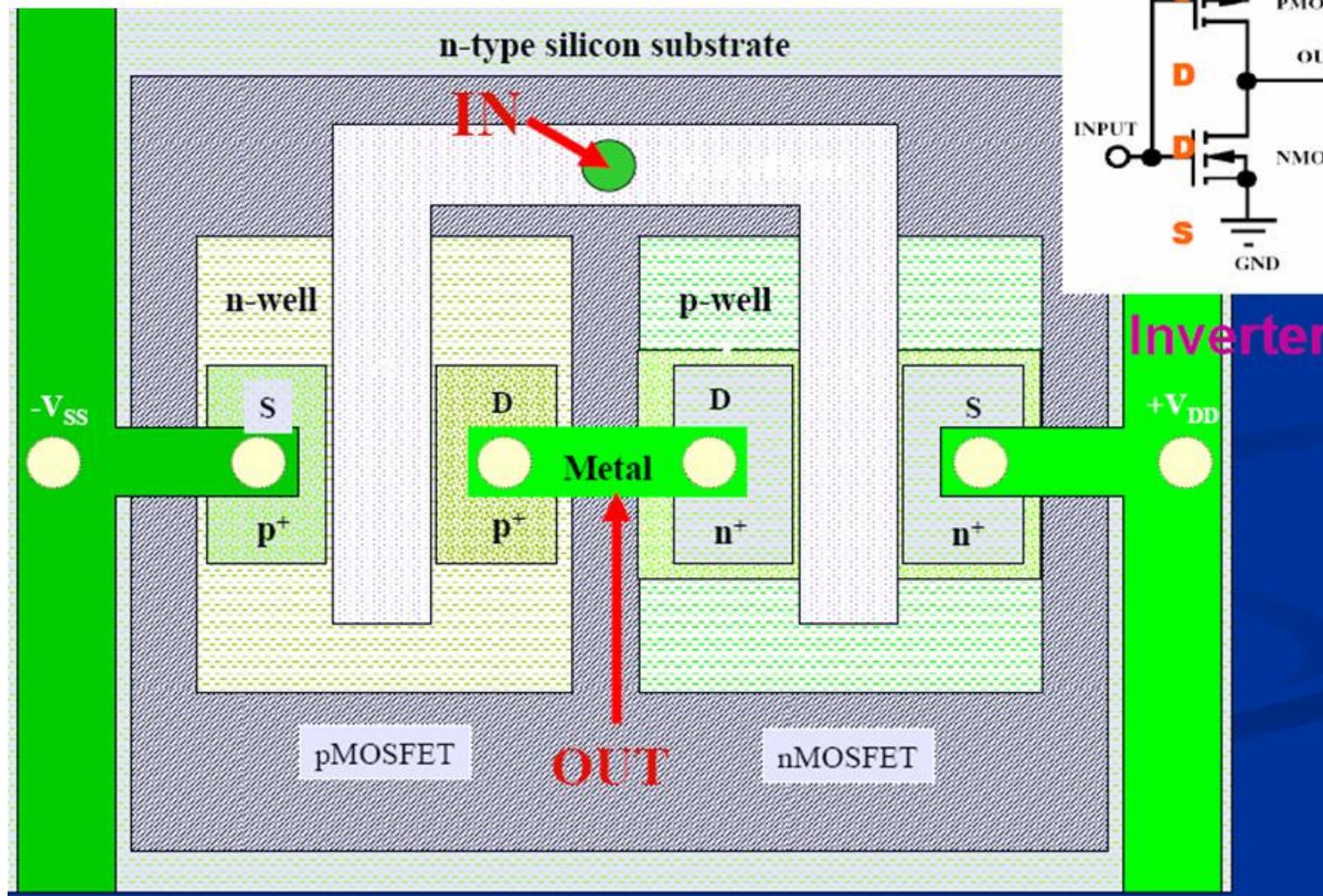
Passivation/protection layer of Si_3N_4 is deposited by PECVD and patterned with Mask #16. Final anneal (400-500°C, 30min, in forming gas – 10% H_2 in N_2) to alloy the metal contacts and reduce electrical charges in the Si/SiO_2 interfaces.

Finish the device

Wire bonding and packaging



Top view of an inverter



90 nm generation transistor and interconnect

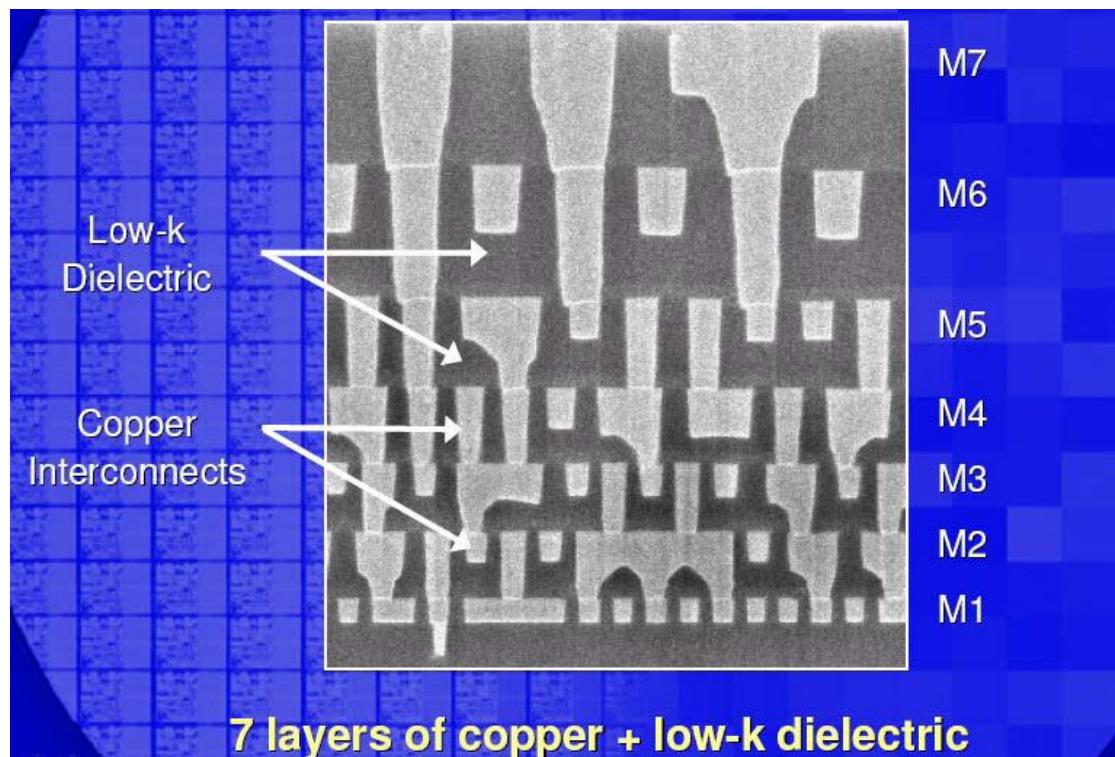
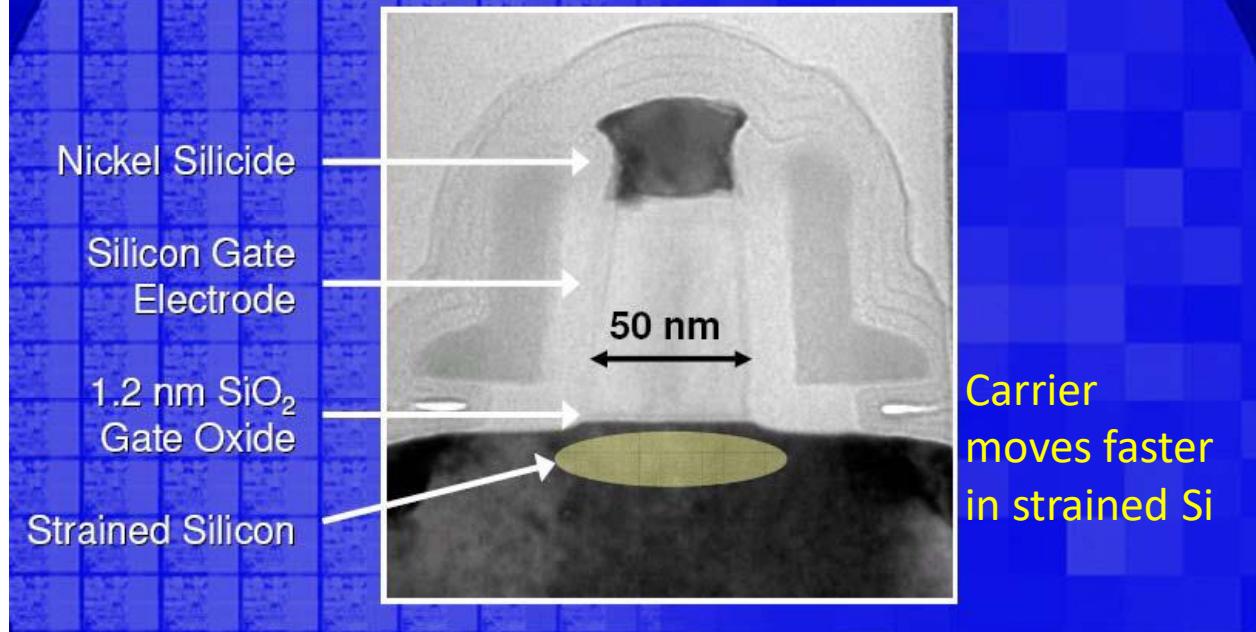
Ni silicide (not Ti silicide).

Only 1.2nm gate oxide.

Strained silicon.

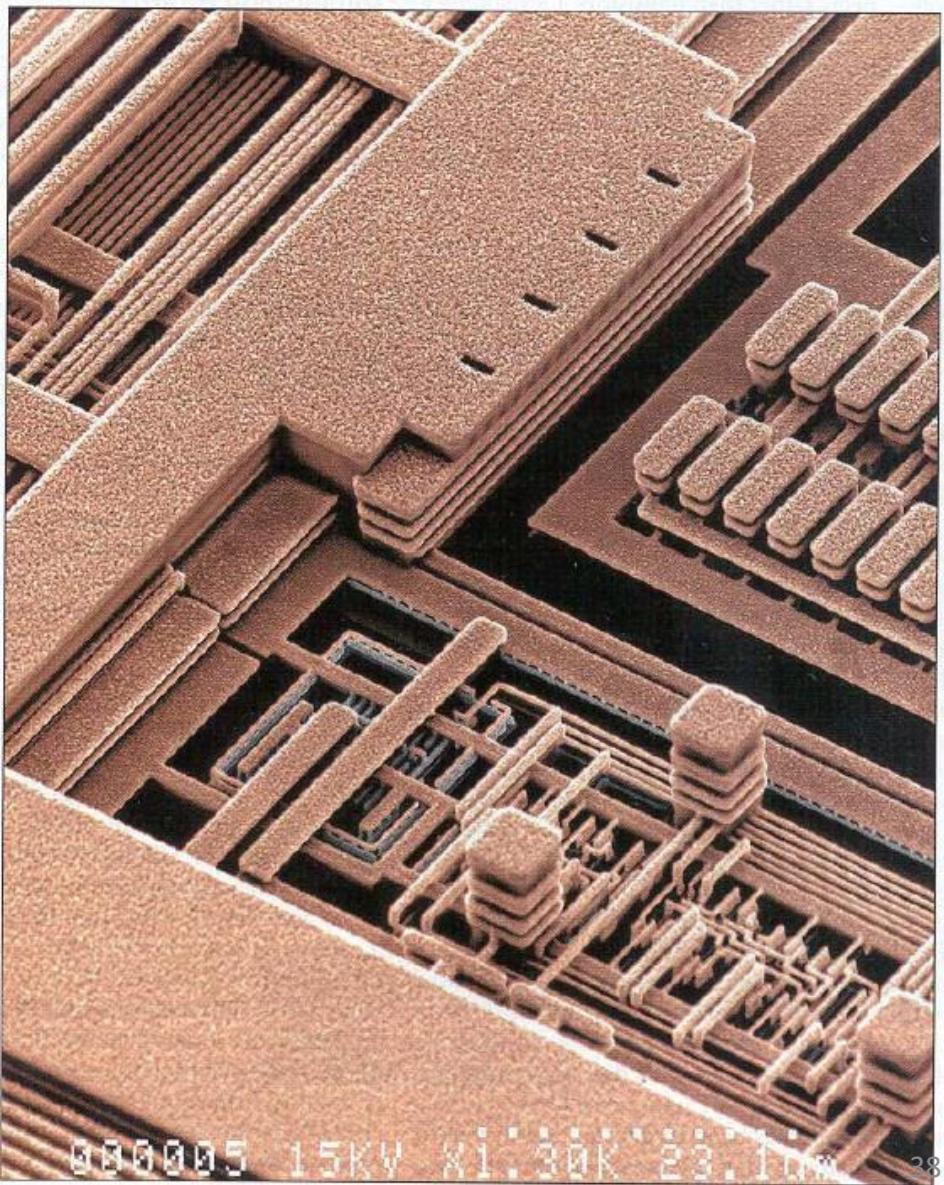
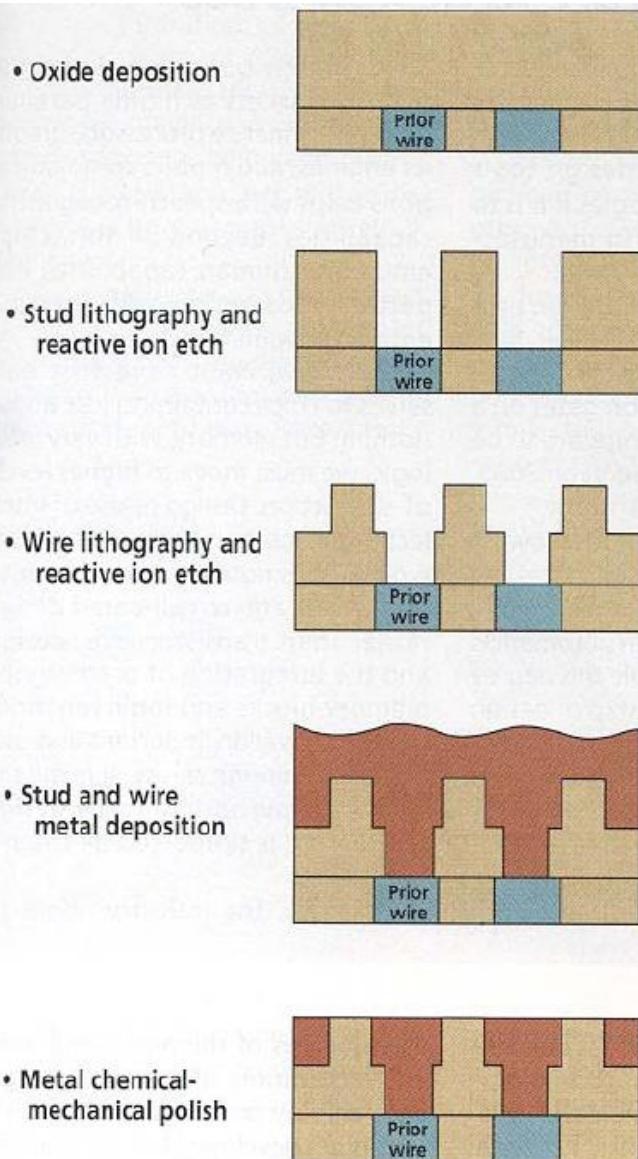
Low-k dielectric (lower ϵ , than SiO_2) to reduce capacitance and RC delay for faster circuit.

Copper interconnect (not Al) by electroplating and chemical mechanical polishing (see next slide).



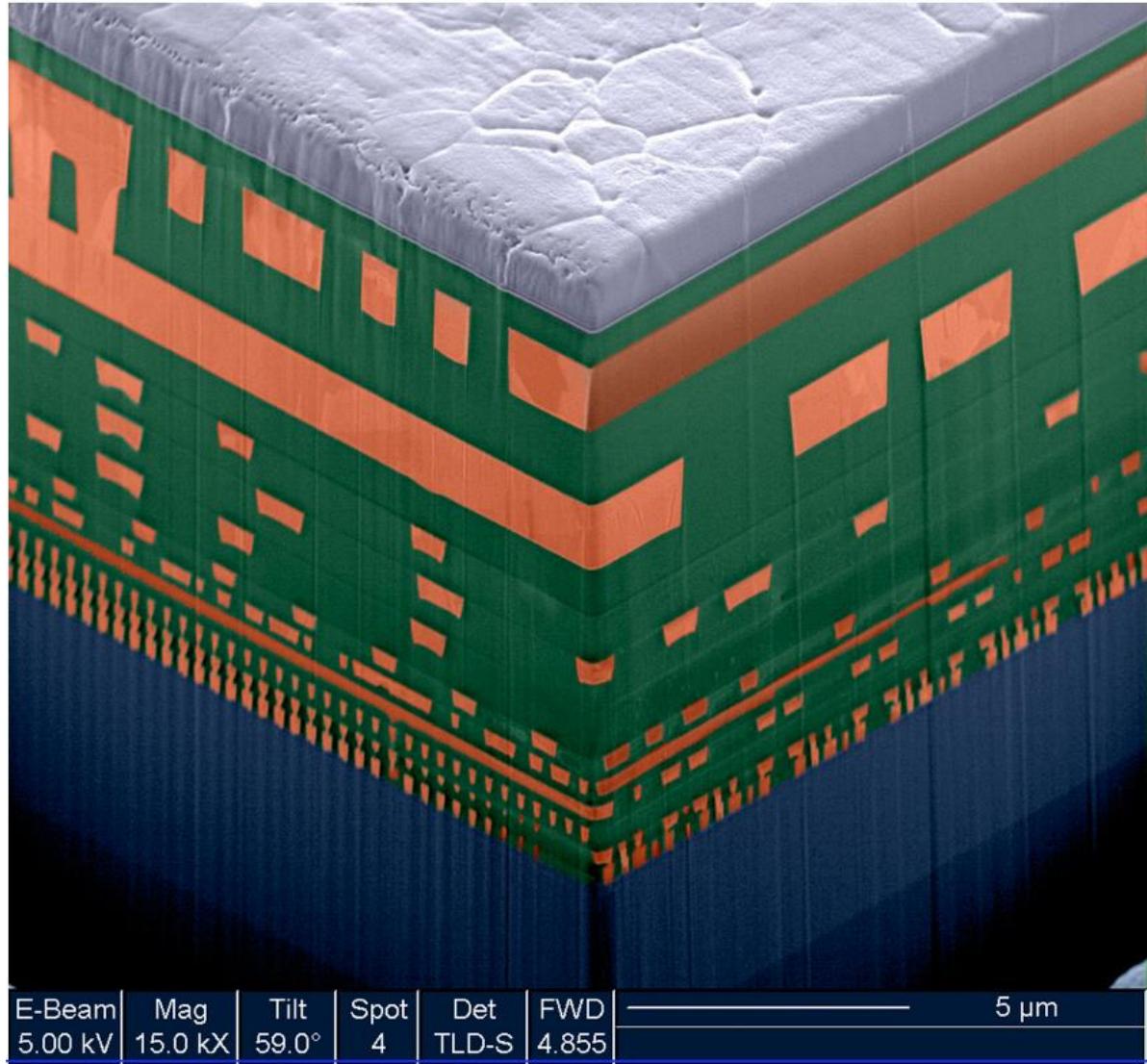
Advanced metallization: Cu based

Dual damascene IC process



CMOS interconnects

Modern processes use several levels of metal, separated by layers of deposited oxide or other low-k materials.



Chapter 3 Crystal growth, wafer fabrication and basic properties of silicon wafers

1. Silicon crystal structure and defects.
2. Czochralski single crystal growth.
3. Growth rate and dopant incorporation for CZ method.
4. Float zone single crystal growth and doping.
5. Wafer fabrication.
6. Measurement methods.

Single crystal silicon

- Three types of solids - amorphous, polycrystalline, mono-crystalline (single-crystal).
- Semiconductor devices and VLSI (very large scale integrated) circuits require high-purity single-crystal semiconductors. Because:
 - Difficult to control properties of amorphous or poly-crystals.
 - By doping, electronic properties (carrier density, mobility, conductivity, carrier lifetime) of a single crystal can be controlled more precisely.
- Amorphous silicon is used in photovoltaic cells, electronic displays (large-area).
- Polycrystalline silicon is used as a gate contact in MOSFETs (VLSI circuits).

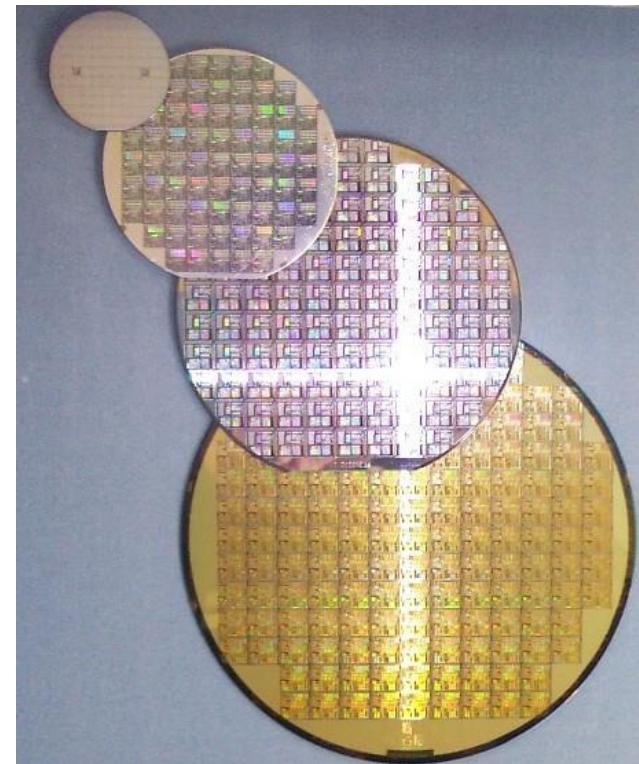
Single crystal Si wafers

Diameters: currently up to 300mm
(500mm?)

Wafer thickness: $\sim 650\mu\text{m}$

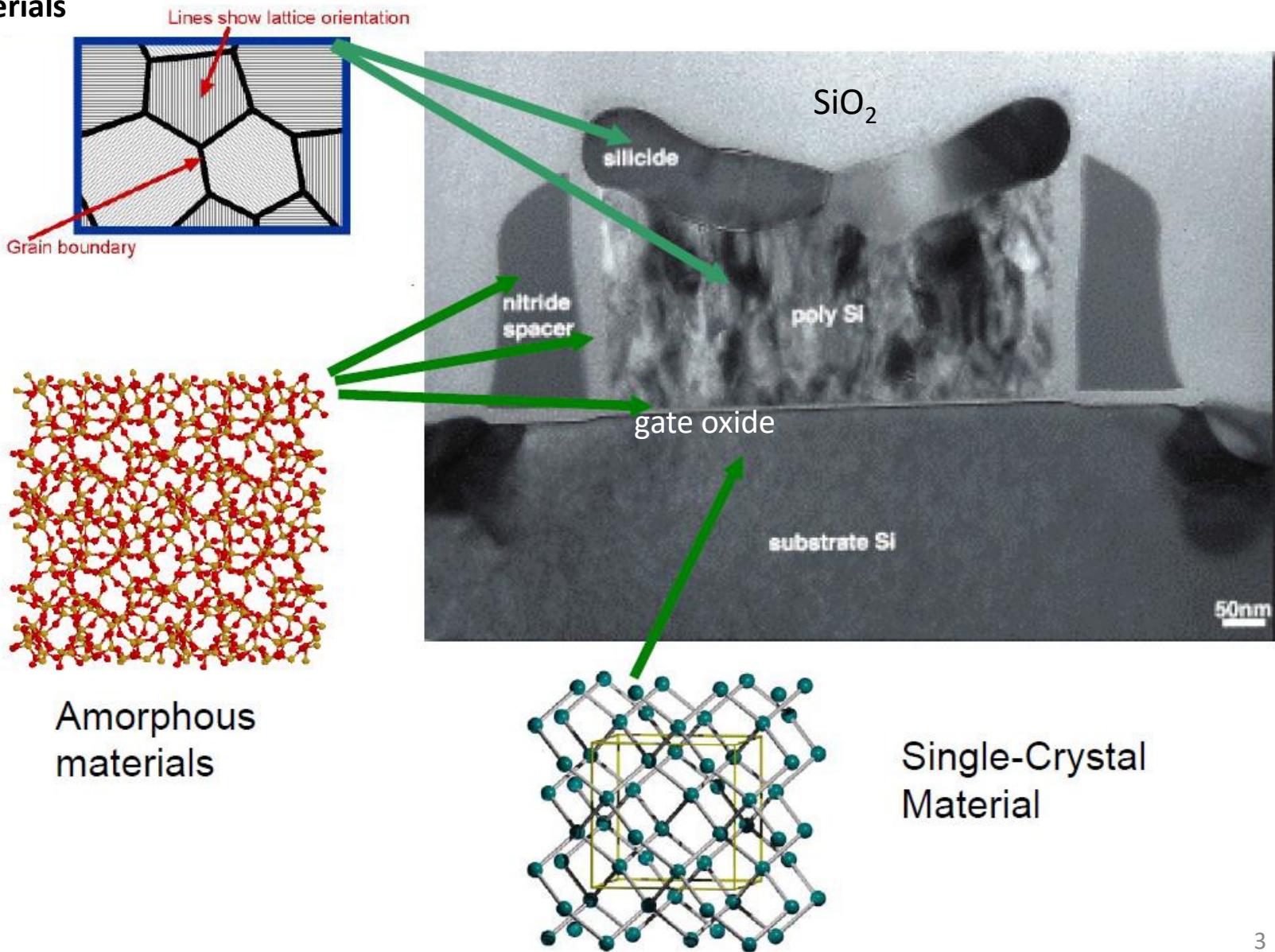
Wafer purity: 150 parts/trillion

Impurities: 99.9999999% Si



Microstructure of electronic materials

Polycrystalline materials



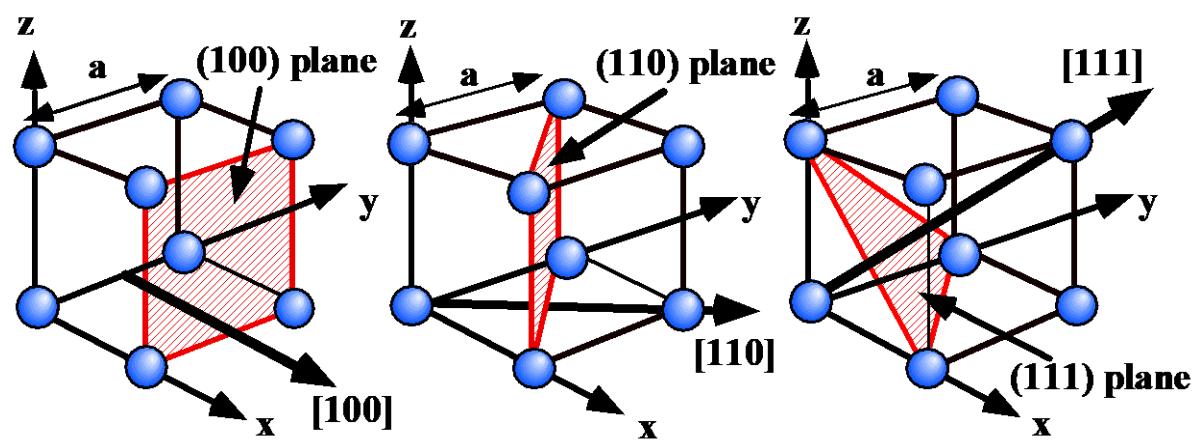
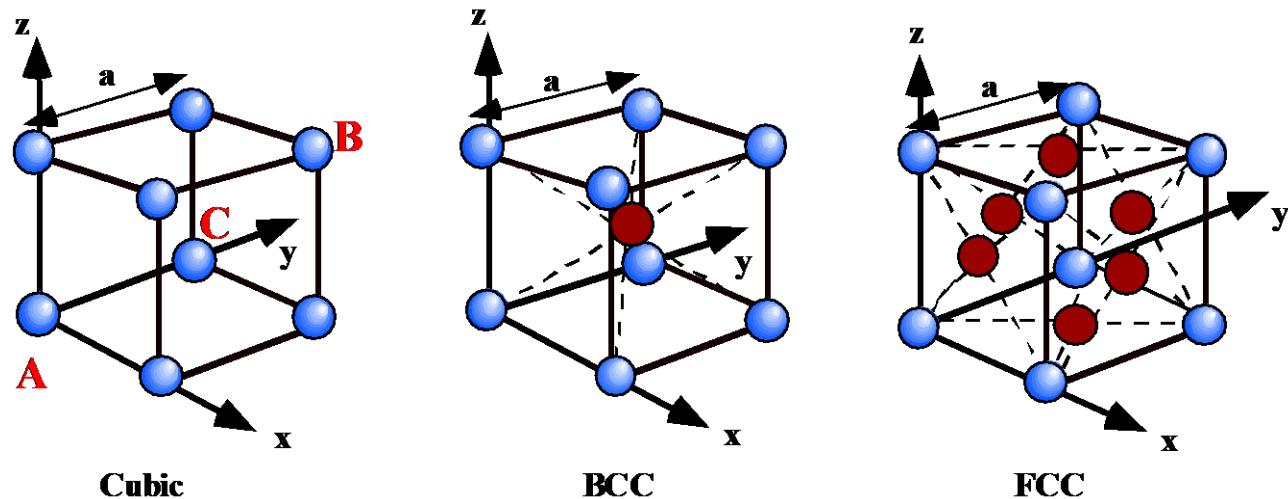
Crystallography - introduction

Crystals are characterized by a unit cell which repeats in the x, y, z directions.

Planes and directions are defined using an x, y, z coordinate system.

[111] direction is defined by a vector having components of 1 unit in x, y and z.

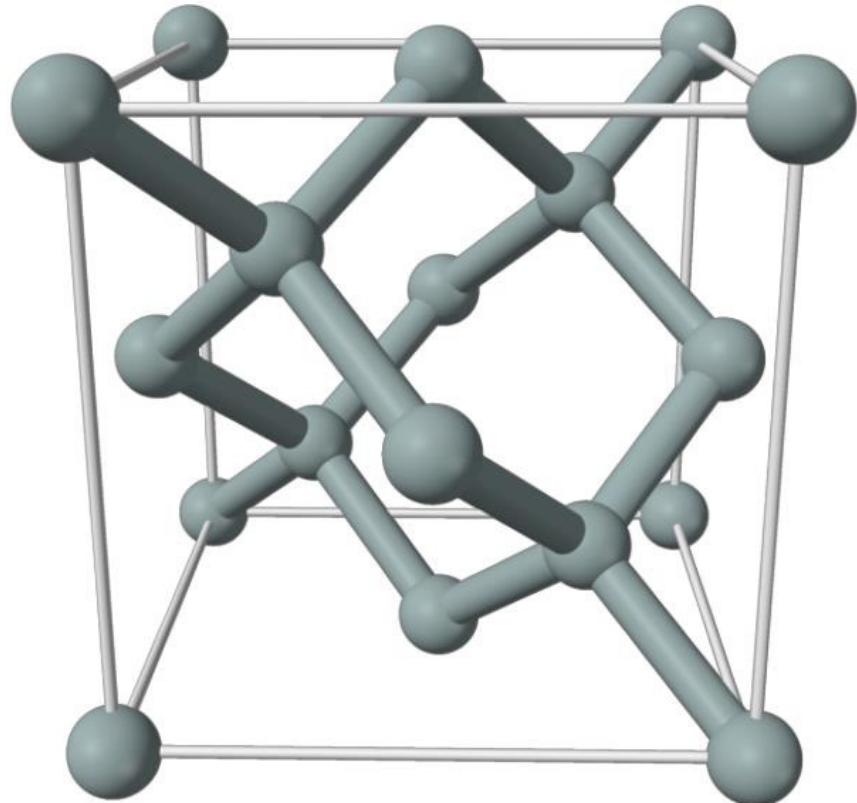
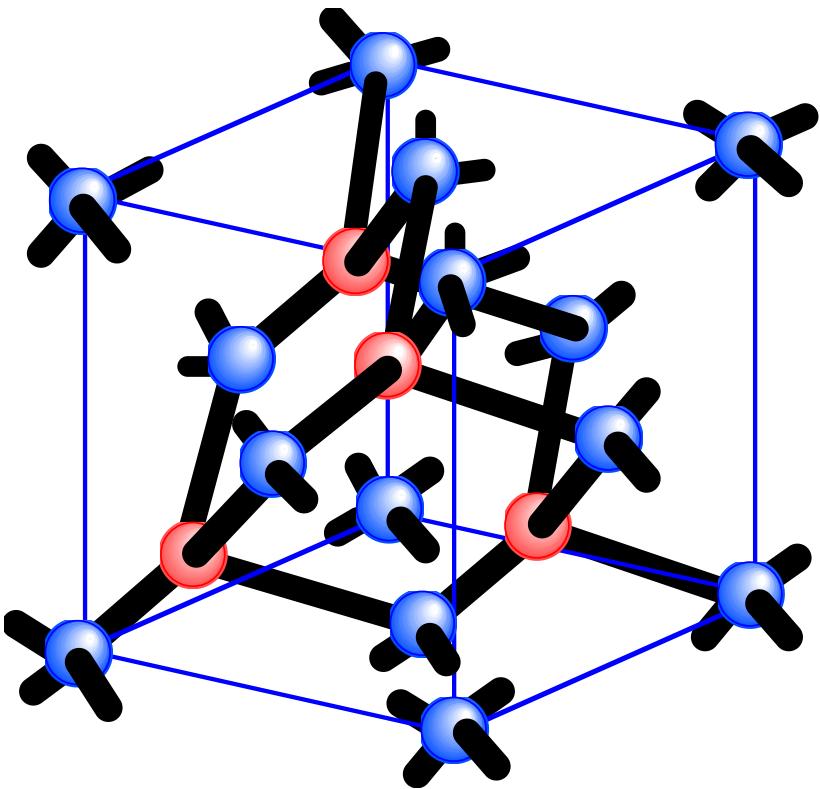
Planes are defined by Miller indices - *reciprocals* of the intercepts of the plane with the x, y and z axes.



For the (110) plane, it has intercepts with x, y, z-axis of 1, 1, ∞ (i.e. no intercept with z-axis). So its Miller indices are $(1/1, 1/1, 1/\infty) = (1,1,0)$.

For any plane (l, m, n) , it is always perpendicular to the direction $[l, m, n]$.
E.g. [111] direction is perpendicular to (111) plane.

Diamond Structure of Si



Silicon has the basic diamond crystal structure:
Two merged FCC cells offset by $a/4$ in x, y and z.

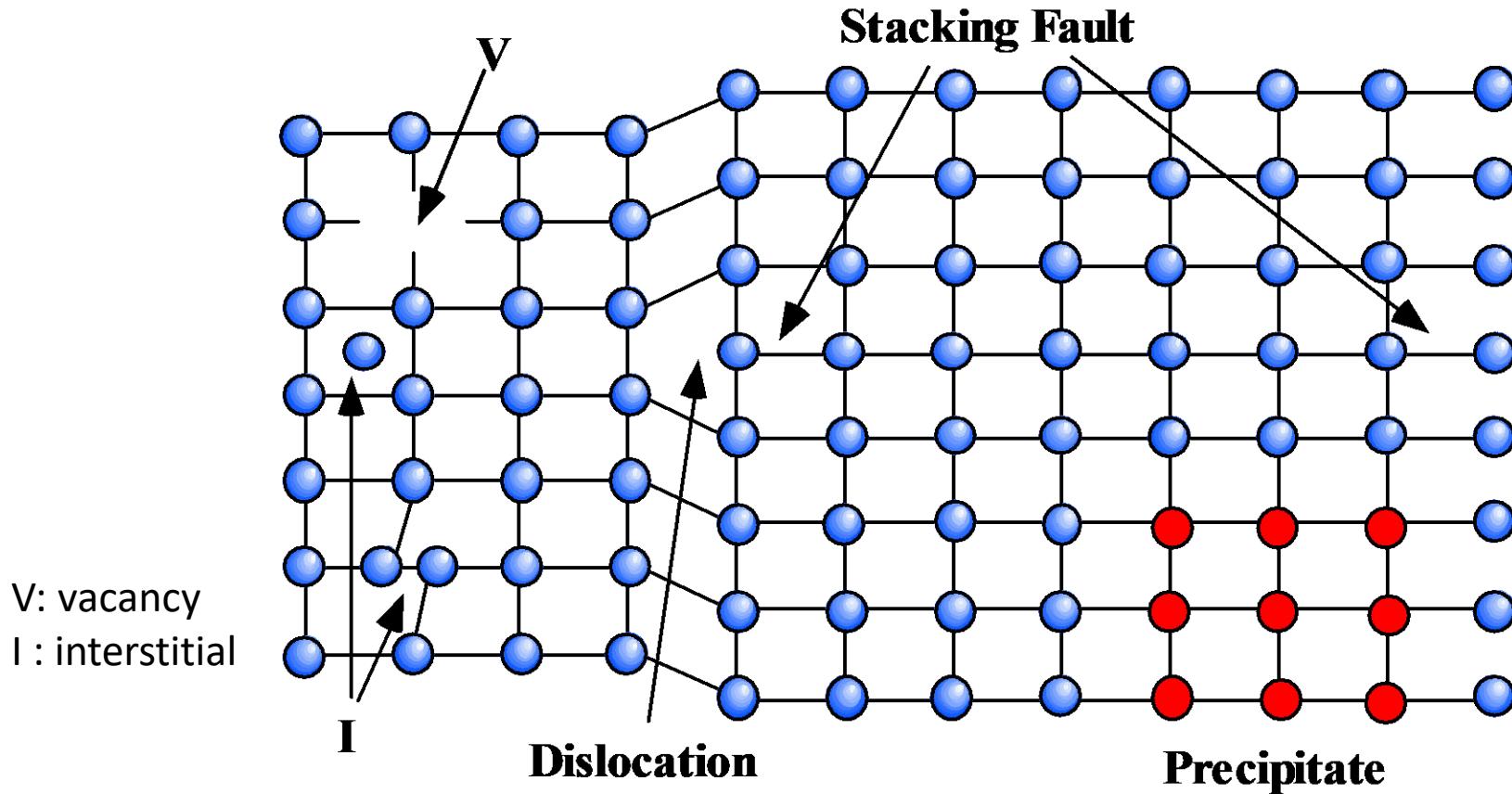
Defect in crystals

Point defects: vacancy, interstitial, substitutional, Frankel defects

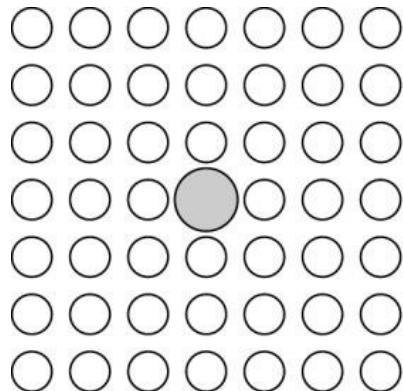
Linear defects: edge and screw dislocations

Planar defects: stacking fault, grain boundaries, surfaces

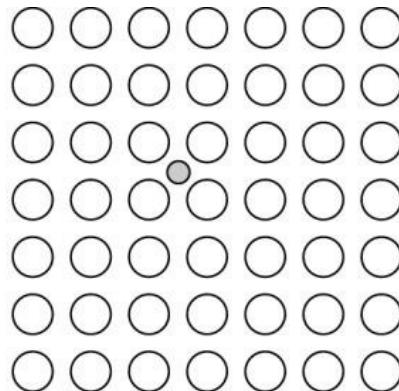
Bulk defects: cracks and pores, precipitate



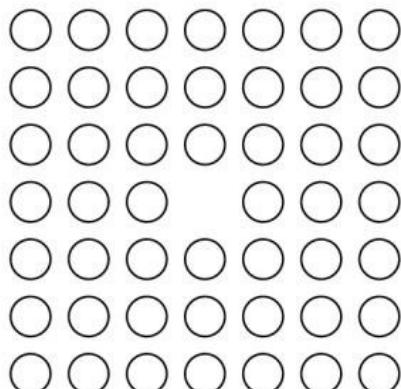
Point defects



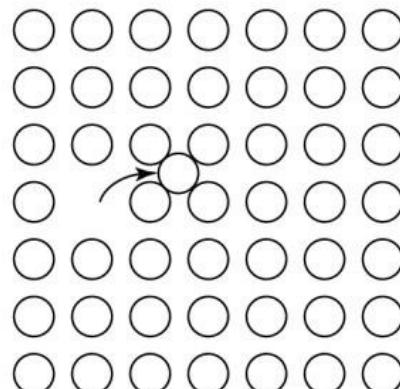
(a)



(b)



(c)



(d)

Point defects.

- a) Substitutional.
- b) Interstitial.
- c) Vacancy (Schottky defect).
- d) Frenkel-type defect
(interstitial - vacancy pair).

Point defects dictate most dopant diffusion mechanisms, and thus determine the impurity profile.

Chapter 3 Crystal growth, wafer fabrication and basic properties of silicon wafers

1. Silicon crystal structure and defects.
2. Czochralski single crystal growth.
3. Growth rate and dopant incorporation for CZ method.
4. Float zone single crystal growth and doping.
5. Wafer fabrication.
6. Measurement methods.

Raw material preparation



Sand



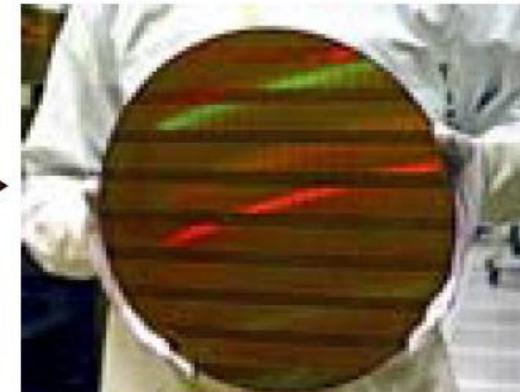
Metallurgical Grade Silicon
(MGS)



Electronics Grade Silicon
(EGS)

Economical value

From sand to ultrapure Si single crystal wafers to CPUs



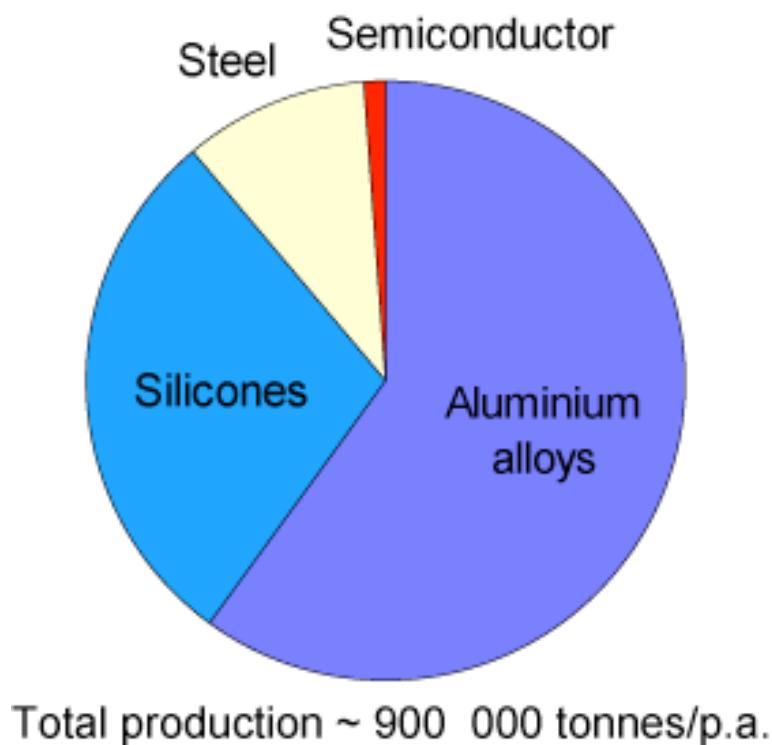
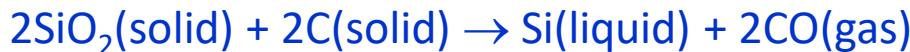
\$35 / ton

\$200 / 300mm wafer

\$300 / CPU

Metallurgical grade silicon

Quartzite (sand, SiO_2) is placed in a hot ($\sim 1800^\circ\text{C}$) furnace with carbon releasing materials, and reacts as shown, forming metallurgic grade silicon (MGS):



Element	Concentration (ppma)
Al	1200-4000
B	37-45
P	27-30
Ca	590
Cr	50-140
Cu	24-90
Fe	1600-3000
Mn	70-80
Mo	<10
Ni	40-80
Ti	150-200
V	100-200
Zr	30

Metallurgical grade silicon (~98% pure) production and typical impurity levels.

Over 50% MGS is used to make Al alloys.

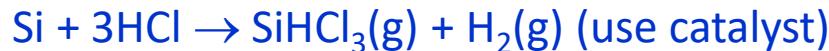
The fraction used for semiconductors is very small.

Electronic grade silicon (polycrystalline)

Basically, the solid Si is first converted into a liquid form (SiHCl_3) for purification, then converted back into solid Si.

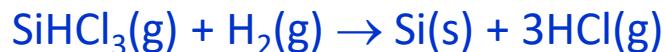
Both reactions occur at high temperatures.

Metallurgical grade silicon is treated with hydrogen chloride to form trichlorosilane:



SiHCl_3 is liquid *at room temperature*, boiling point 32°C. Multiple distillation of the liquid removes the unwanted impurities (99.9999% pure).

The purified SiHCl_3 is then used in a hydrogen reduction reaction to prepare the electronic grade Si (EGS):



(this is the reverse reaction of the above reaction)

EGS is the raw material for Si single crystal production.

Si single crystal growth



Two methods used: Czochralski (CZ) and float zone (FZ).

Jan Czochralski (cho-HRAL-skee) (1885 - 1953) was a Polish chemist who invented the Czochralski process, which is used to grow single crystals and is used in the production of semiconductor wafers.

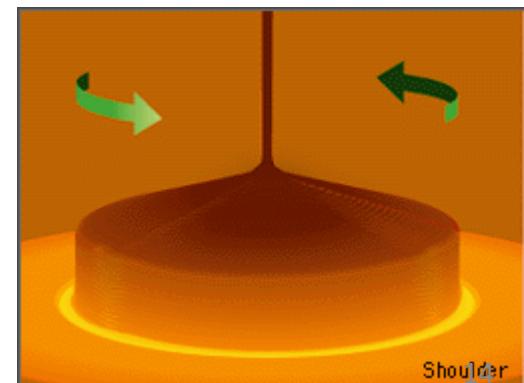
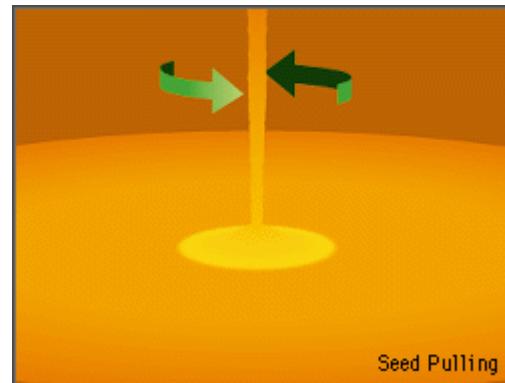
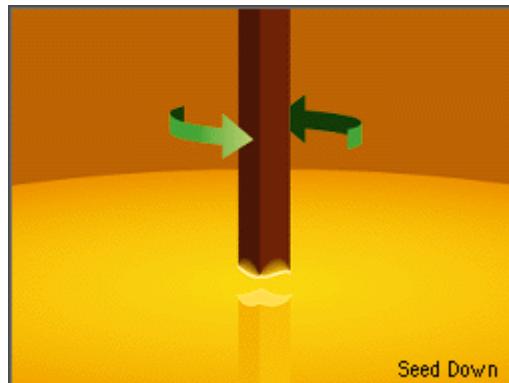
He discovered the Czochralski method in 1916 when he accidentally dipped his pen into a crucible of molten tin rather than his inkwell. He immediately pulled his pen out to discover that a thin thread of solidified metal was hanging from the nib. The nib was replaced by a capillary, and Czochralski verified that the crystallized metal was a single crystal.

http://en.wikipedia.org/wiki/Jan_Czochralski

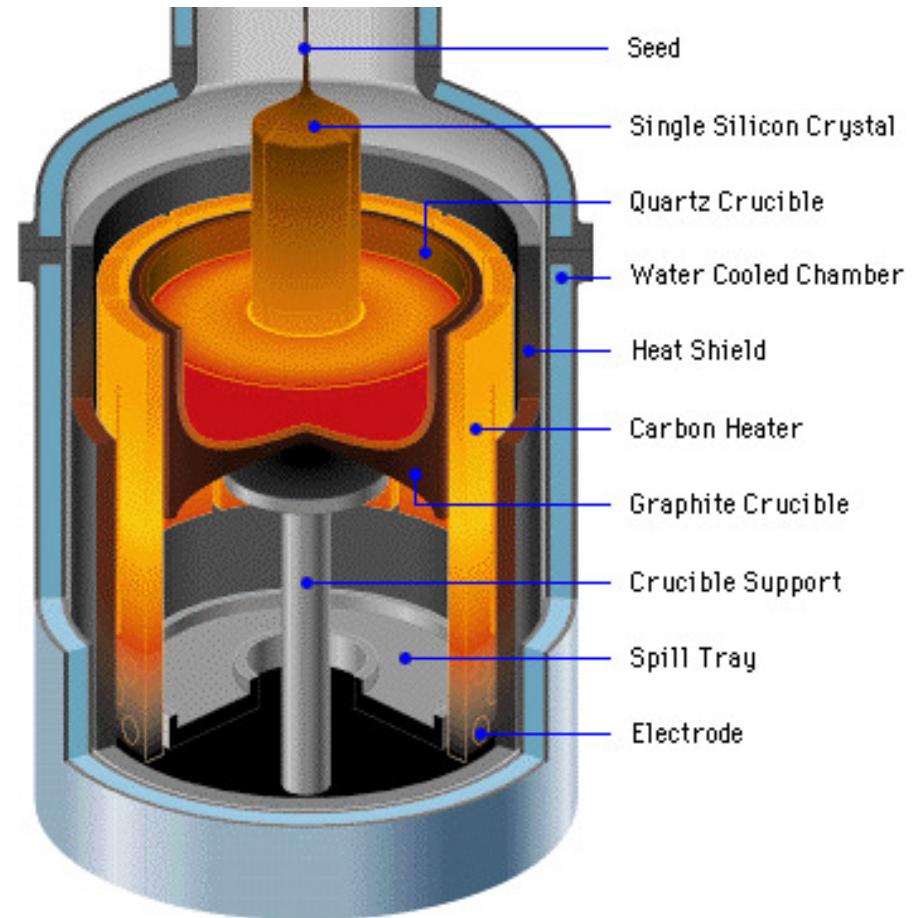
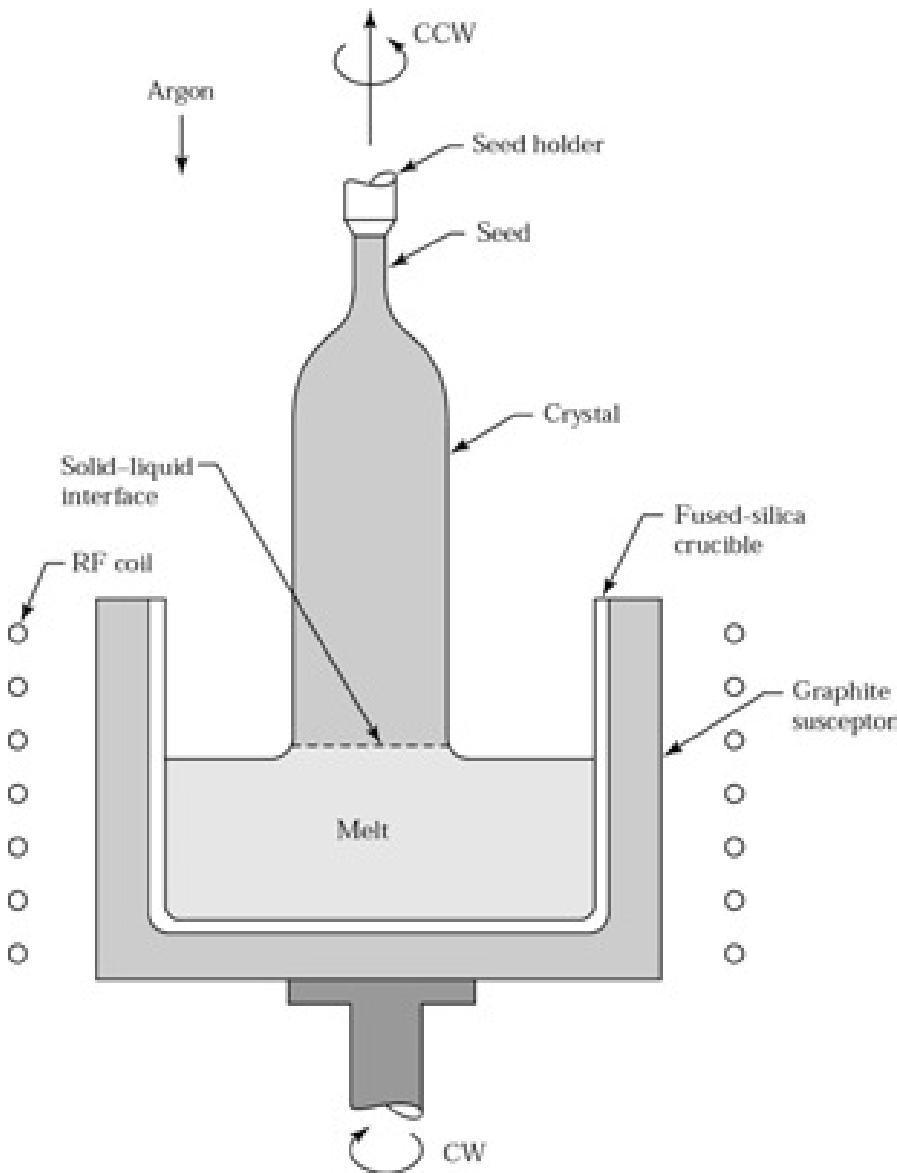


Czochralski method (CZ)

- It is widely employed for Si, GaAs, and InP.
- The EGS is broken into small pieces and placed in an SiO_2 crucible.
- In an argon ambient, the crucible is heated to just above 1417°C.
- Dopant is added to the melt to intentionally dope the resulting crystal.
- A single crystal seed is then lowered into the melt (crystal orientation and wafer diameter determined by seed orientation and pull rate), and withdrawn slowly.
- Melt flows up the seed and cools as crystal begins to grow.
- Seed rotated about its axis to produce a circular cross-section crystal. The rotation inhibits the natural tendency of the crystal to grow along certain orientations to produce a faceted crystal.
- Long ingots (boules) ~100 kg, with very good circular cross-section are produced.
- The oxygen and carbon (from graphite furnace components), contribute about 10^{17} - $10^{18}/\text{cm}^3$ contaminants.



Czochralski method (CZ)

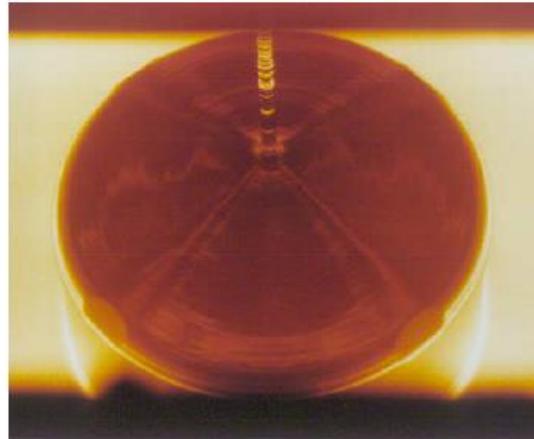


Czochralski growth of silicon

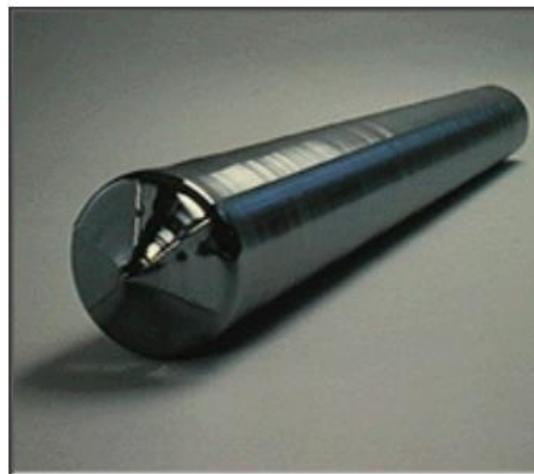
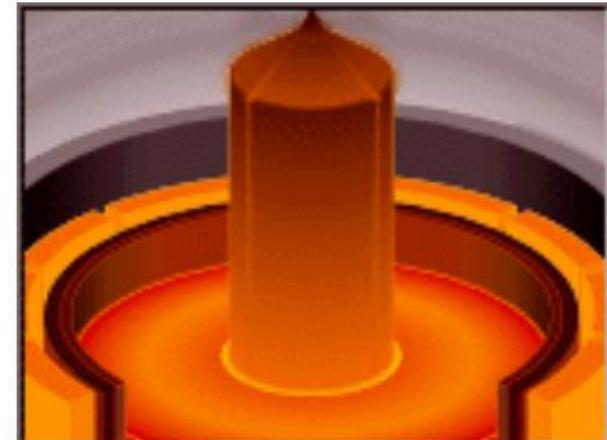
A commercial CZ puller



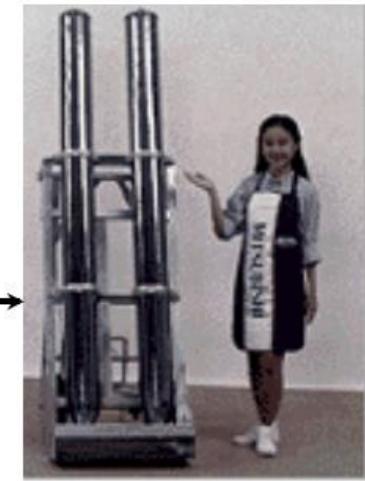
Early in the growth process



Later in the growth process

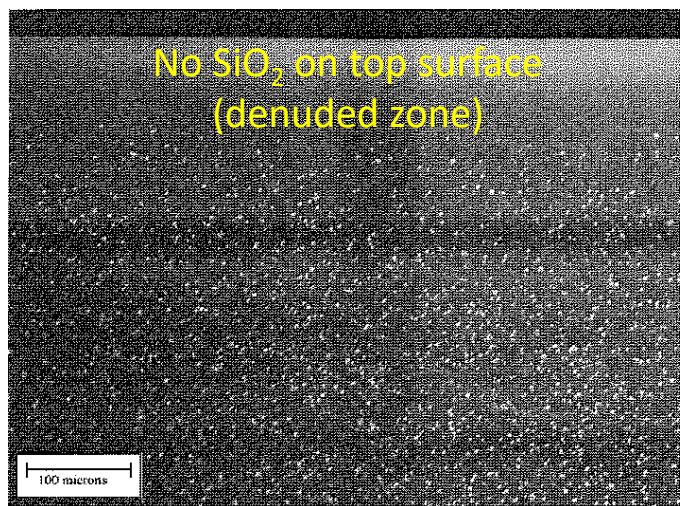


finished
boules



Oxygen and carbon in CZ silicon

- The CZ growth process inherently introduces O (from SiO_2 crucible) and C (from graphite susceptor/supporter).
- Typically, $C_O \approx 10^{18} \text{ cm}^{-3}$ and $C_C \approx 10^{16} \text{ cm}^{-3}$.
- The O in CZ silicon often forms small SiO_2 precipitates in the Si crystal under normal processing conditions.
- O and these precipitates can actually be very useful: provide mechanical strength, internal gettering (see next chapter).



SiO_2 precipitates (white dots) in bulk Si

Chapter 3 Crystal growth, wafer fabrication and basic properties of silicon wafers

1. Silicon crystal structure and defects.
2. Czochralski single crystal growth.
3. Growth rate and dopant incorporation for CZ method.
4. Float zone single crystal growth and doping.
5. Wafer fabrication.

Czochralski process - crystal growth rate

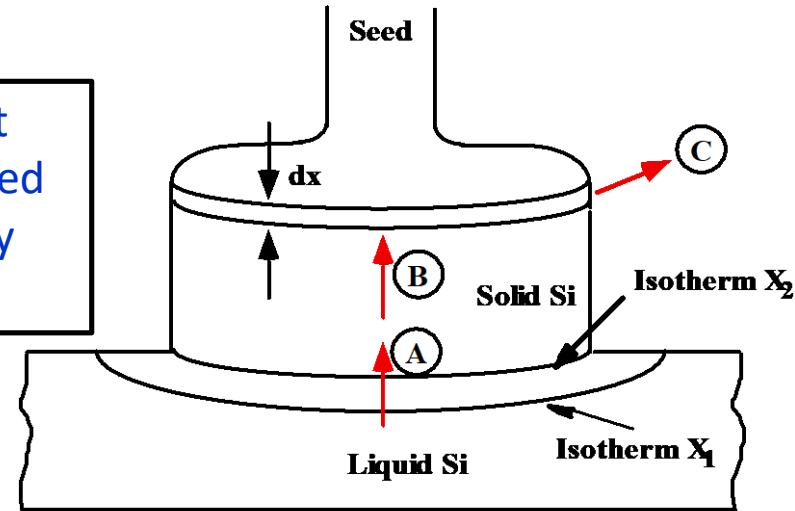
We wish to find a relationship between pull rate and crystal diameter.

Heat balance:

latent heat of fusion
(i.e. crystallization) +
heat conducted from
melt to crystal
(A)

=
heat
conducted
through the
crystal
(B)

=
heat
radiated
away
(C).



Freezing occurs between isotherms X_1 and X_2 .

$$(1) \quad L \frac{dm}{dt} + k_L \frac{dT}{dx_1} A_1 = k_S \frac{dT}{dx_2} A_2$$

$\underbrace{\qquad\qquad\qquad}_{(A)} \quad \underbrace{\qquad\qquad\qquad}_{(B)}$

$A_{1,2}$ = Cross-sectional area

L = latent heat of fusion

$\frac{dm}{dt}$ = amount of freezing per unit time

k_L = thermal conductivity of liquid

$\frac{dT}{dx_1}$ = thermal gradient at isotherm x_1

k_S = thermal conductivity of solid

$\frac{dT}{dx_2}$ = thermal gradient at x_2

Czochralski process - crystal growth rate

The rate of growth of the crystal is

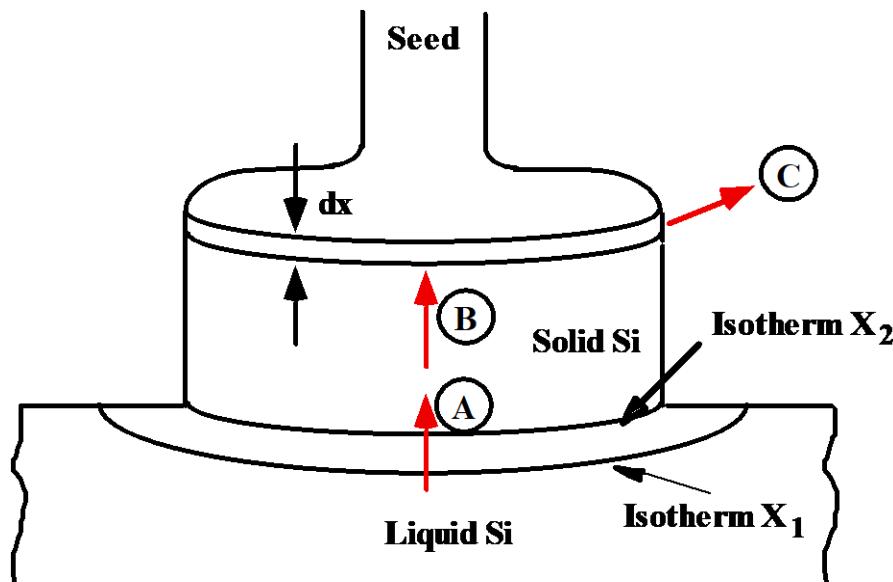
$$\frac{dm}{dt} = v_p A N \quad (2)$$

where v_p is the pull rate and N is the density.

Neglecting the middle term in Eqn. (1) we have:

$$v_{p\text{MAX}} = \frac{k_s}{LN} \frac{dT}{dx_2} \quad (3)$$

In order to replace dT/dx_2 , we need to consider the heat transfer processes.



Heat radiation from the crystal (C) is given by the Stefan-Boltzmann law*:

$$dQ = (2\pi r dx)(\sigma \epsilon T^4) \quad (4)$$

Heat conduction up the crystal is given by

$$Q = k_s (\pi r^2) \frac{dT}{dx} \quad (5)$$

$2\pi r dx$ = radiation surface.

σ = Boltzman constant (NOT k_B)

k_s = thermal conductivity of the solid.

$\epsilon=1$ for perfect blackbody.

$\epsilon<1$ for "grey-body".

*Stephen-Boltzmann law:

http://en.wikipedia.org/wiki/Stefan%20Boltzmann_law

Czochralski process - crystal growth rate

- Differentiating (5), we have

$$\frac{dQ}{dx} = k_s (\pi r^2) \frac{d^2 T}{dx^2} + (\pi r^2) \frac{dT}{dx} \frac{dk_s}{dx} \approx k_s (\pi r^2) \frac{d^2 T}{dx^2} \quad (6)$$

- Substituting (6) into (4), we have

$$\frac{d^2 T}{dx^2} - \frac{2\sigma\varepsilon}{k_s r} T^4 = 0 \quad (7)$$

- k_s varies roughly as $1/T$, so if k_M is the thermal conductivity at the melting point,

$$k_s = k_M \frac{T_M}{T} \quad (8)$$

$$\therefore \frac{d^2 T}{dx^2} - \frac{2\sigma\varepsilon}{k_M r T_M} T^5 = 0 \quad (9)$$

- Solving this differential equation, evaluating it at $x = 0$ and substituting the result into (3), we obtain (see text):

$$v_{P\text{MAX}} = \frac{1}{LN} \sqrt{\frac{2\sigma\varepsilon k_M T_M^5}{3r}} \quad (10)$$

$v_{p\text{max}}$ = maximum pull rate, inversely proportional to the square root of crystal radius.

Crystal growth rate: example

Example Calculate the maximum pull rate for a 6"-diameter CZ crystal.

Answer Numerical values to substitute into Eq. (3.31) are contained in the appendix

$$v_{P\text{MAX}} = \frac{1}{LN} \sqrt{\frac{2\sigma\varepsilon k_M T_M^5}{3r}}$$

For perfect blackbody: $\varepsilon=1$

For "grey" body: $\varepsilon<1$

$$= \frac{1}{(430 \text{ cal gm}^{-1})(2.328 \text{ gm cm}^{-3})}.$$

$$\sqrt{\frac{2\left(5.67 \times 10^{-5} \frac{\text{erg}}{\text{cm}^2 \text{ sec K}^4}\right)(0.55)\left(0.048 \frac{\text{cal}}{\text{sec cm K}}\right)(1690 \text{ K})^5\left(2.39 \times 10^{-8} \frac{\text{cal}}{\text{erg}}\right)}{3(7.62 \text{ cm})}}$$

$$= 0.00656 \text{ cm sec}^{-1} \text{ or } 23.6 \text{ cm hr}^{-1}$$

~4 hours to pull one meter long boule.

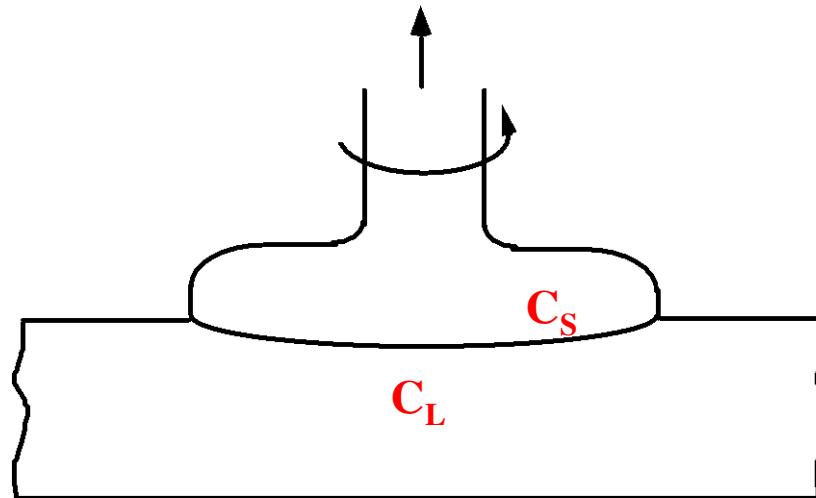
However, this equation is not so accurate, and in practice feedback is used to adjust the pulling rate in order to maintain constant diameter.

Dopant incorporation during crystal growth

- Dopants are added to the melt to provide a controlled N or P doping level in the wafers.
- However, the dopant incorporation process is complicated by dopant segregation.
- Generally, impurities “prefer to stay in the liquid” as opposed to being incorporated into the solid.
- This process is known as segregation. The degree of segregation is characterized by the segregation coefficient, k_o , for the impurity.

$$k_o = \frac{C_s}{C_L}$$

C_s and C_L are the impurity concentration just on the either side of the solid/liquid interface.



Dopant behavior during crystal growth

$$k_o = \frac{C_s}{C_L}$$

Dopant	k_o
As	0.3
Bi	7×10^{-4}
C	0.07
Li	10^{-2}
O	0.5
P	0.35
Sb	0.023
Al	2.8×10^{-3}
Ga	8×10^{-3}
B	0.8
Au	2.5×10^{-5}

Most k_o values are < 1 which means the impurity prefers to stay in the liquid.

Thus as the crystal is pulled, dopant concentration will increase.

In other words, the distribution of dopant along the ingot will be graded.

Distribution coefficient: example

Example Find the concentration of phosphorous (P) atoms in the melt to obtain Si doped with 10^{16} atoms/cm³ (Czochralski growth) $k_d = 0.35$ for P in Si.

$$k_d = \frac{c_s}{c_L} \Rightarrow c_L = \frac{10^{16} \text{ atoms / cm}^3}{0.35} = 2.86 \times 10^{16} \text{ atoms / cm}^3$$

How many grams of P should be added if the initial load in the crucible is 5 kg of Si? (density of Si = 2.33g/cm³)

$$v = \frac{m}{d} = \frac{5000 \text{ g}}{2.3 \text{ g / cm}^3} = 2145.9 \text{ cm}^3$$

In the total melt volume, we want 2.86×10^{16} atoms/cm³. The number of atoms is:

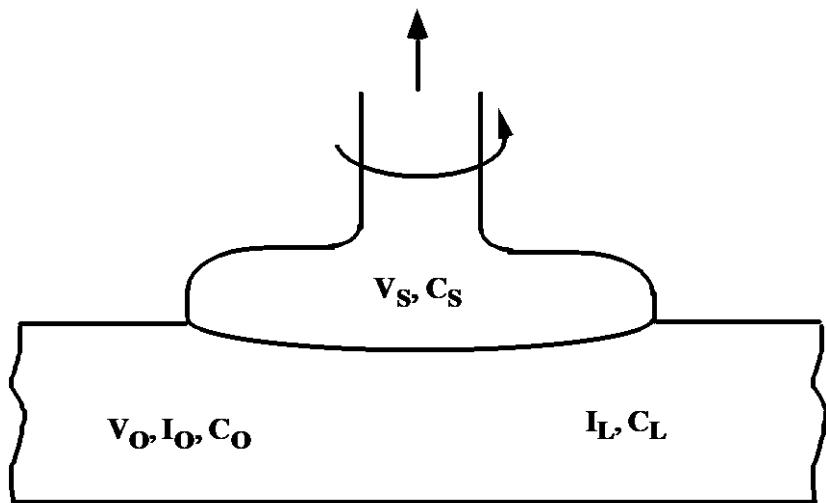
$$2.86 \times 10^{16} \text{ atoms / cm}^3 [V_{Si} + V_P] \quad \text{but } VP \ll VSi, \text{ so } VC \approx VSi$$

$$\text{P atoms} = 2.86 \times 10^{16} \text{ atoms / cm}^3 (2145.0 \text{ cm}^3) = 6.135 \times 10^{19} \\ = 1.019 \times 10^{-4} \text{ mol}$$

$$Z_p = 31 \text{ g}$$

$$\frac{6.135 \times 10^{19} \text{ atoms} \times 31 \text{ g / mole}}{6.023 \times 10^{23} \text{ atoms / mole}} = 3.159 \times 10^{-3} \text{ g} = 3.16 \text{ mg}$$

Dopant incorporation during crystal growth



V_0 = initial volume

I_0 = number of impurities

C_L = impurity concentration in the melt

C_0 = initial impurity concentration in the melt

$$C_0 = I_0/V_0$$

$$\text{By definition: } C_s/C_L = k_0$$

If during growth, an additional volume dV_S freezes, the impurities incorporated into dV are given by:

$$\therefore \int_{I_0}^{I_L} \frac{dI_L}{I_L} = -k_0 \int_0^{V_S} \frac{dV_S}{V_0 - V_S}$$

$$\therefore I_L = I_0 \left(1 - \frac{V_S}{V_0}\right)^{k_0}$$

$$dI_L = -k_0 C_L dV_S = -k_0 \frac{I_L}{V_0 - V_S} dV_S$$

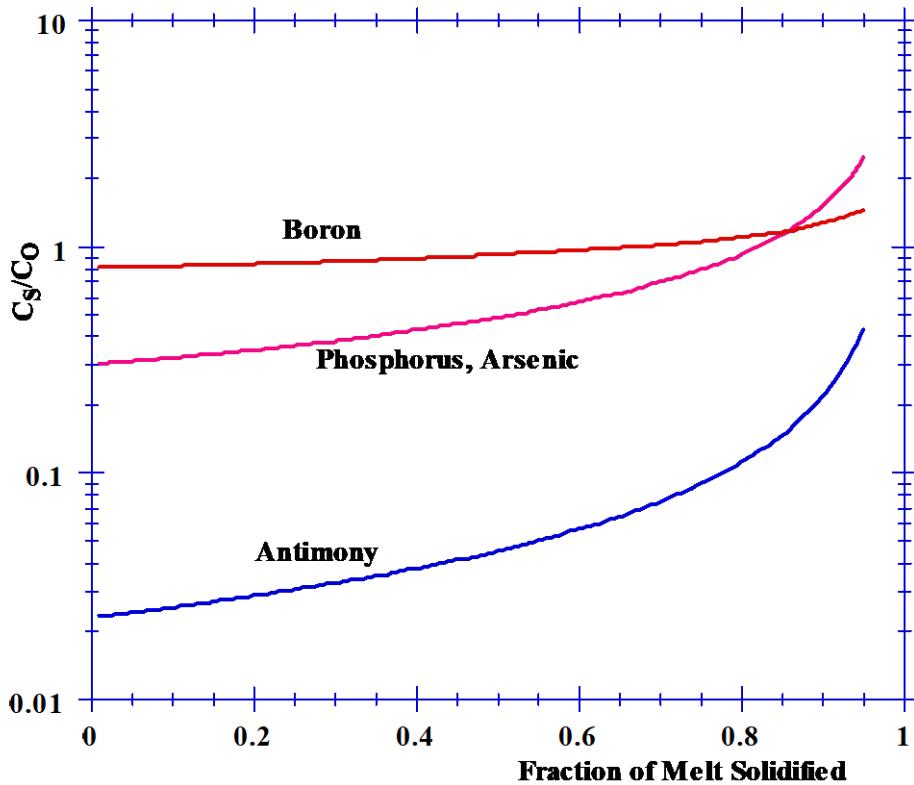
$$C_L = \frac{I_L}{V_L} = \frac{I_0 \left(1 - \frac{V_S}{V_0}\right)^{k_0}}{V_0 - V_S} = \frac{I_0}{V_0} \frac{(1-f)^{k_0}}{1-f} = C_0 (1-f)^{k_0-1}$$

The impurity level in the crystal (C_s) $C_s = -\frac{dI_L}{dV_S}$

$$\therefore C_s = C_0 k_0 (1-f)^{k_0-1}$$

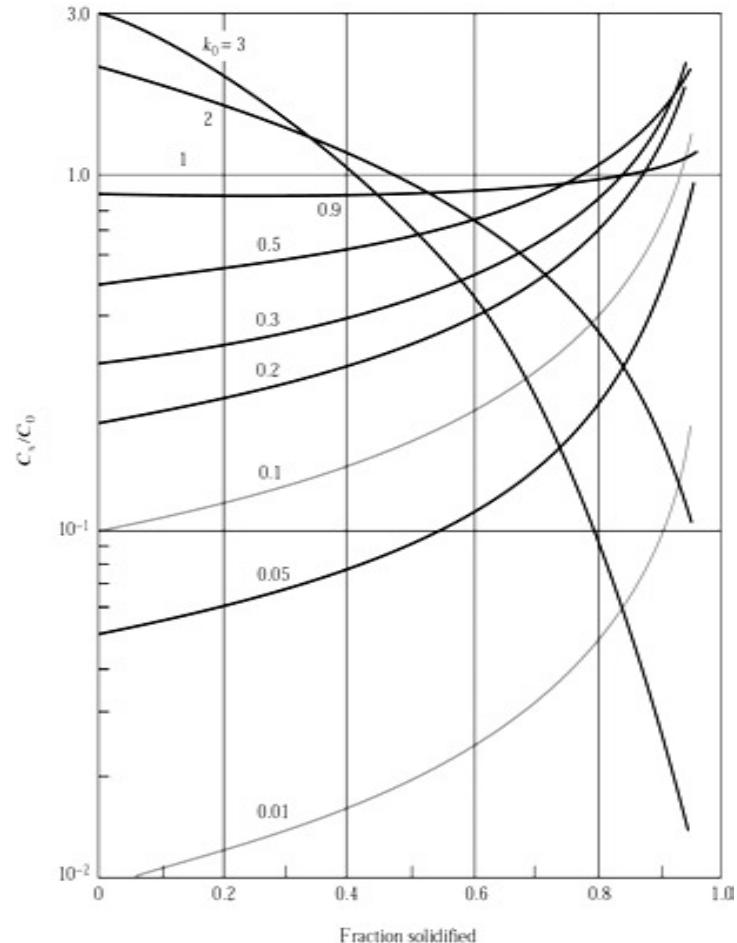
$f = V_s/V_0$, the ratio of solid silicon to the initial liquid volume.

Doping concentration versus position along the grown CZ crystal for common dopants in silicon



Note the relatively flat profile produced by boron with a k_o close to 1.

Dopants with $k_o \ll 1$ produce much more doping variation along the crystal.



Doping concentration in a solid as a function of the fraction solidified.
(but no dopant has $k_o > 1$)

Chapter 3 Crystal growth, wafer fabrication and basic properties of silicon wafers

1. Silicon crystal structure and defects.
2. Czochralski single crystal growth.
3. Growth rate and dopant incorporation for CZ method.
4. Float zone single crystal growth and doping.
5. Wafer fabrication.
6. Measurement methods.

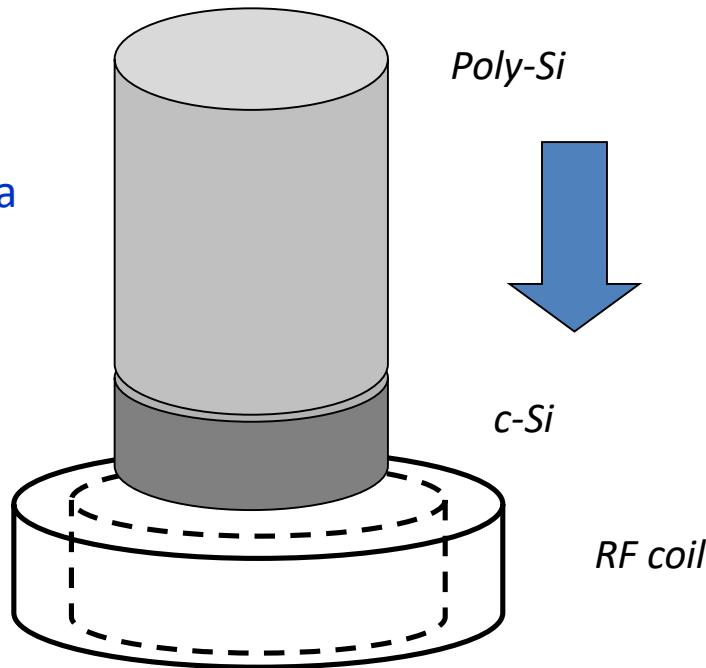
Float-zone crystal growth: overview

- For CZ-grown Si, impurities (O and C) can be introduced from the melt contacting the SiO_2 crucible and from graphite susceptor/supporter.
- This limits the resistivity to $\sim 20\Omega\text{cm}$, while intrinsic Si is $230\text{k}\Omega\text{cm}$.
- These crystals are more expensive and have very low oxygen and carbon and thus, are not suitable for the majority of silicon IC technology.
- Carrier concentrations down to 10^{11} atoms/cm³ have been achieved.
- It is far less common, and is reserved for situations where oxygen and carbon impurities cannot be tolerated.
- Float-zone does not allow as large Si wafers as CZ does (200mm and 300mm) and radial distribution of dopant in FZ wafer is not as uniform as in CZ wafer.
- It is good for solar cells, power electronic devices (thyristors and rectifiers) that use the entire volume of the wafer not just a thin surface layer, etc.

Float-zone crystal growth process

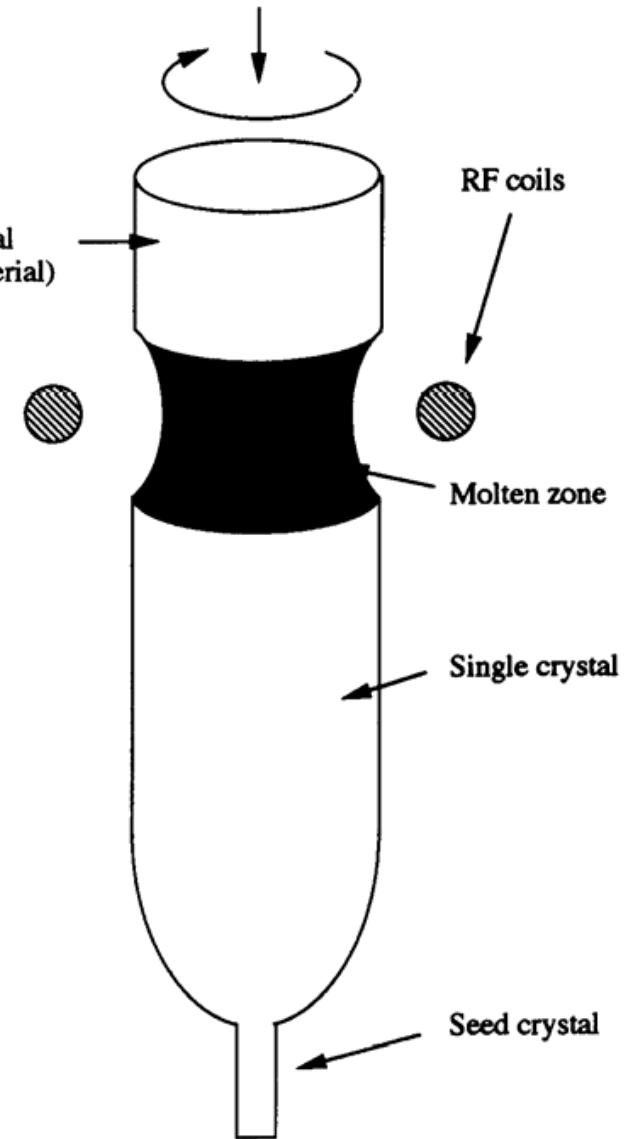
- Polycrystalline silicon is converted into single-crystal by zone heating (zone melting).
- The entire poly-Si rod from the EGS process is extracted as a whole.
- The rod is clamped at each end, with one end in contact with a single crystal seed.
- An RF heating coil induces eddy currents (power I^2R) in the silicon, heating it beyond its melting point in the vicinity of the coil.
- The "floating" melt zone is about 2cm wide/high.
- The seed crystal touches the melt zone and is pulled away, along with a solidifying Si boule following the seed. The crystalline direction follows that of the seed single crystal.
- Limited to about a 4" wafer, as the melt zone will collapse - it is only held together by surface tension (and RF field levitation).

Melt is not held in a container, it is “float”, thus the name “float zone”.



Float-zone: zone refining

- Dopants/impurities prefer to stay in the liquid than in the solid.
- Thus, the impurities generally stay in the melt zone, and don't solidify in the boule.
- That is, segregation (and evaporation) of impurities in the melt zone help purify the Si further.
- One can "purify" FZ wafers further by successively passing the coil along the boule. The impurities then segregate towards the end of the boule.
- Of course, if neglecting impurity evaporation, the total amount of impurity is the same. Yet the impurity at the lower part is much lower than, and at the upper part approaches to, the original impurity concentration.



Doping in FZ growth

(Doping can of course be achieved if the starting material poly-crystalline Si rod is doped. But due to zone-refining, the doping is not uniform along the boule)

Gas doping:

Dopants are introduced in gaseous form during FZ growth.

n-doping: PH_3 (Phosphine), AsCl_3

p-doping: B_2H_6 (Diborane), BCl_3

Good uniformity along the length of the boule.

Pill doping:

Drill a small hole in the top of the EGS rod, and insert the dopant.

If the dopant has a small segregation coefficient, most of it will be carried with the melt as it passes the length of the boule.

Resulting in only a small non-uniformity.

Ga and In doping work well this way.

Zone refining

Zone length is L . The rod has initial uniform impurity concentration of C_0 .

If the molten zone moves upwards by dx , the number of impurities in the liquid ($=I$) will change since some will be dissolved into the melting liquid at the top ($=C_0 dx$) and some will be lost to the freezing solid on the bottom ($=C_S dx = k_0 C_L dx$). Thus: (assume cross-sectional area =1)

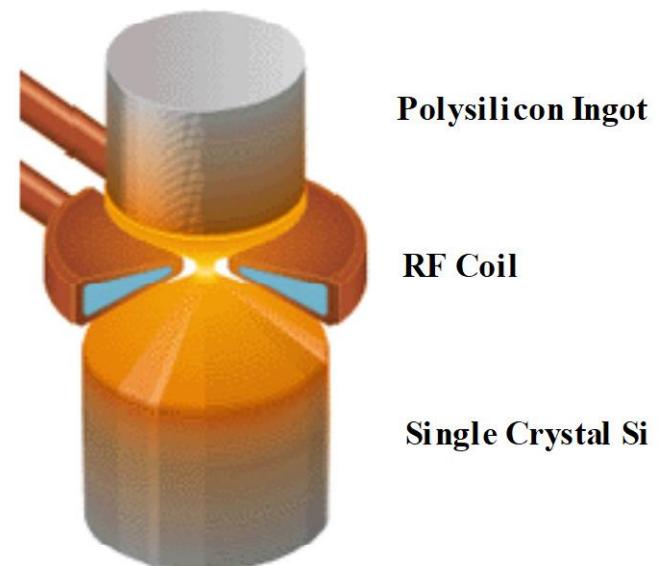
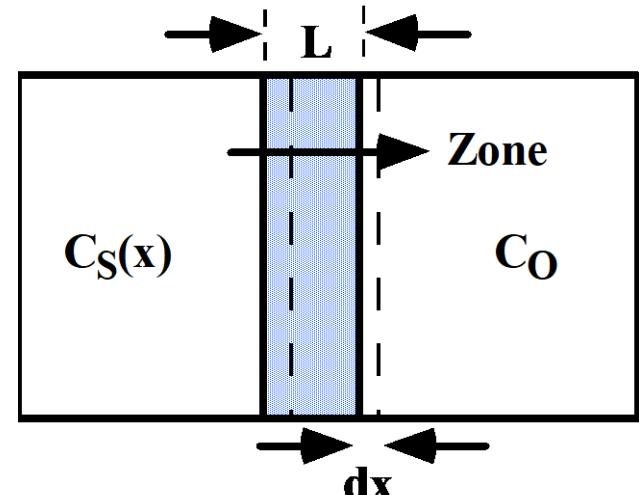
$$dI = (C_0 - k_0 C_L)dx, \text{ but } C_L = I / L$$

$$\int_0^x dx = \int_{I_0}^I \frac{dI}{C_0 - \frac{k_0 I}{L}}$$

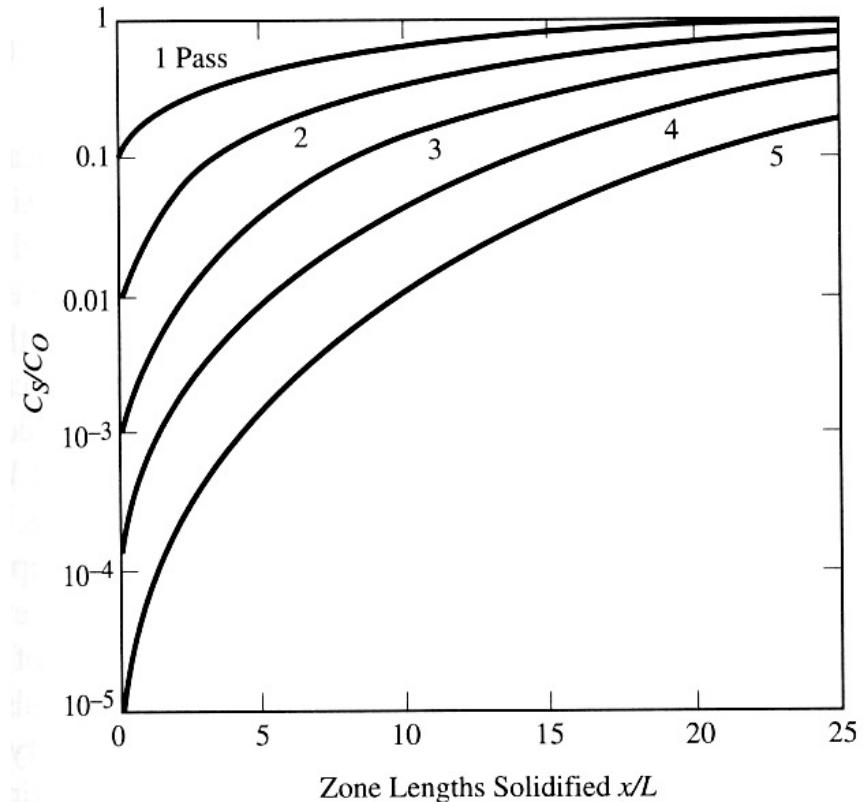
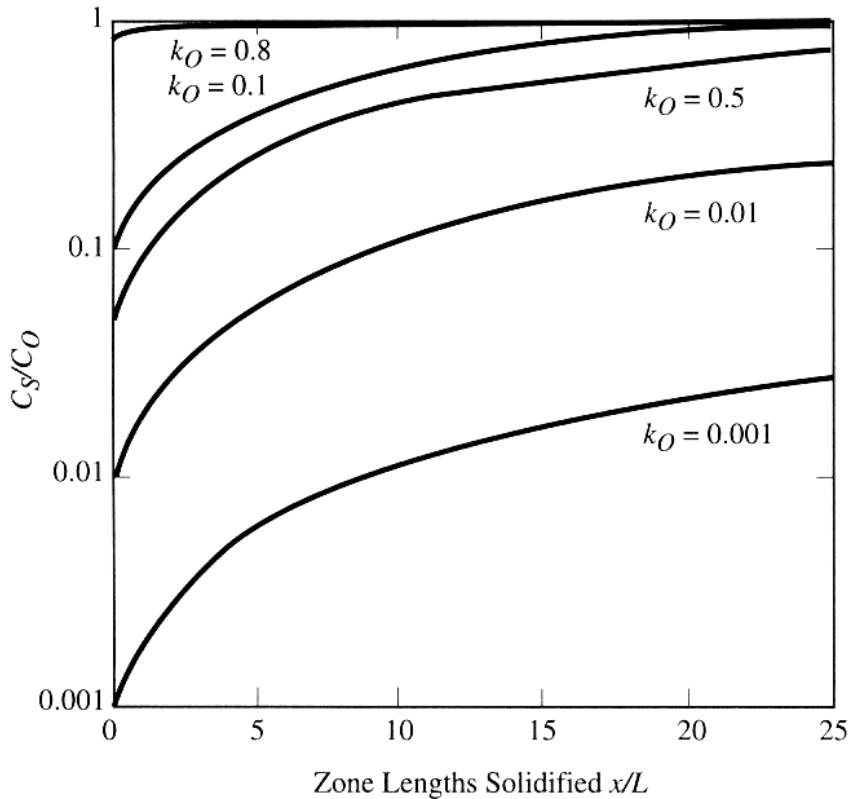
$$I = \frac{C_0 L}{k_0} - \left(\frac{C_0 L}{k_0} - I_0 \right) e^{-\frac{k_0 x}{L}}$$

$$C_S = k_0 C_L = k_0 \frac{I}{L}; I_0 = C_0 L$$

$$C_S = C_0 \left\{ 1 - \left(1 - k_0 \right) e^{-\frac{k_0 x}{L}} \right\}$$



Floating zone crystal growth – zone refining



Impurity during float-zone growth or zone refining. One pass of the molten zone through the solid.

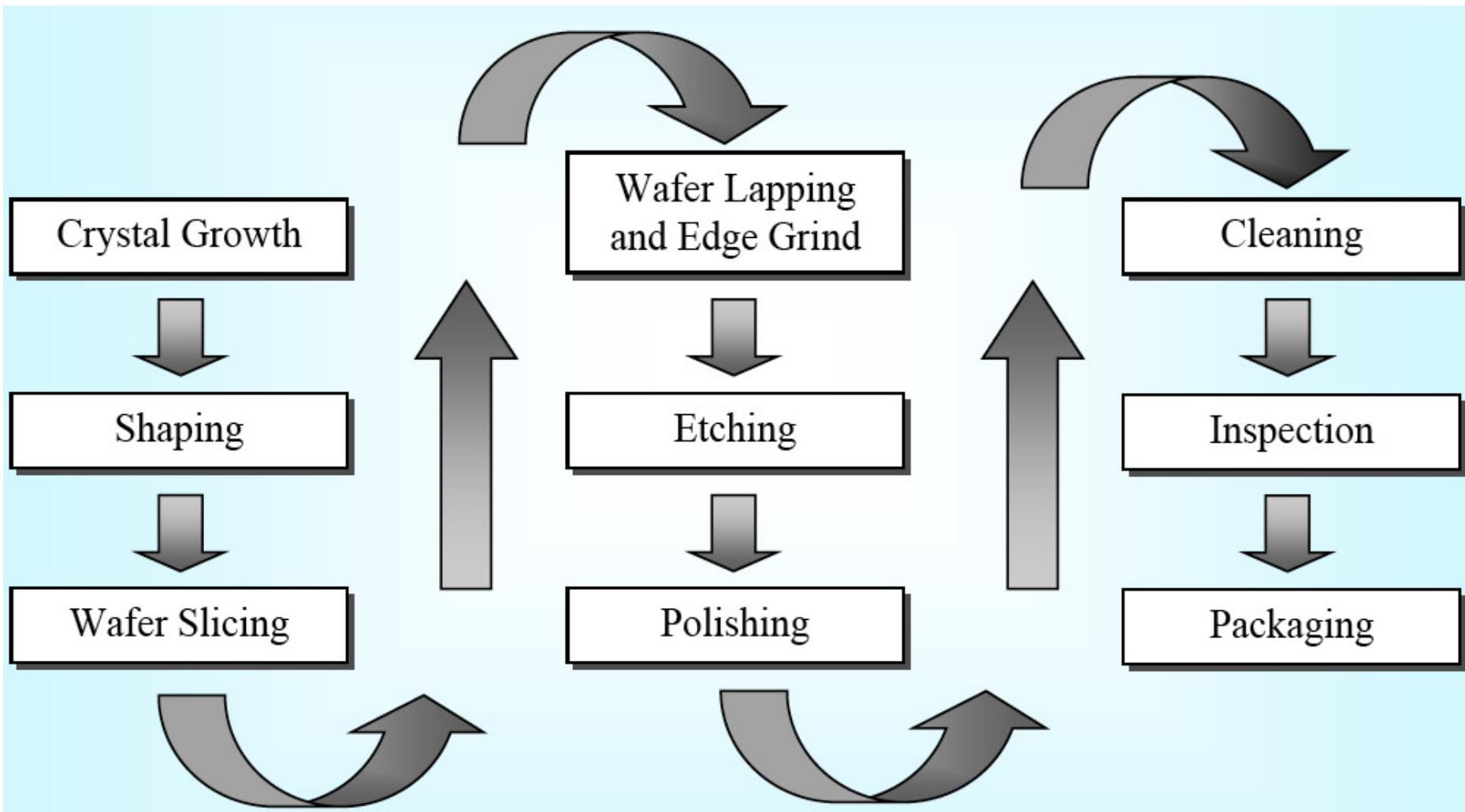
L is the length of the molten zone ($\sim 2\text{cm}$)

Zone refining with multiple passes, $k_0=0.1$. L is the length of the molten zone.

Chapter 3 Crystal growth, wafer fabrication and basic properties of silicon wafers

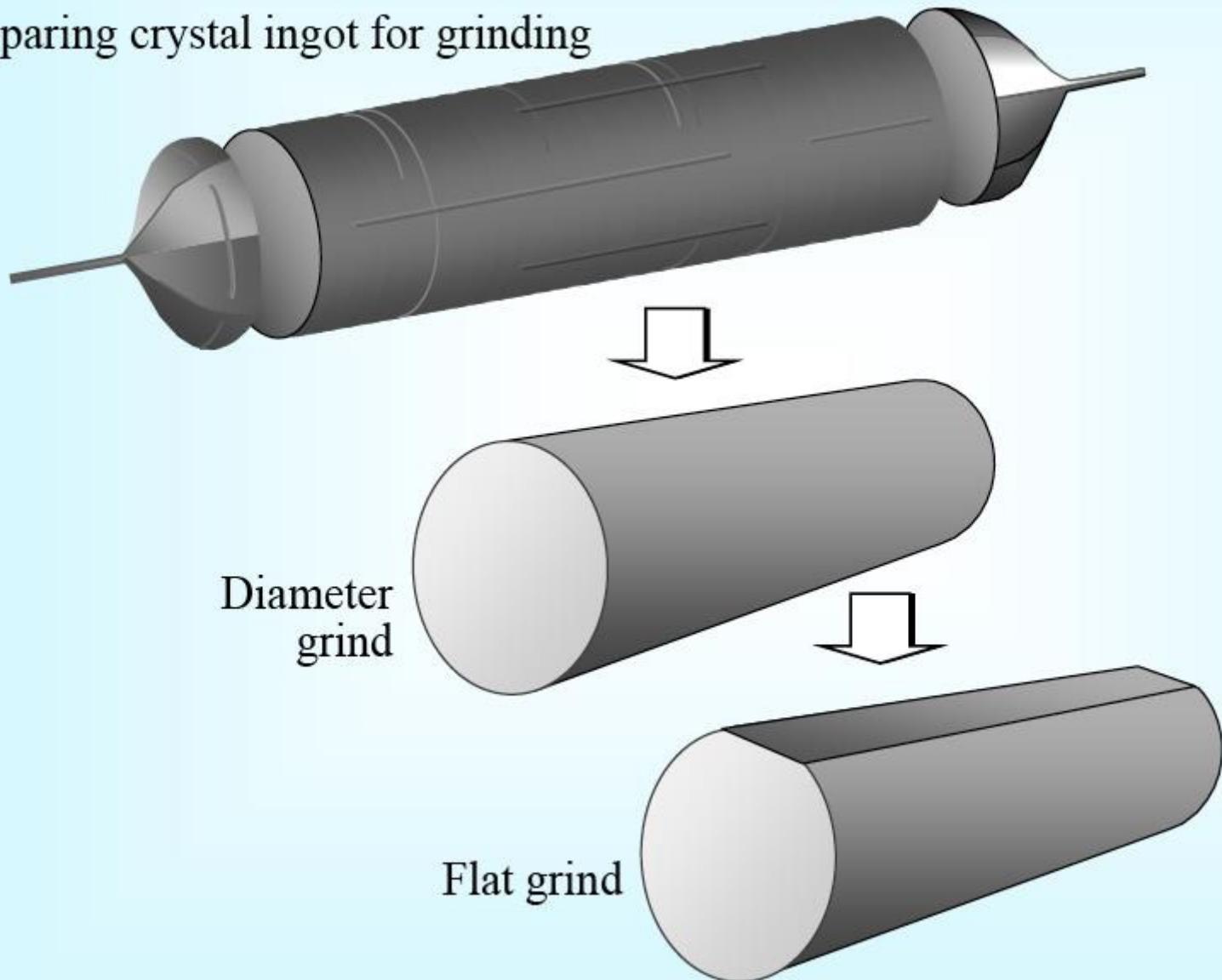
1. Silicon crystal structure and defects.
2. Czochralski single crystal growth.
3. Growth rate and dopant incorporation for CZ method.
4. Float zone single crystal growth and doping.
5. Wafer fabrication.

Steps for wafer preparation

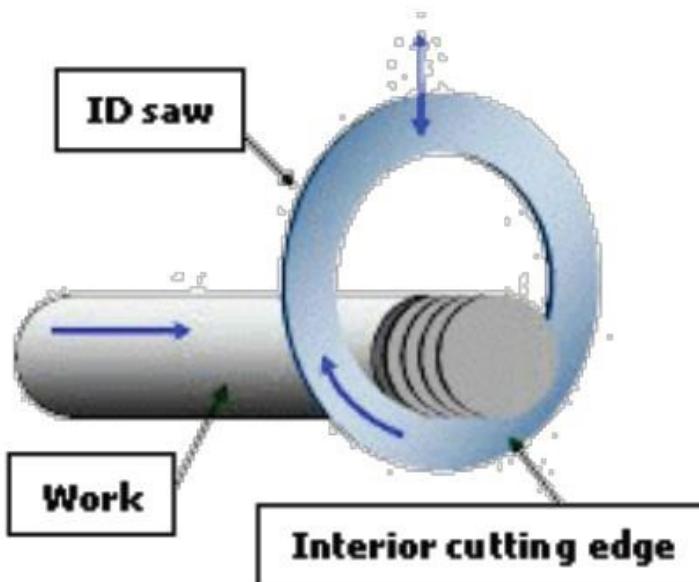


Ingot grinding

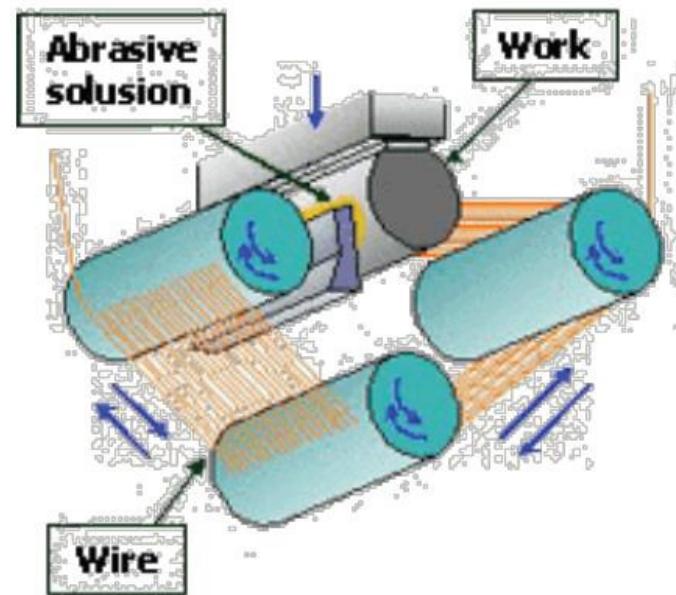
Preparing crystal ingot for grinding



Wafer slicing



Traditional method of slicing



Wire saw for large wafers



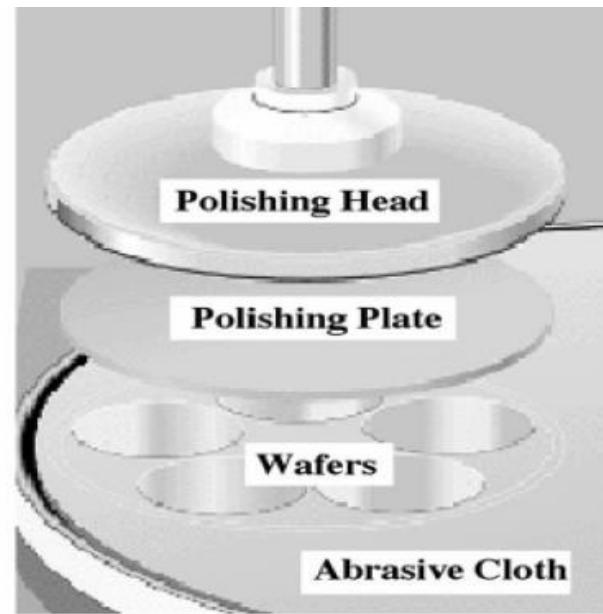
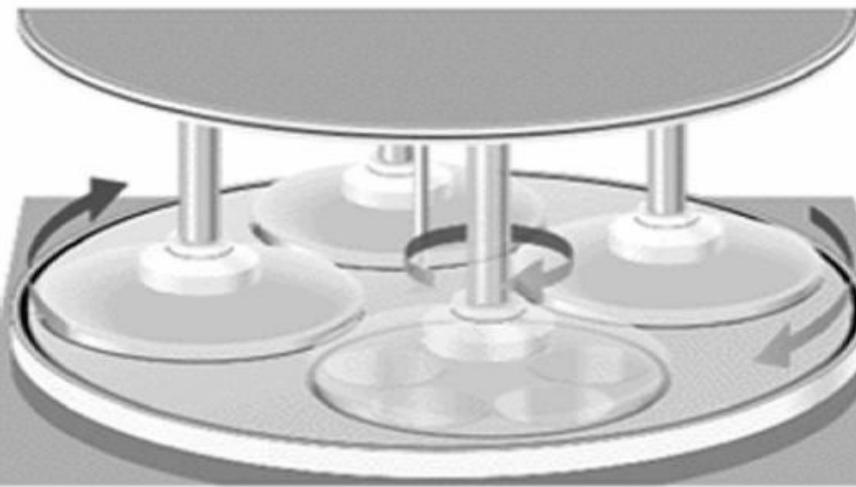
The saw blade itself is about $400\mu\text{m}$ thick, together with the loss at the seed and tail end of the crystal, only 50% of the boule ends up in wafer form.

After slicing, mechanical lapping and wet chemical etching is performed before final chemical mechanical polishing. The wet etching is typically:

$$3\text{Si} + 4\text{HNO}_3 + 18\text{HF} \rightarrow 3\text{H}_2\text{SiF}_6 + 4\text{NO} + 8\text{H}_2\text{O}$$

Wafer polishing

Chemical mechanical polishing

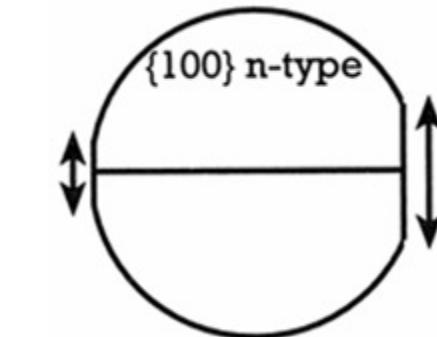
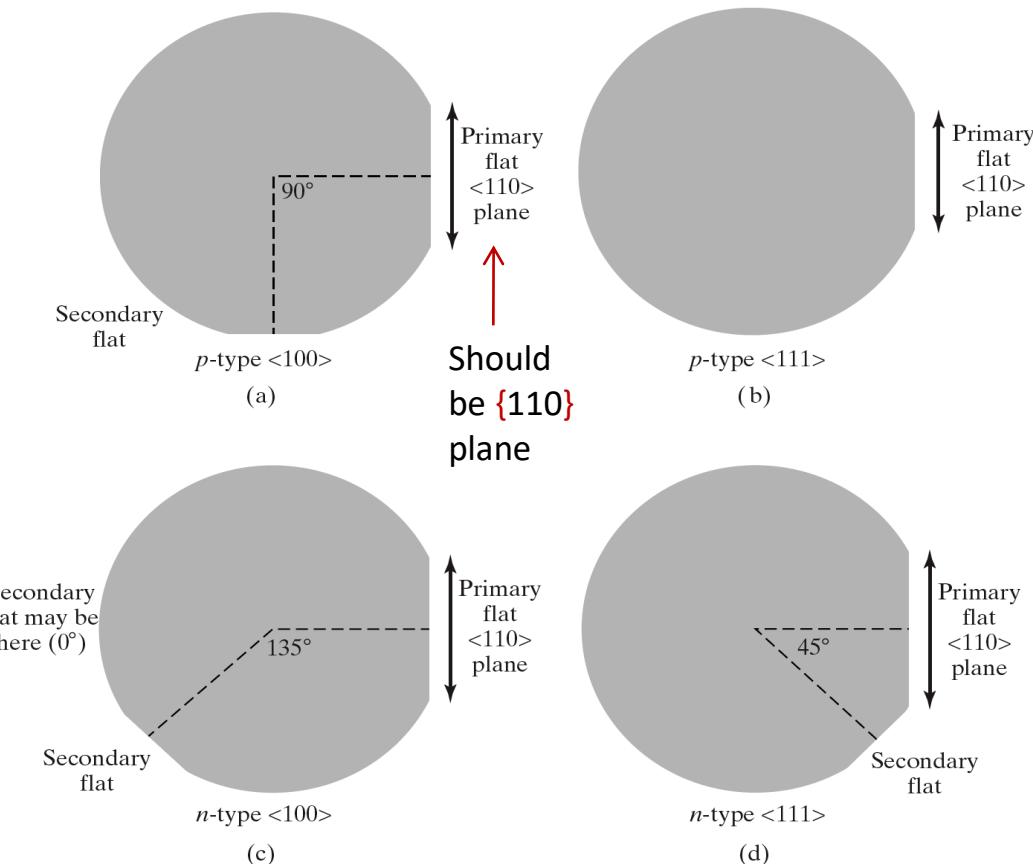


Slurry consists of nano-particles (10nm SiO_2 or Al_2O_3) and chemicals (NaOH).

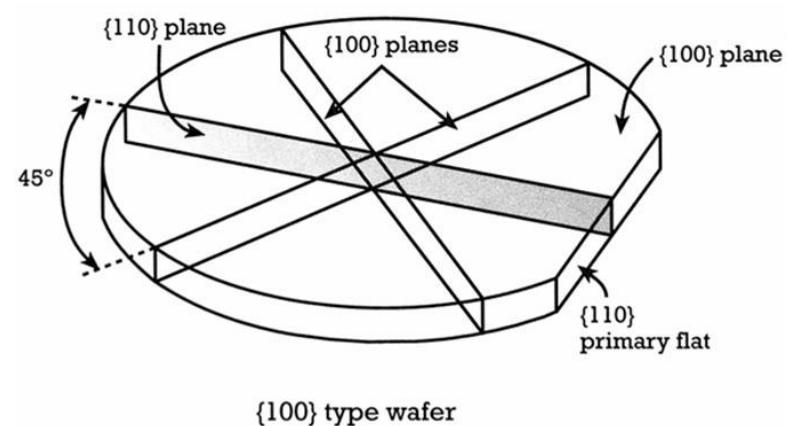


The rotation and pressure generates heat that drives a chemical reaction in which OH^- radicals from the NaOH oxidize the silicon. The SiO_2 particles abrade the oxide away.

Common (not always) wafer surface orientation



Another flat configuration
for $\{100\}$ n-type wafer



$\{100\}$ wafer usually breaks along $\{110\}$ plane (actually Si cleaves naturally along $\{111\}$ plane, which meet the surface at an angle of 54.7° , the angle between $<001>$ and $<111>$). Sometimes (not often) $\{100\}$ wafers break along $\{100\}$ plane. ($\{100\} = (100)+(010)+(001)$)⁴⁰

Chapter 4 Clean room, wafer cleaning and gettering

1. Introduction.
2. Clean room.
3. Wafer cleaning.
4. Gettering.
5. Measurement methods.

Effect of defect and contamination on semiconductor industry

Year of 1st DRAM Shipment	1997	1999	2003	2006	2009	2012
Minimum Feature Size	250nm	180nm	130nm	100nm	70nm	50nm
Wafer Diameter (mm)	200	300	300	300	450	450
DRAM Bits/Chip	256M	1G	4G	16G	64G	256G
DRAM Chip Size (mm ²)	280	400	560	790	1120	1580
Microprocessor Transistors/chip	11M	21M	76M	200M	520M	1.40B
Critical Defect Size	125nm	90nm	65nm	50nm	35nm	25nm
Starting Wafer Total LLS (cm⁻²)	0.60	0.29	0.14	0.06	0.03	0.015
DRAM GOI Defect Density (cm⁻²)	0.06	0.03	0.014	0.006	0.003	0.001
Logic GOI Defect Density (cm⁻²)	0.15	0.15	0.08	0.05	0.04	0.03
Starting Wafer Total Bulk Fe (cm⁻³)	3x10¹⁰	1x10¹⁰	Under 1x10¹⁰	Under 1x10¹⁰	Under 1x10¹⁰	Under 1x10¹⁰
Metals on Wafer Surface After Cleaning (cm⁻²)	5x10⁹	4x10⁹	2x10⁹	1x10⁹	< 10⁹	< 10⁹
Starting Material Recombination Lifetime (μsec)	≥ 300	≥ 325	≥ 325	≥ 325	≥ 450	≥ 450

LLS: localized light scatters (use laser to detect and count particles)

GOI: gate oxide integrity, by electrical measurement

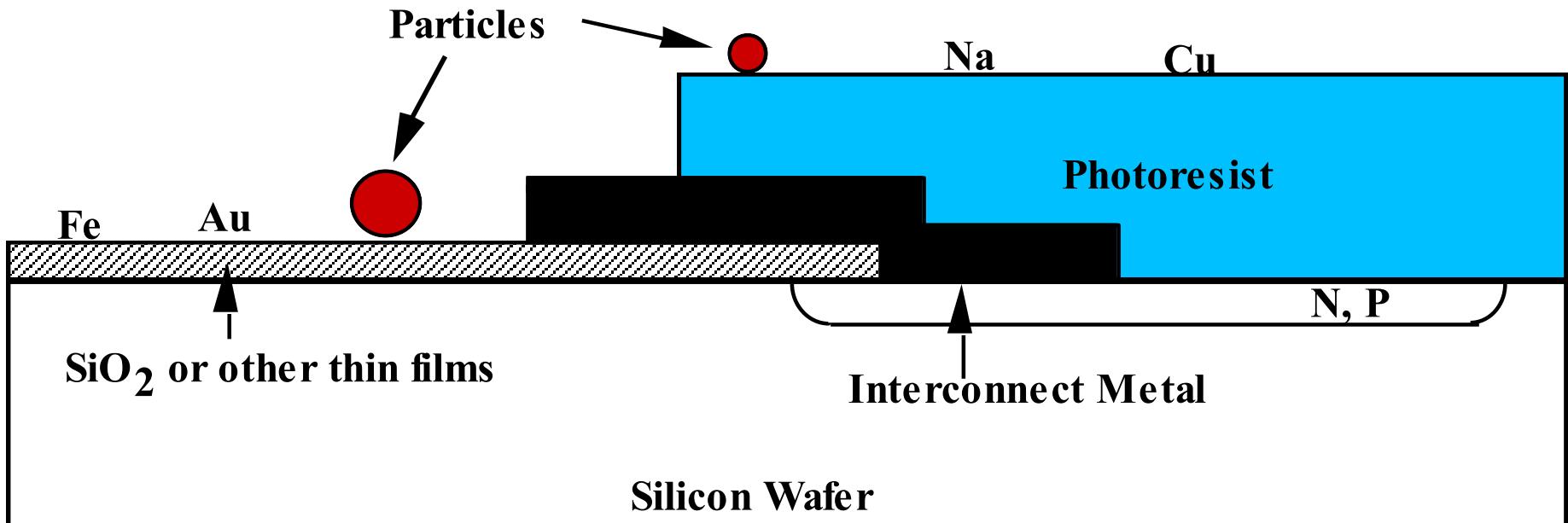
10⁹/cm² ~ 0.0001% monolayer

Importance of unwanted impurities increases with shrinking geometries of devices.

75% of the yield loss is due to defects caused by particles (1/2 of the min feature size).

Type of contaminants

Contaminants: particles, organic films, photoresist, heavy metals or alkali ions.



Effects on MOSFET

MOSFET threshold voltage is given by:

$$V_{TH} = V_{FB} + 2\phi_f + \frac{\sqrt{2\varepsilon_S q N_A (2\phi_f)}}{C_O} + \frac{q Q_M}{C_O}$$

If $t_{ox}=10\text{nm}$, then a 0.1V V_{th} shift can be caused by Na^+ or K^+ of $Q_M=2.15\times10^{11}\text{ ions/cm}^2$ ($<0.1\%$ monolayer or 10ppm in the oxide). $\varepsilon_0=8.85\times10^{-12}\text{F/m}$, $\varepsilon_{ox}=3.9$

Chapter 4 Clean room, wafer cleaning and gettering

1. Introduction.
2. Clean room.
3. Wafer cleaning.
4. Gettering.
5. Measurement methods.

Clean factory is the first approach against contamination

Modern IC factories employ a three tiered approach to controlling unwanted impurities:

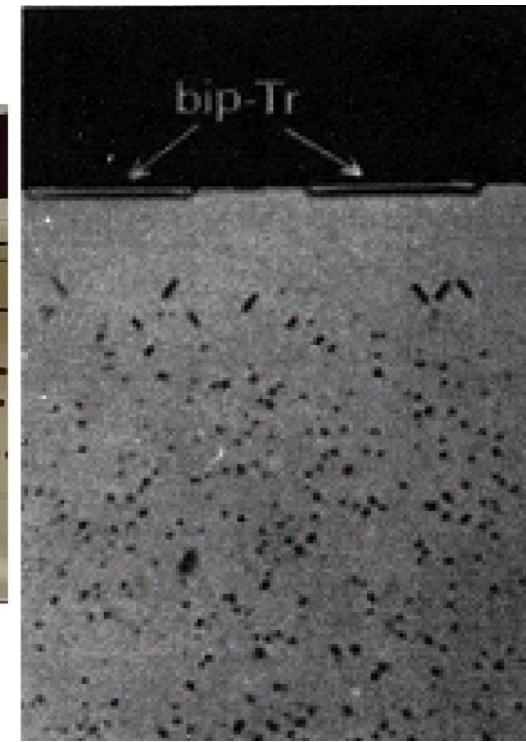
1. clean factories
2. wafer cleaning
3. gettering



Clean factory



Wafer cleaning



Gettering

Clean room

Factory environment is cleaned by:

- HEPA filters and recirculation for the air.
- “Bunny suits” for workers.
- Filtration of chemicals and gases.
- Manufacturing protocols.

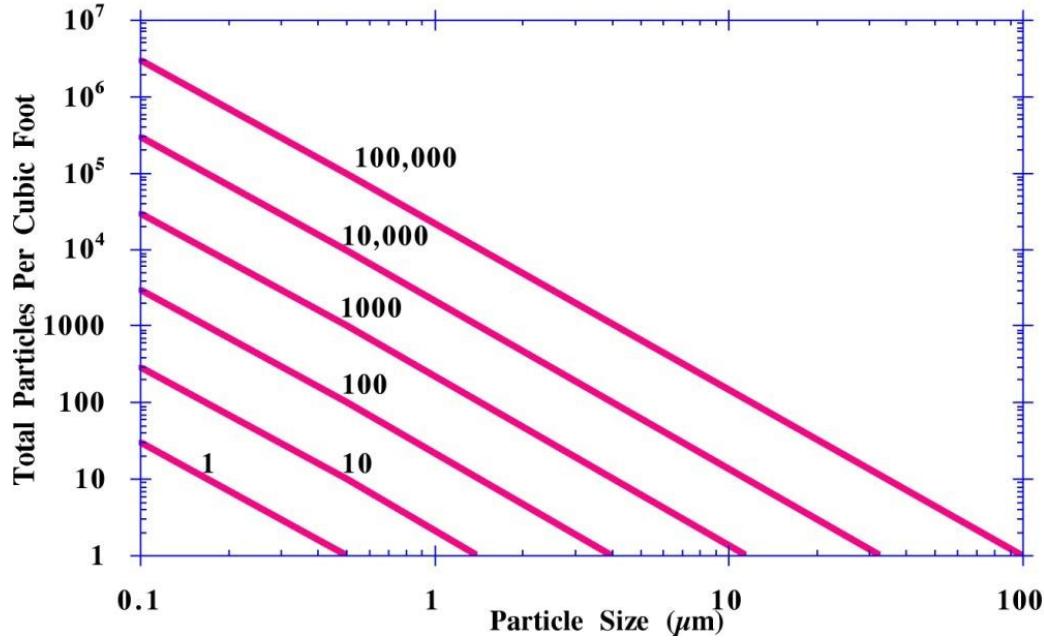
HEPA: High Efficiency Particulate Air

- HEPA filters composed of thin porous sheets of ultrafine glass fibers ($<1\mu\text{m}$ diameter).
- It is 99.97% efficient at removing particles from air.
- Room air forced through the filter at 50cm/sec.
- Large particles trapped, small ones stick to the fibers due to electrostatic forces.
- The exit air is typically better than class 1.



Class of a clean room

- Air quality is measured by the “class” of the facility.
- Class 1-100,000 mean number of particles, greater than $0.5\mu\text{m}$, in a cubit foot of air.
- A typical office building is about class 100,000.
- The particle size that is of most concern is 10nm – 10 μm . Particles <10nm tend to coagulate into large ones; those >10 μm are heavy and precipitate quickly.
- Particles deposit on surfaces by Brownian motion (most important for those < $0.5\mu\text{m}$) and gravitational sedimentation (for larger ones).



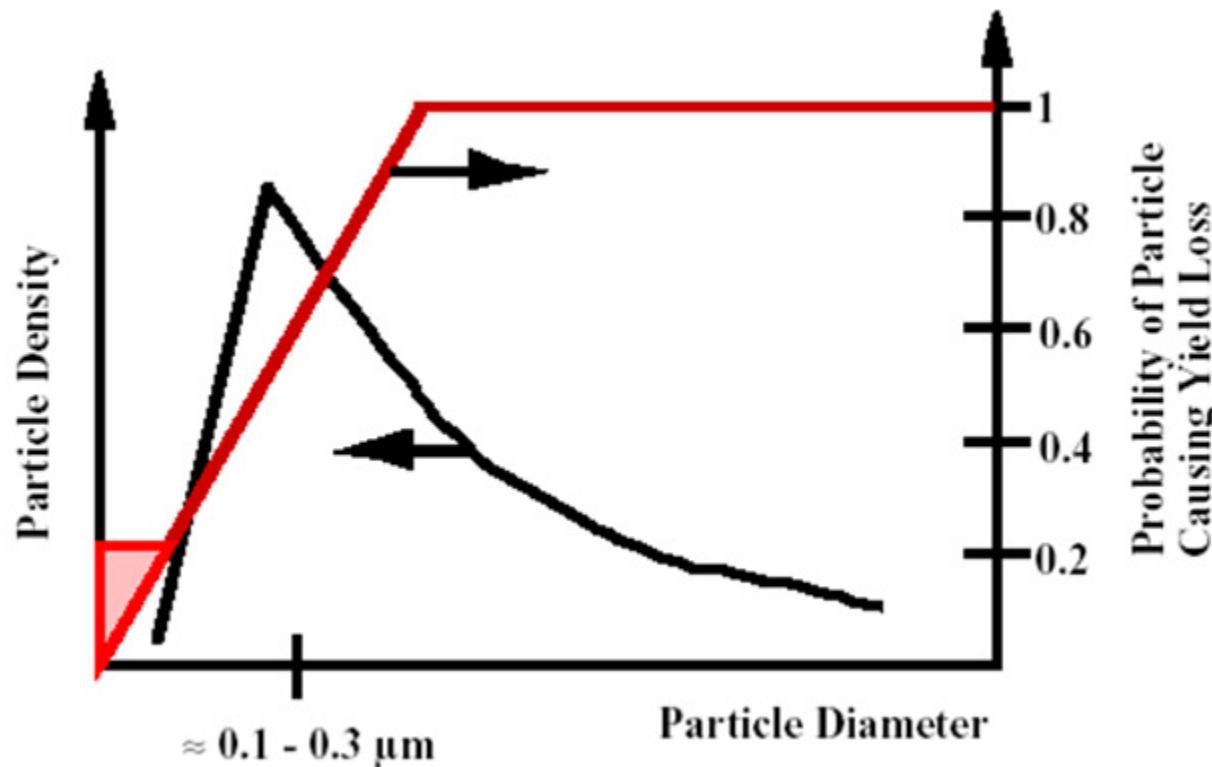
Particle diameter (μm)

Class	0.1	0.3	0.5	5.0
1	35	3	1	
10	350	30	10	
100		300	100	
1000			1000	7
10000			10000	70
100000			100000	700

by definition

Particle contamination and yield

- Generally, particles on the order of the technology minimum features size or larger will cause defect.
- $\approx 75\%$ yield loss in modern VLSI fabrication facilities is due to particle contamination.
- Particles on the order of $0.1\text{-}0.3\mu\text{m}$ are the most troublesome: larger particles precipitate easily; smaller one coagulate into larger particles.



Chapter 4 Clean room, wafer cleaning and gettering

1. Introduction.
2. Clean room.
3. Wafer cleaning.
4. Gettering.
5. Measurement methods.

Modern wafer cleaning

- Cleaning involves removing particles, organics and metals from wafer surfaces.
- Particles are largely removed by ultrasonic agitation during cleaning.
- Organics (photoresist) are removed in O₂ plasma or in H₂SO₄/H₂O₂ (Piranha) solutions.
- The “RCA clean” is used to remove metals and any remaining organics.

A cassette of wafers



Typical person emit 5-10 million particle per minute.

Most modern IC plants use robots for wafer handling.

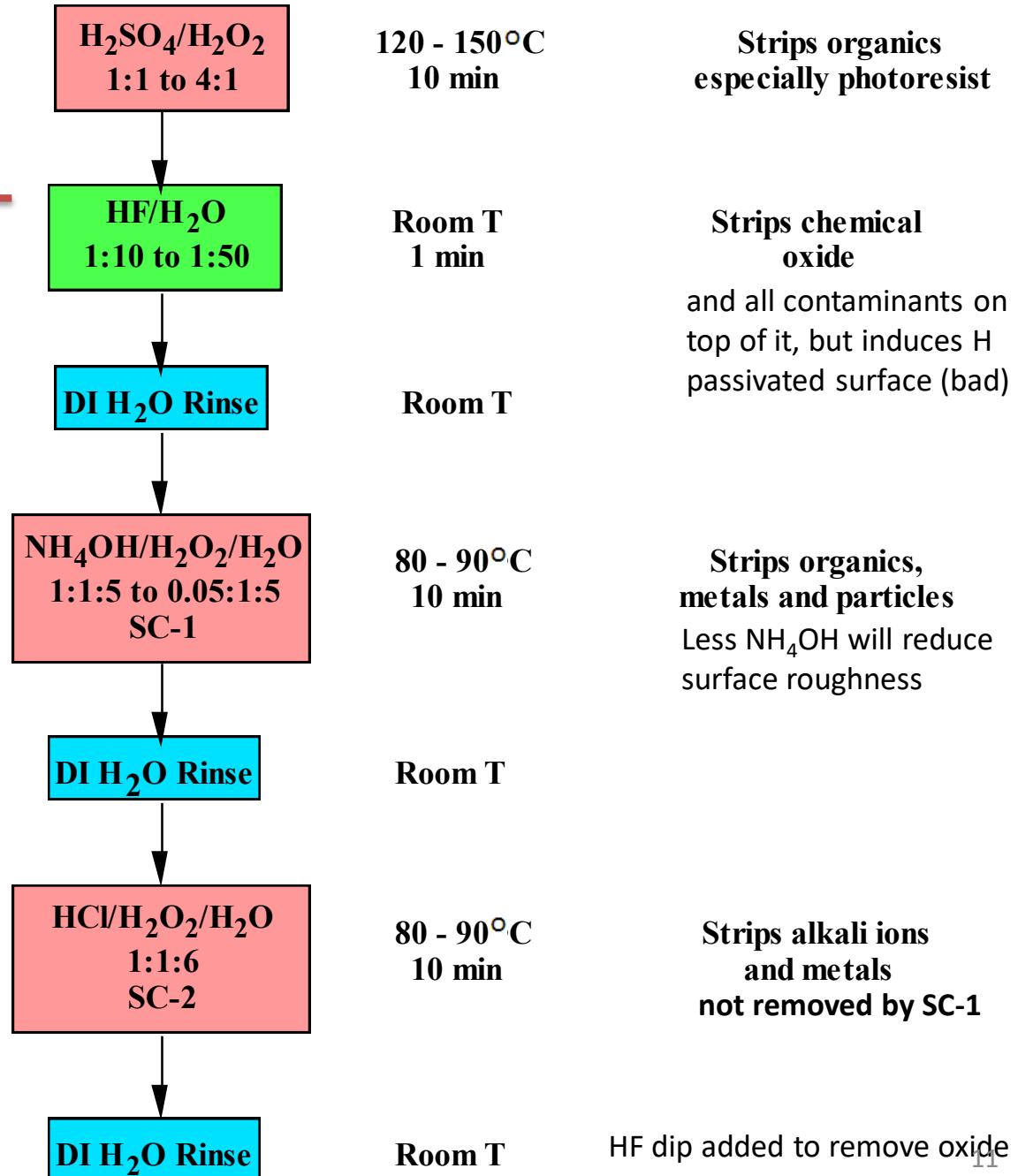
Standard RCA cleaning procedure

RCA clean is “standard process” used to remove organics, heavy metals and alkali ions.

Ultrasonic agitation is used to dislodge particles.

SC: Standard Cleaning

RCA: Radio Corporation of America, now makes TV, stereos...



DI water: de-ionized water

Standard cleaning (SC)

SC-1:

$\text{NH}_4\text{OH}(28\%):\text{H}_2\text{O}_2(30\%):\text{H}_2\text{O}=1:1:5 - 1:2:7$; 70-80°C, 10min, high pH.

- Oxidize organic contamination (form CO_2 , H_2O_2 ...)
- Form complex such as $\text{Cu}(\text{NH}_3)_4^{+2}$ with metals (IB, IIB, Au, Ag, Cu, Ni, Zn, Cd, Co, Cr).
- Slowly dissolve native oxide and grow back new oxide, which removes particles on oxide.
- But NH_4OH etches Si and make the surface rough, thus less NH_4OH is used today.

SC-2:

$\text{HCl}(73\%):\text{H}_2\text{O}_2(30\%):\text{H}_2\text{O}=1:1:6 - 1:2:8$; 70 - 80°C; 10min, low pH.

- Remove alkali ions and cations like Al^{+3} , Fe^{+3} and Mg^{+2} that form NH_4OH insoluble hydroxides in basic solutions like SC-1.
- These metals precipitate onto wafer surface in the SC-1 solution, while they form soluble complexes in SC-2 solution.
- SC-2 also complete the removal of metallic contaminates such as Au that may not have been completely removed by SC-1 step.

Ultrasonic cleaning and DI water

RCA cleaning with ultrasonic agitation is more effective in removing particles.

Ultrasonic cleaning:

- Highly effective for removing surface contaminants
- Mechanical agitation of cleaning fluid by high-frequency vibrations (between 20 and 45 kHz) to cause cavitation - formation of low pressure vapor bubbles that scrub the surface.
- Higher frequencies (>45kHz) form smaller bubbles, thus less effective.
- However, **megasonic** (1MHz) cleaning is also found effective in particle removal.

DI (de-ionized) water is used for wafer cleaning.

One monitors DI water by measuring its resistivity, which should be $>18\text{M}\Omega\text{cm}$.



$$\begin{aligned} \text{Diffusivity of: } & \text{H}^+ \approx 9.3 \times 10^{-5} \text{ cm}^2 \text{s}^{-1} \rightarrow \mu_{\text{H}^+} = qD/kT = 3.59 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1} \\ \text{of: } & \text{OH}^- \approx 5.3 \times 10^{-5} \text{ cm}^2 \text{s}^{-1} \rightarrow \mu_{\text{OH}^-} = qD/kT = 2.04 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1} \end{aligned}$$

$$\rho = \frac{1}{q([H^+] \mu_{H^+} + [OH^-] \mu_{OH^-})} = 18.5 \text{ M}\Omega\text{cm}$$

Chapter 4 Clean room, wafer cleaning and gettering

1. Introduction.
2. Clean room.
3. Wafer cleaning.
4. Gettering.
5. Measurement methods.

Gettering

- For the alkali ions, gettering generally uses dielectric layers on the topside; PSG for trapping, or Si_3N_4 layer for blocking them from getting into the device region.
- For metal ions, gettering generally uses traps on the wafer backside or in the wafer bulk. Here gettering works because the metals (Au...) do not “fit” in the silicon lattice easily because of their very different atomic size, thus they prefer to stay at defect sites.
- Therefore, the idea of gettering is to create such defect sites outside of active device region.
- Backside = *external* gettering: roughing/damaging the backside of the wafer, or depositing a poly-silicon layer, to provide a low energy “sink” for impurities.
- Bulk = *intrinsic* (or internal) gettering: using internal defects to trap impurities, thus moving them away from the active region of the wafer.

PSG: phosphosilicate glass, is a $\text{P}_2\text{O}_5/\text{SiO}_2$ glass that is normally deposited by CVD, usually contains 5% by weight phosphorus.

PSG traps alkali ions (Na^+ , K^+) and form stable compounds.

At higher than room temperature, alkali ions can diffuse into PSG from device region and trapped there.

Problems with PSG: it affects electric fields since dipoles exist in PSG, and it absorb water, leading to Al corrosion.

Gettering

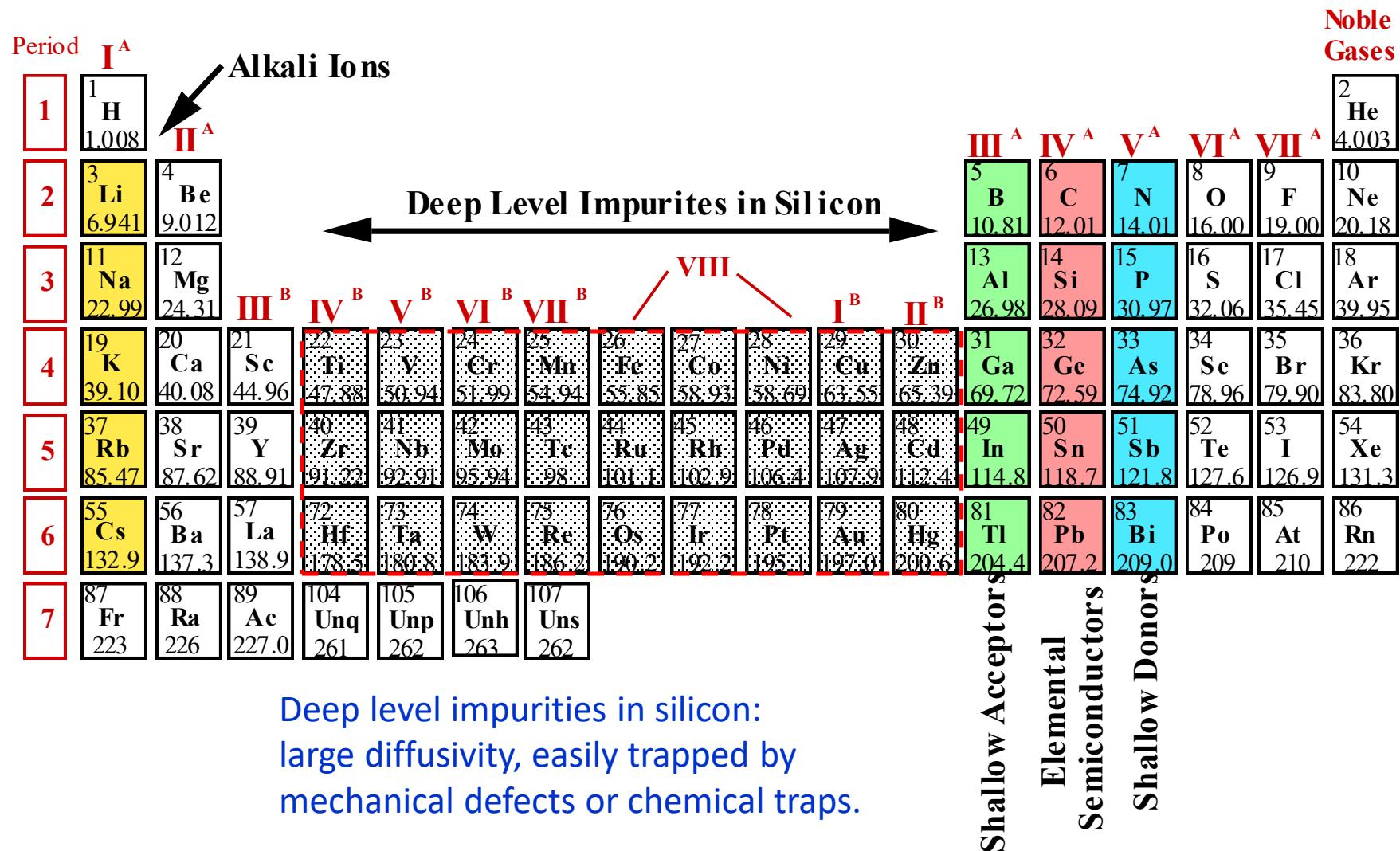
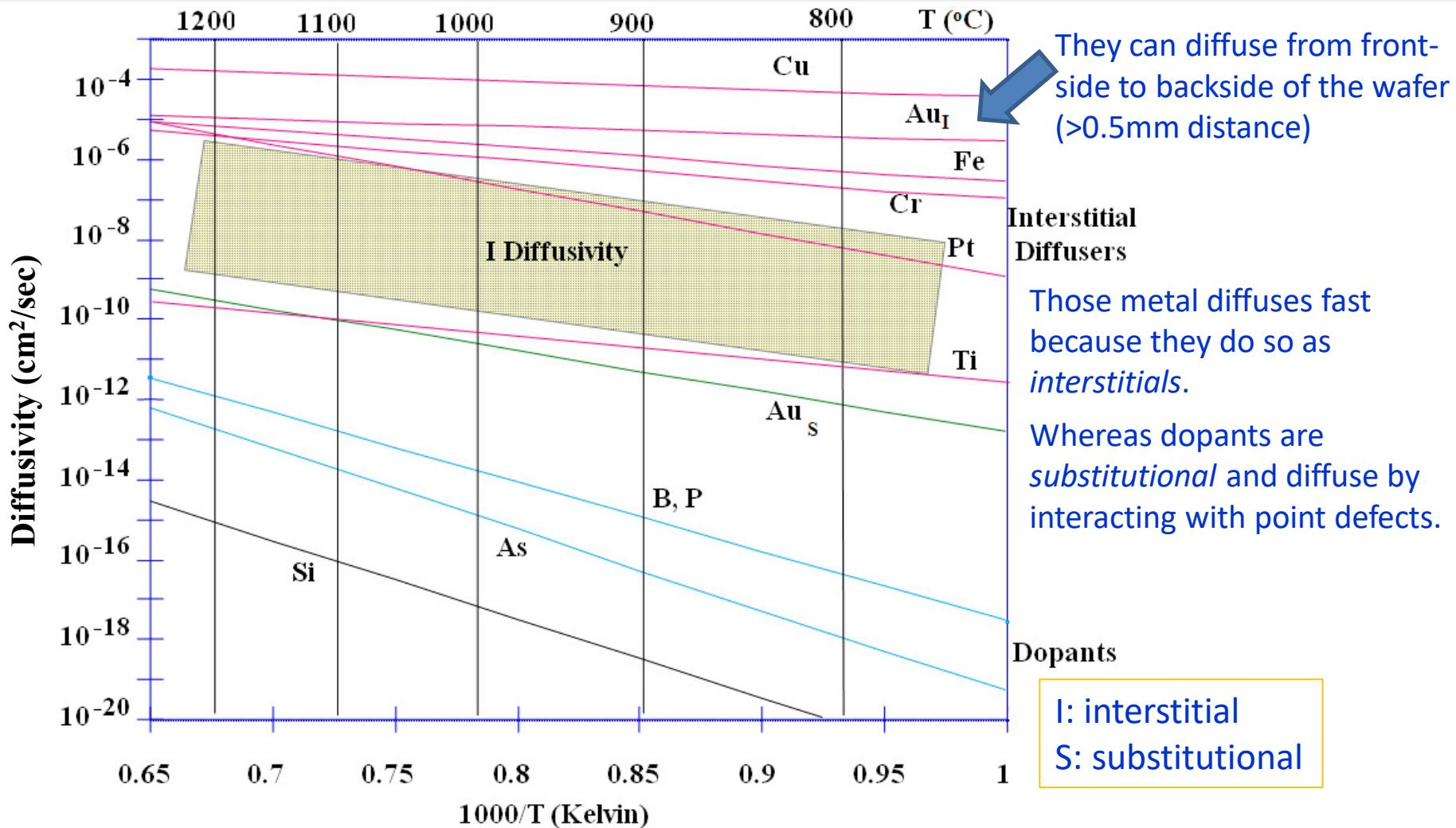


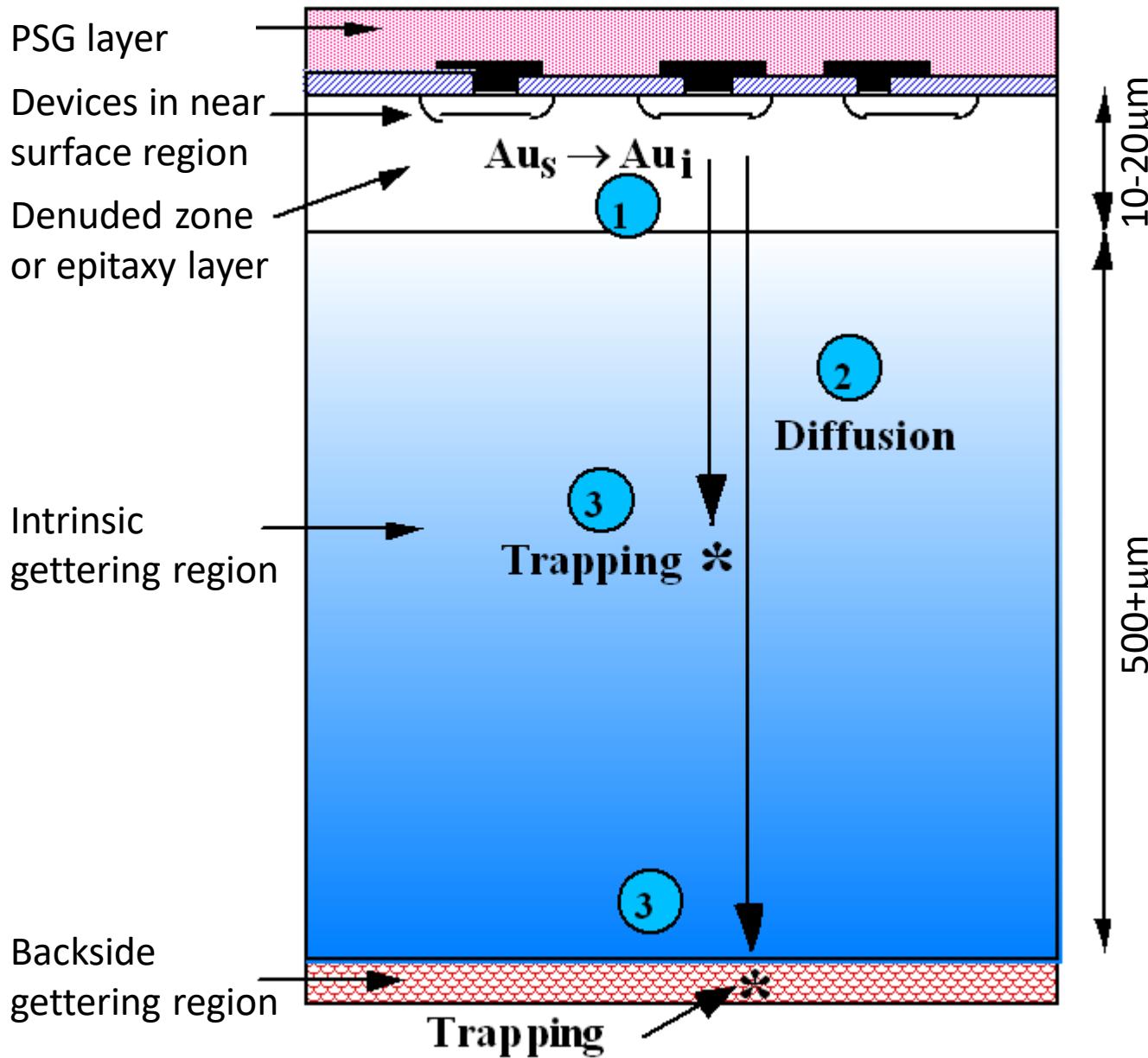
Figure 4-6 Periodic table indicating the elements that are of most concern in gettering.

Fast diffusion of various impurities



Heavy metal gettering relies on metal's very high diffusivity (when in interstitial sites) in silicon, and its preference to segregate to "trap" sites.

Gettering mechanism



Gettering mechanism

Gettering consists of:

1. Making metal atoms mobile.
2. Migration of these atoms to trapping sites.
3. Trapping of atoms.

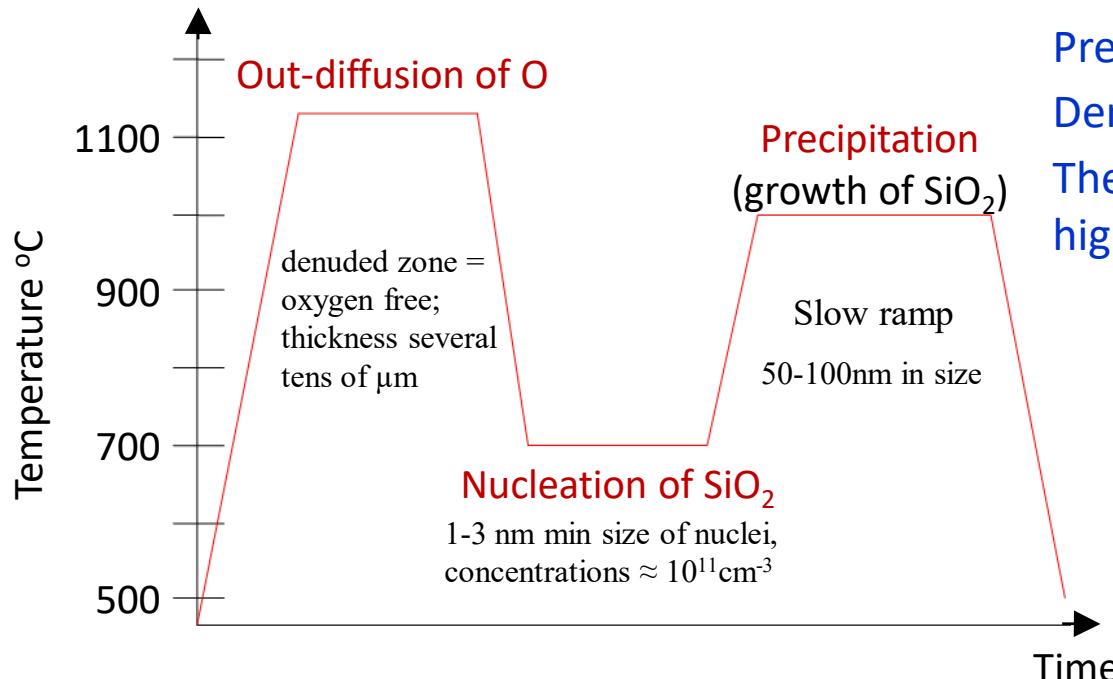
Step 1 generally happens by kicking out the substitutional (s) atom into an interstitial (i) site. One possible reaction is: (I = interstitial Si) $Au_s + I \leftrightarrow Au_i$

Step 2 usually happens easily once the metal is interstitial since most metals diffuse rapidly in this form.

Step 3 happens because heavy metals segregate preferentially to damaged regions (dislocation or stacking fault) or to N⁺ regions, or pair with effective getters like P (AuP).

Step 1 can be facilitated by introducing large amount of Si interstitials, by such as high density phosphorus diffusion, ion implantation damage or SiO₂ precipitation.

Intrinsic gettering



Precipitates (size) grow @ high T
Density of nucleation sites grow @ low T
Therefore, low T to increase density, and high T to grow its size.

Oxygen diffusivity:

$$D_0 = 0.13 \exp\left(\frac{-2.53}{kT}\right) \text{cm}^2 \text{sec}^{-1}$$

$D_0 \gg D_{\text{dopants}}$ but $D_0 \ll D_{\text{metals}}$

In intrinsic gettering, the metal atoms segregate to dislocations (formed because of volume mismatch of SiO_2 and host Si lattice) around SiO_2 precipitates.

15 to 20 ppm oxygen wafers are required:

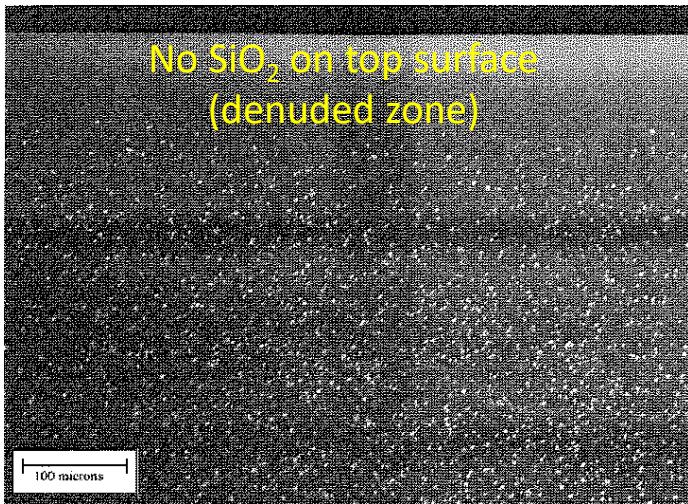
<10 ppm - precipitate density is too sparse to be an effective getterer.

>20 ppm - wafers tend to warp during the high temperature process.

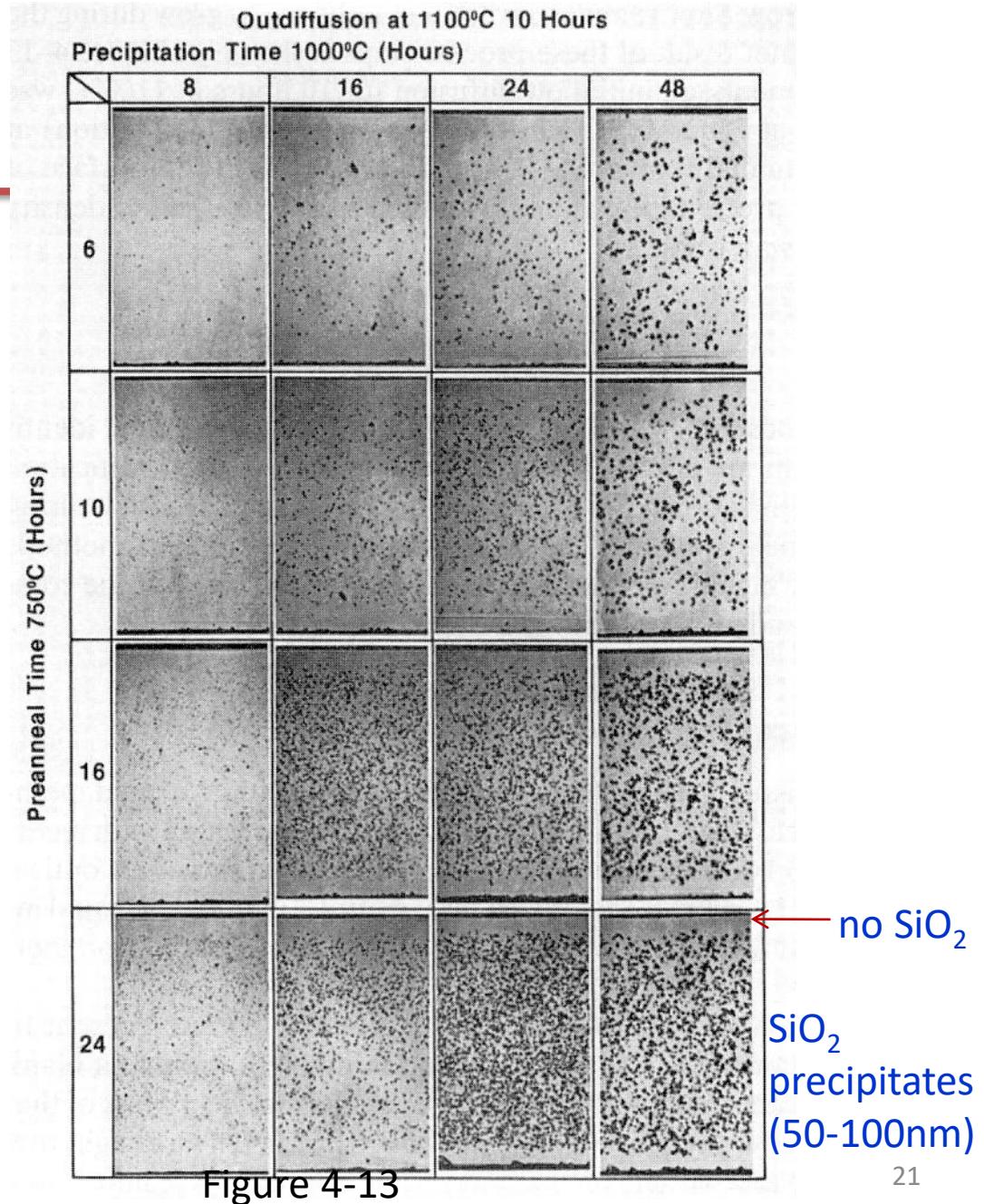
Note: devices that use the entire wafer as the active region (solar cells, thyristors, power diodes, etc...) can not use this technique, but can use extrinsic gettering.

Today, most wafer manufacturers perform this intrinsic gettering task that is better controlled.

Intrinsic Gettering: SiO_2 precipitates



SiO_2 precipitates (white dots) in bulk Si



Chapter 4 Clean room, wafer cleaning and gettering

1. Introduction.
2. Clean room.
3. Wafer cleaning.
4. Gettering.
5. Measurement methods.

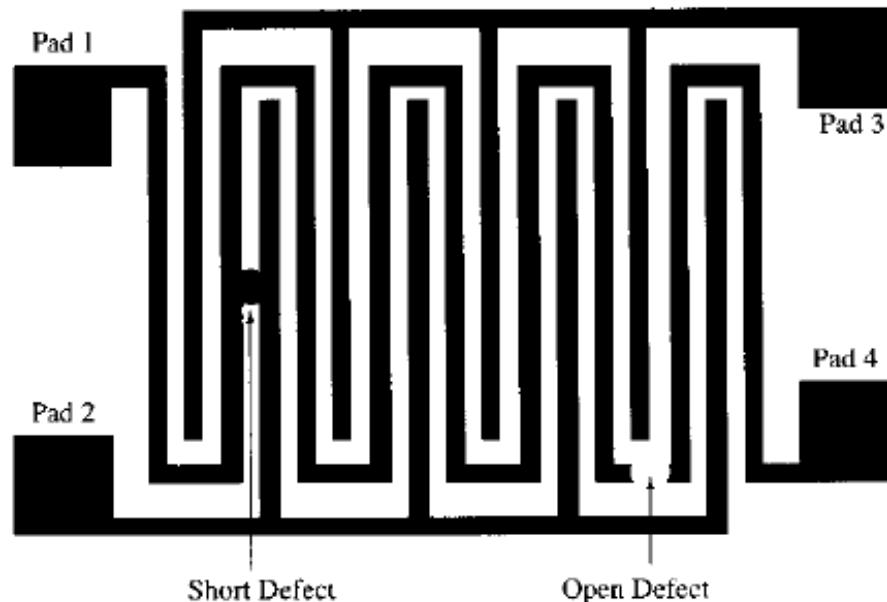
Particle contamination detection

Un-patterned wafers (blank)

- Count particles in microscope
- Laser scanning systems that give maps of particles down to $\approx 0.2 \mu\text{m}$

Patterned wafers

- Optical system compares a die with a “known good reference” die (adjacent die, chip design - its appearance)
- Image processing identifies defects
- Test structure (not in high volume manufacturing)



Test structures design
to detect defects

Figure 4-14 Layout (top view) of a typical electrical test structure designed to detect defects in an interconnect layer. Typical open and short defects are illustrated.

Monitoring the wafer cleaning efficiency

Concentrations of impurities determined by surface analysis.
Excite → Identify (unique atomic signature) → Count concentrations

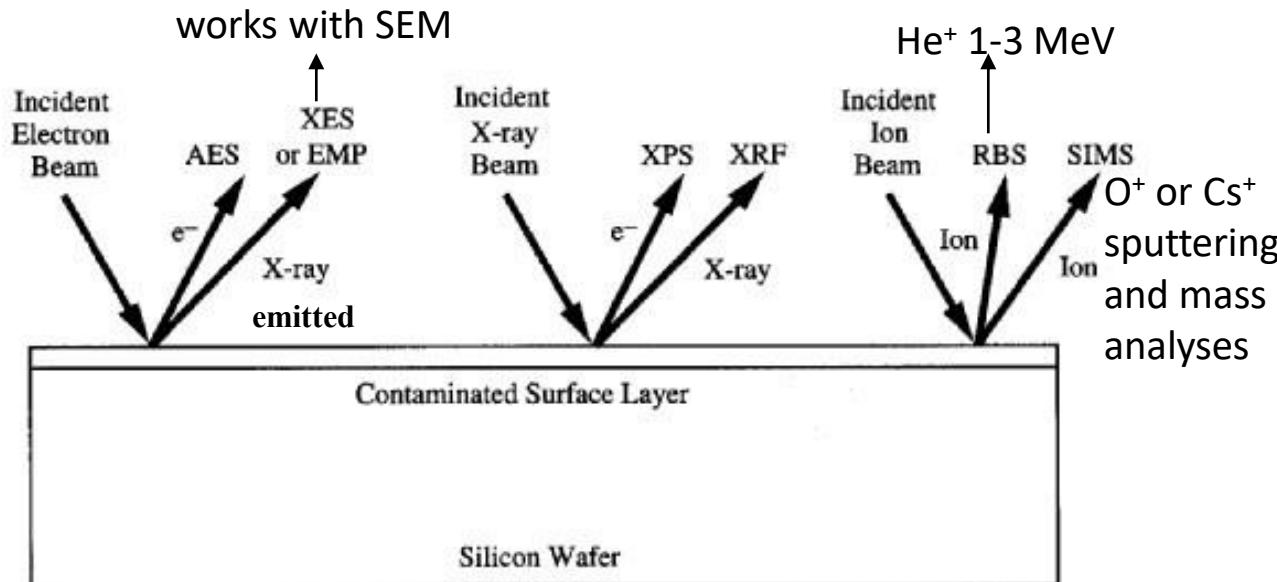


Figure 4-16 Surface analysis techniques used to identify and quantify contamination in IC manufacturing.

Primary beam	electron	good lateral resolution (e can be focused, but not x-ray)
Detected beam	electron	good depth resolution and surface sensitivity
	X-ray	poor depth resolution and poor surface sensitivity
	ions (SIMS)	excellent
	ions (RBS)	good depth resolution, reasonable sensitivity (0.1 atomic%)

Chapter 5 Lithography

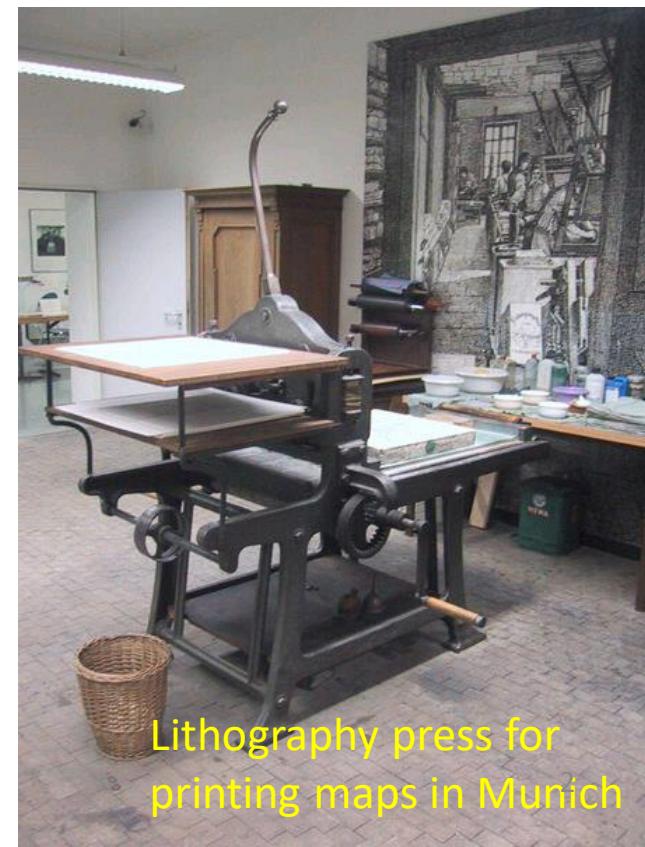
1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

History

- Historically, lithography is a type of printing technology that is based on the chemical repellence of oil and water.
- Photo-litho-graphy: *latin*: light-stone-writing.
- In 1826, Joseph Nicephore Niepce in Chalon France takes the first photograph using bitumen of Judea on a pewter plate, developed using oil of lavender and mineral spirits.
- In 1935 Louis Minsk of Eastman Kodak developed the first negative photoresist.
- In 1940 Otto Suess developed the first positive photoresist.
- In 1954, Louis Plambeck, Jr., of Du Pont, develops the Dycryl polymeric letterpress plate.



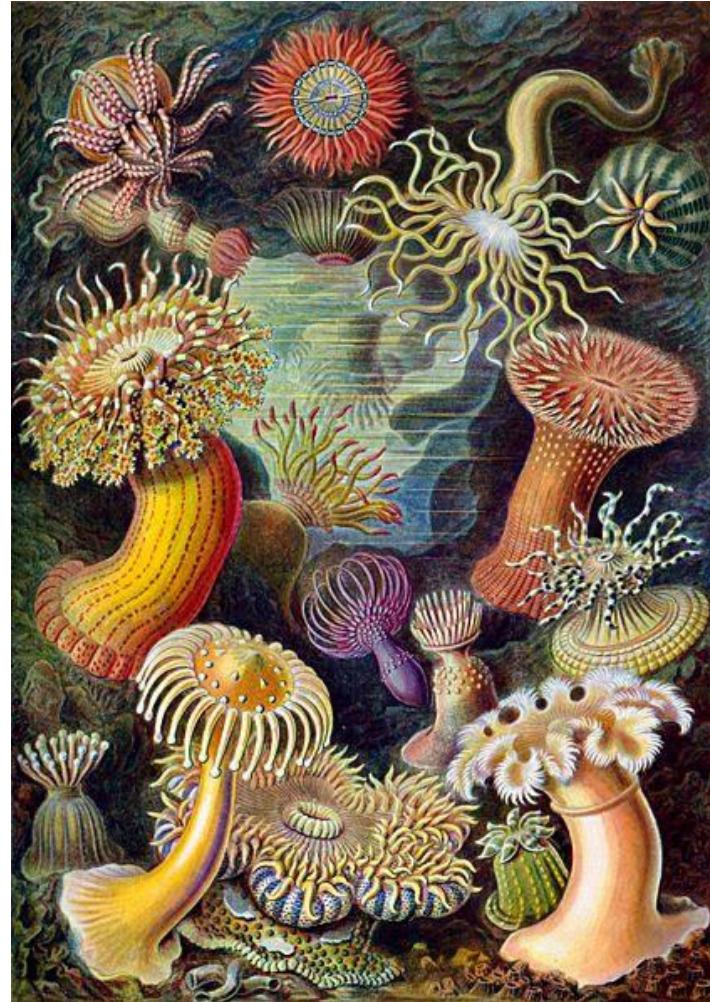
Lithography stone and mirror-image print of a map of Munich.



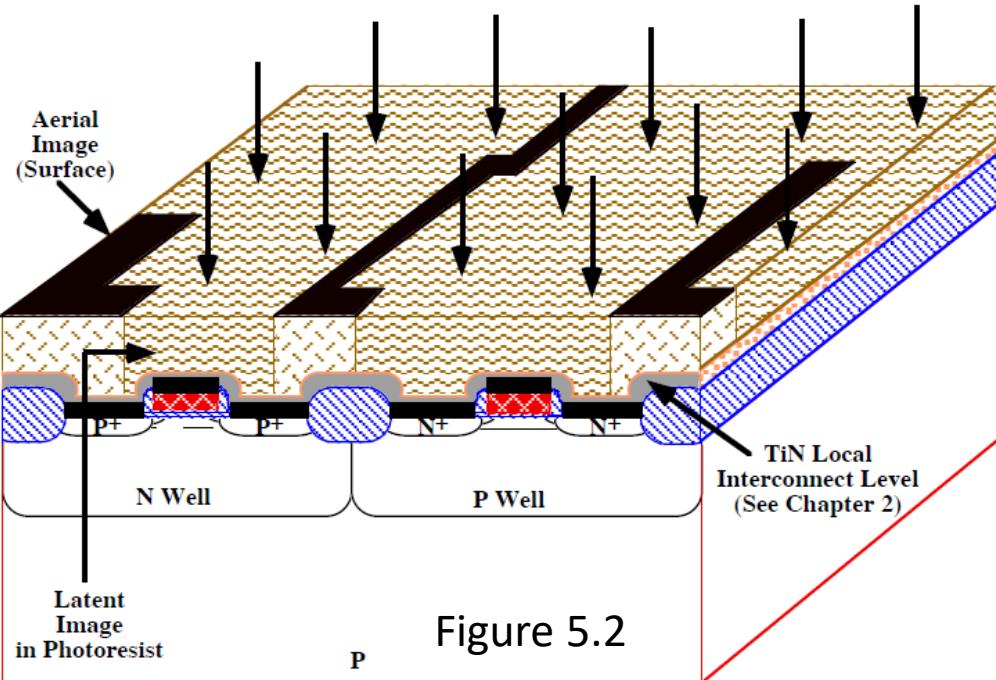
Lithography press for printing maps in Munich

Lithography for art: the print principle

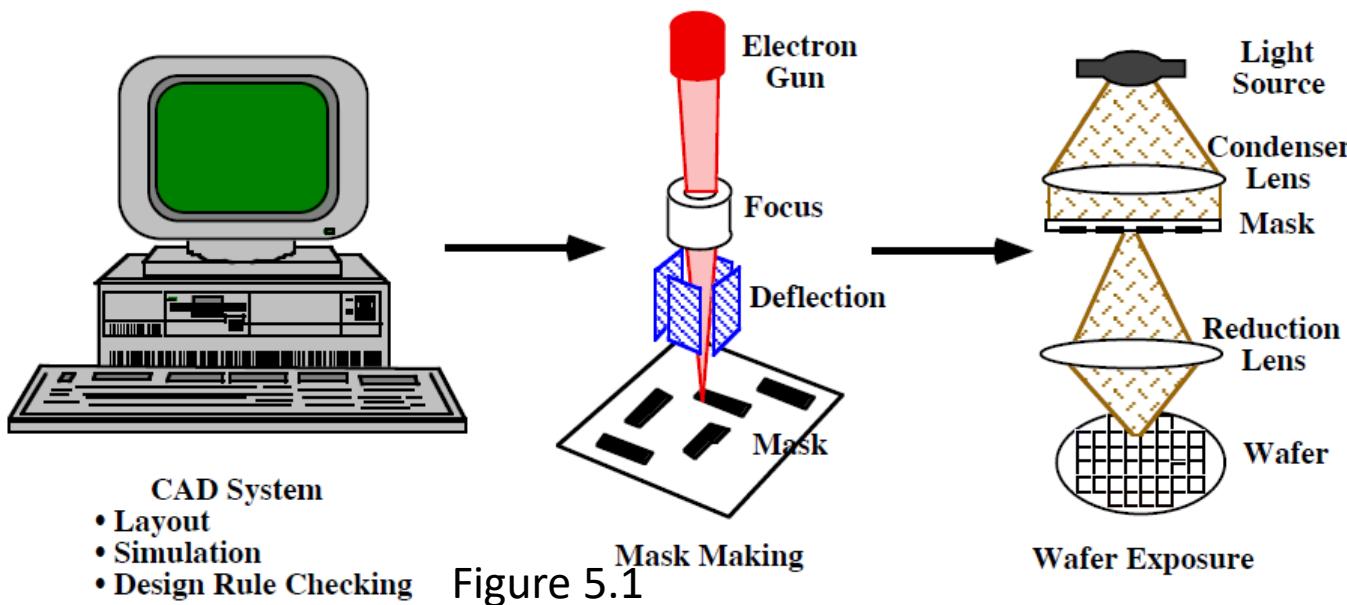
- Lithography is a printing process that uses chemical processes to create an image.
- For instance, the positive part of an image would be a hydrophobic chemical, while the negative image would be water.
- Thus, when the plate is introduced to a compatible ink and water mixture, the ink will adhere to the positive image and the water will clean the negative image.



Photolithography for IC manufacturing



- In IC manufacturing, lithography is the single most important technology.
- ~35% of wafer manufacturing costs comes from lithography.
- The SIA roadmap is driven by the desire to continue scaling device feature sizes.
- Placement/alignment accuracy ~1/3 of feature size.



Patterning process consists of:
Mask design
Mask fabrication
Wafer exposure

Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Light source: mercury arc lamp

Traditionally Hg vapor lamps have been used which generate many spectral lines from a high intensity plasma inside a glass lamp.

Electrons are excited to higher energy levels by collisions in the plasma, and photons are emitted when the energy is released. (electron effective temperature 40000K in a plasma!!)

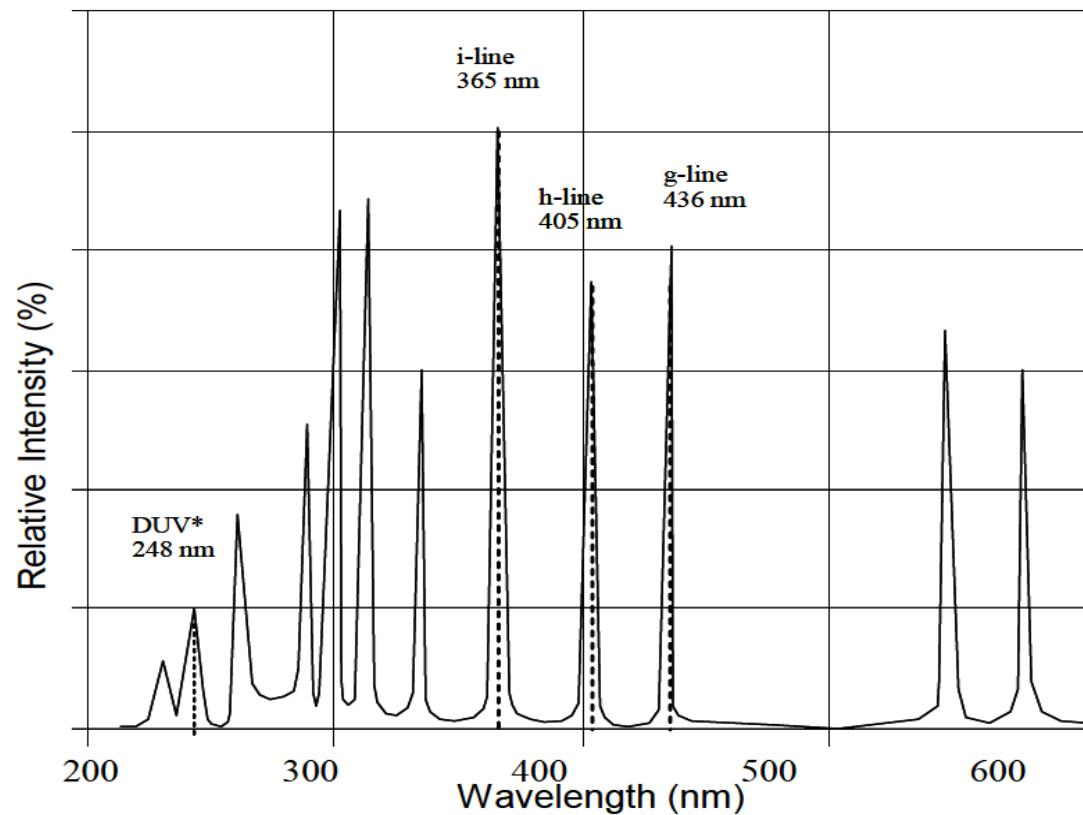
g line $\lambda=436\text{ nm}$

i line $\lambda=365\text{ nm}$

(used for $0.5\mu\text{m}$ and $0.35\mu\text{m}$ lithography generation)



High pressure Hg-vapor lamps
Order \$1000, lasts ~1000 hours.



- Filters can be used to limit exposure wavelengths.
- Intensity uniformity has to be better than several % over the collection area.
- Needs spectral exposure meter for routine calibration due to aging.

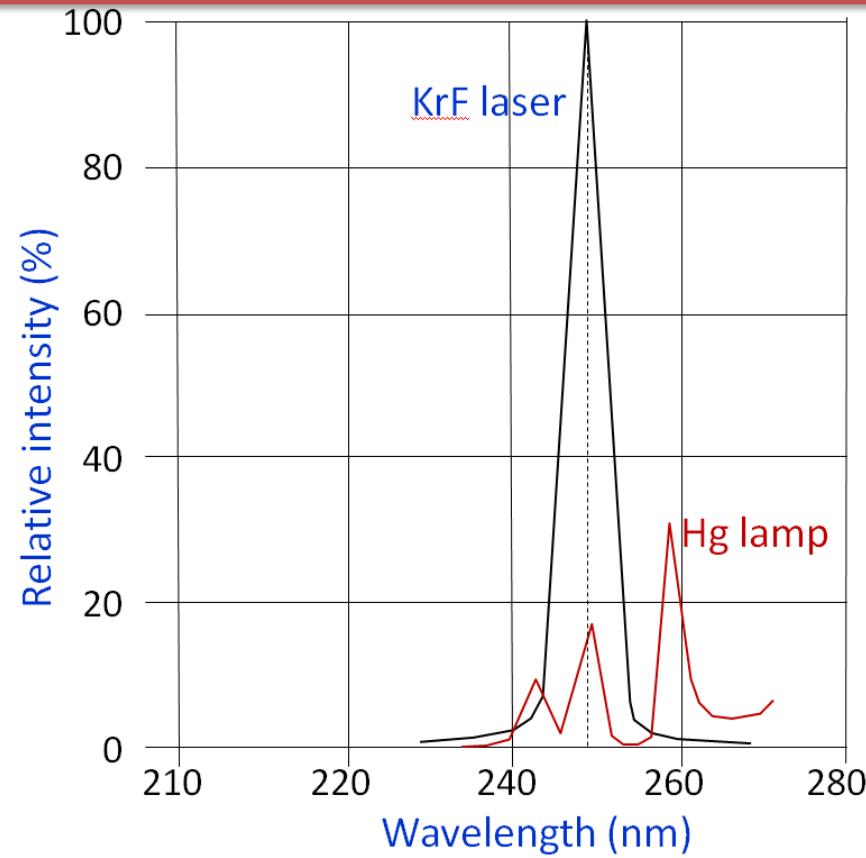
Light source: excimer laser

Decreasing feature size (to $<0.35\mu\text{m}$) requires shorter λ .

Brightest sources in deep UV are excimer lasers.

Excimer laser:

- In excimer lasers, two elements, e.g. a noble gas and a halogen (from a halogen containing compound), which can react and “bind” together only in the excited state but not in their ground states, are present.
- Providing energy will therefore drive the reaction, creating the excimer.
- When the excitation energy is removed, the excimer dissociates and releases the energy at the characteristic wavelength.
- A pulsed excitation is used to repeat the process.



KrF $\lambda = 248\text{ nm}$ (used for $0.25\mu\text{m}$ lithography generation)

ArF $\lambda = 193\text{ nm}$ (currently used for 45nm node/generation production)

Eximer = Excited dimer
 $\text{Xe}^* + \text{Cl}_2 \rightarrow \text{XeCl}^* + \text{Cl}$
 $\text{XeCl}^* \rightarrow \text{XeCl} + \text{DUV}$
DUV = deep UV, 308nm for XeCl laser
 $\text{XeCl} \rightarrow \text{Xe} + \text{Cl}$
Here “*” means excited state

Light sources: summary

Material	Wavelength h (nm)	Max. Output (mJ/pulse)	Frequency (pulses/sec)	Pulse Length (ns)	CD Resolution (μm)
KrF	248	300 – 1500	500	25	≤ 0.25
ArF	193	175 – 300	400	15	≤ 0.18
F ₂	157	6	10	20	≤ 0.15

CD: critical dimension

Source	λ (nm)	Max. Output (mJ/pulse)	Frequency (pulses/s)	Common Nomenclature
Hg lamp	436			G line
Hg lamp	405			H line
Hg lamp	365			I line
KrF (laser)	248	10	2000	DUV
ArF (laser)	193	10	2000	193 DUV
F ₂ (laser)	157	40	500	

Note: the numbers in the two tables are different, so they must be for different systems

Photomask

Types:

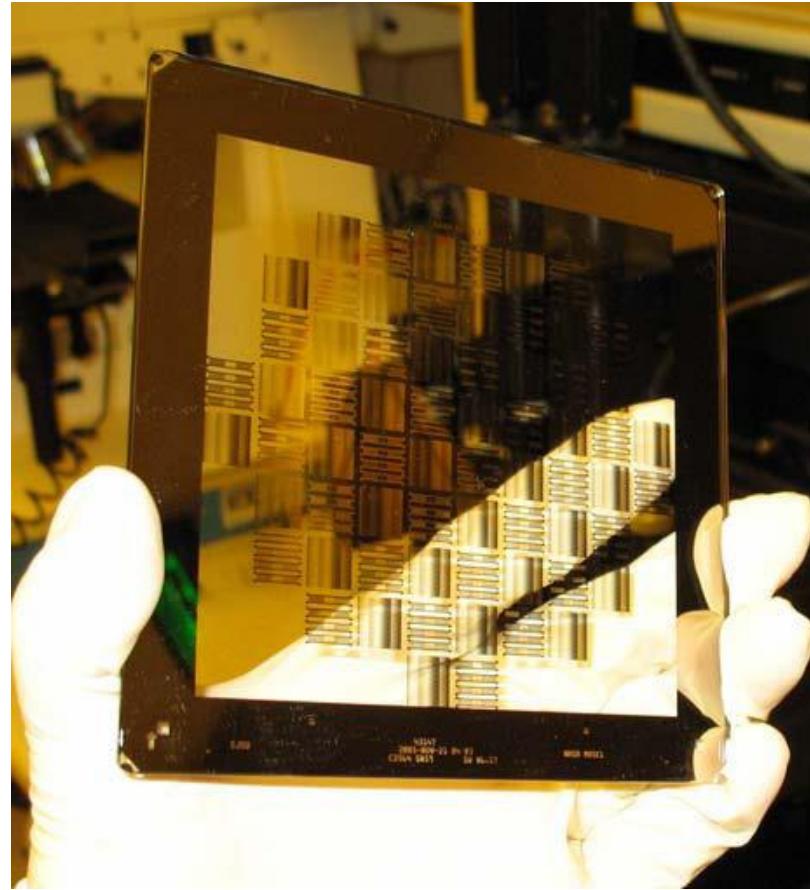
- Photographic emulsion on soda lime glass (cheap).
- Cr on soda lime glass and on quartz glass (most popular).
(Quartz has low thermal expansion coefficient and low absorption of light, but more expensive; needed for deep UV lithography).
- Transparency by laser printer, more and more popular for MEMS (resolution down to few μm with a 20000 dpi printer, very cheap).

Polarity:

- Light-field, mostly clear, drawn feature is opaque.
- Dark-field, mostly opaque, drawn feature is clear.

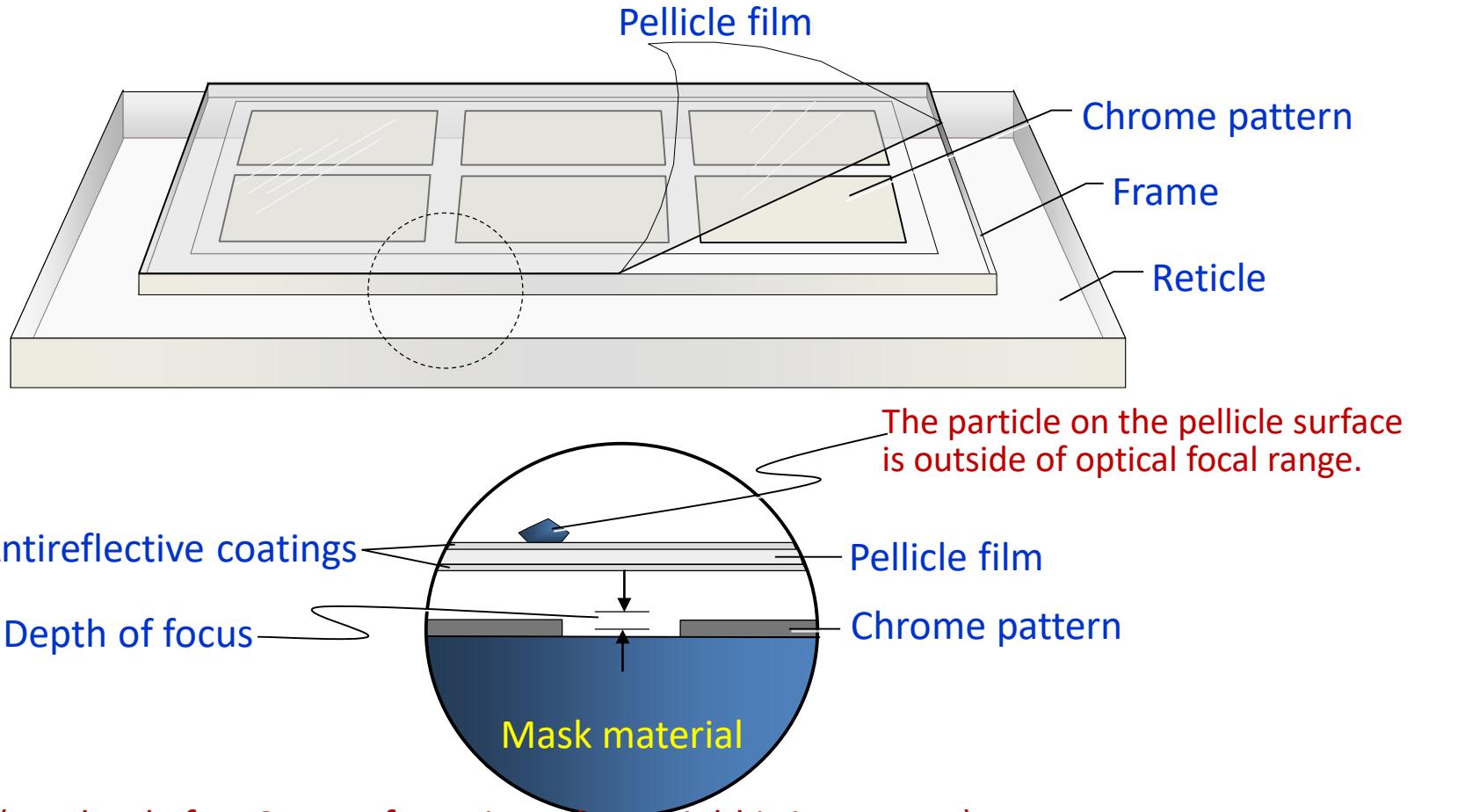
Three potential mask improvements:

Pellicle, antireflective coatings, phase-shift masks.
(we want 100% transmission, no reflection)



Light-field photomask

Pellicle on a reticle (IC word for mask)

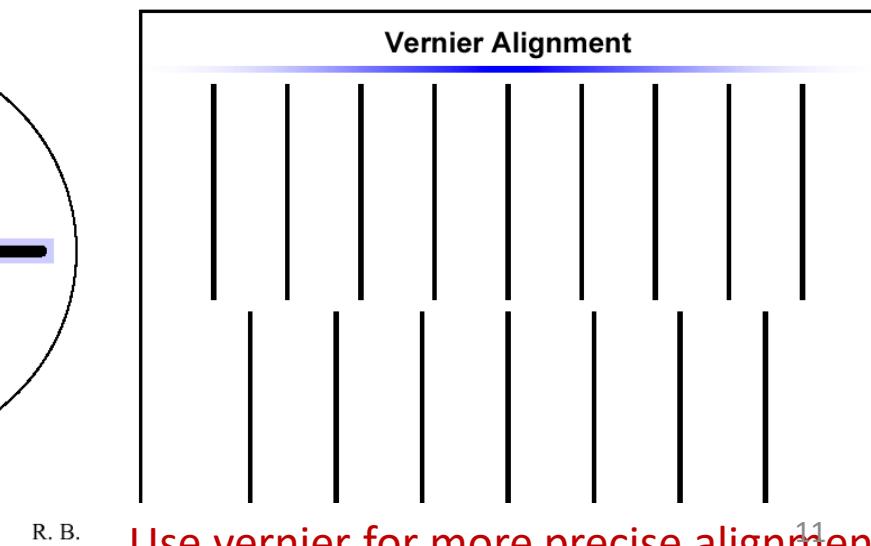
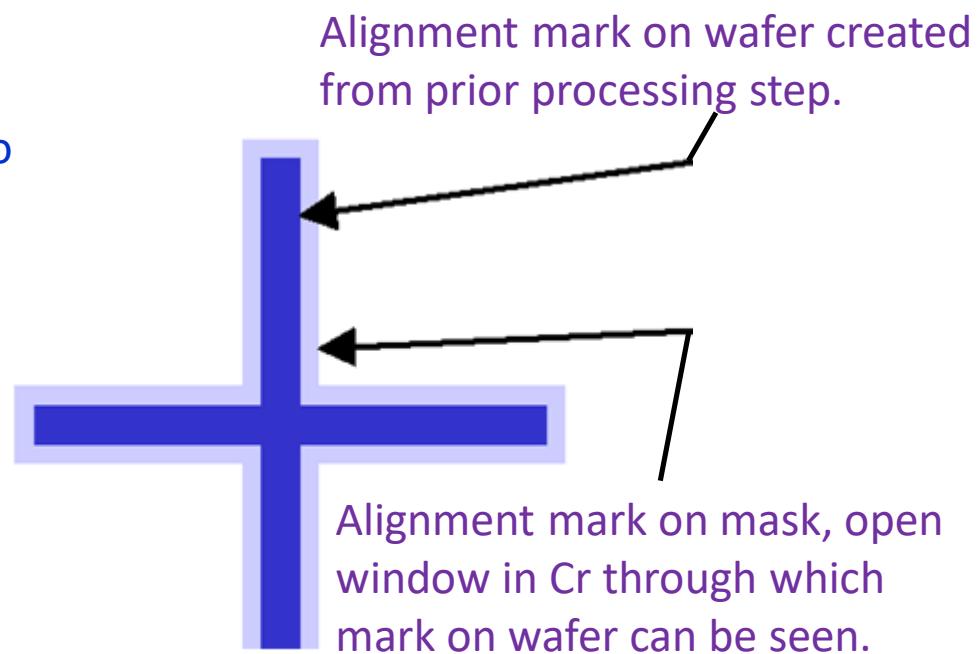
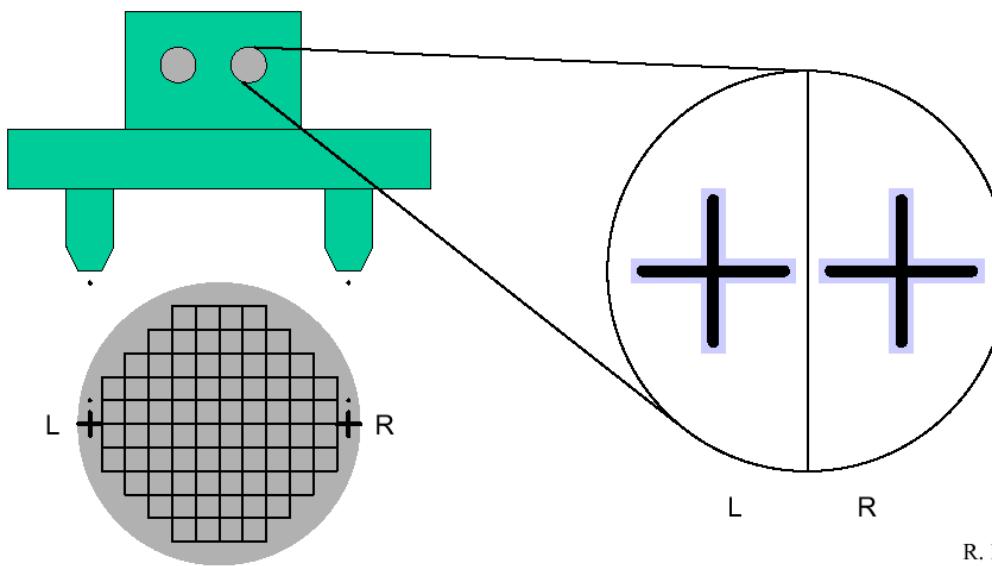


Pellicle: (used only for IC manufacturing where yield is important)

- A thin coating of transparent material similar to Mylar is stretched over a cylindrical frame on either side of the mask.
- The frame stands off the membrane at a distance of **~1 cm** from the surface of the mask.
- Purpose of pellicle is to ensure that particles that fall in the mask are kept outside of the focal plane of the optical system.

Mask to wafer alignment

- 3 degrees of freedom between mask and wafer: x , y , ϕ (angle)
- Use alignment marks on mask and wafer to register patterns prior to exposure.
- Modern steppers use automatic pattern recognition and alignment systems, which takes 1-5 sec to align and expose.
- Normally requires at least two alignment mark sets on opposite sides of wafer or stepped region, and use a split-field microscope to make alignment easier.



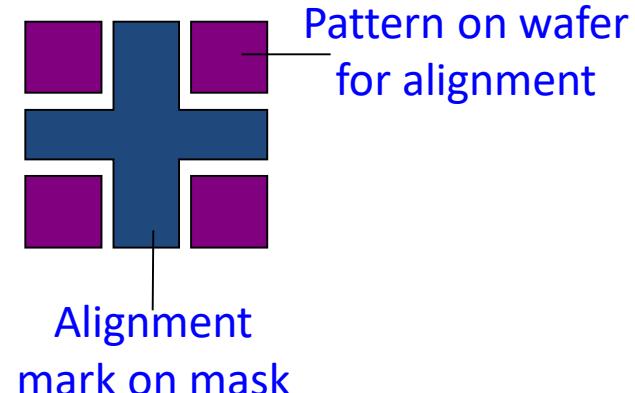
Alignment problems: thermal expansion

$$R = r \cdot (\Delta T_m \cdot \alpha_m - \Delta T_{si} \cdot \alpha_{si})$$

↗ ↑
run-out wafer
error radius

$\Delta T_m, \Delta T_{si}$ = change of mask and wafer temperature.

α_m, α_{si} = coefficient of thermal expansion of mask & silicon.



For example, for thermal expansion of 2 ppm/ $^{\circ}\text{C}$ (silicon 2.6, fused silica/quartz 0.5 ppm/ $^{\circ}\text{C}$), assume temperature change of 1 $^{\circ}\text{C}$, then the distance between two features separated by 50mm will change by 2 ppm or 100nm, which is too large for IC production but OK for most R&D.



Chapter 5 Lithography

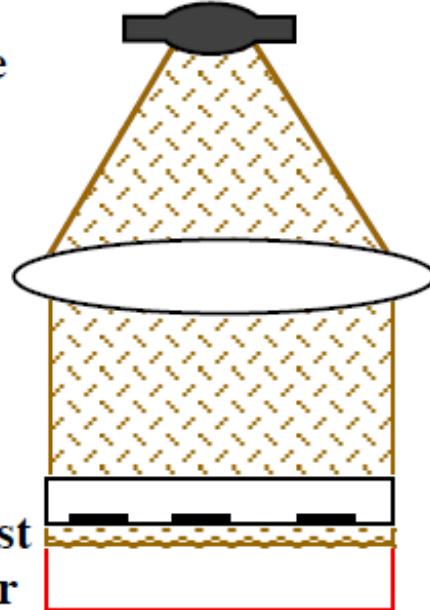
1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Three basic methods of wafer exposure

1:1 Exposure Systems

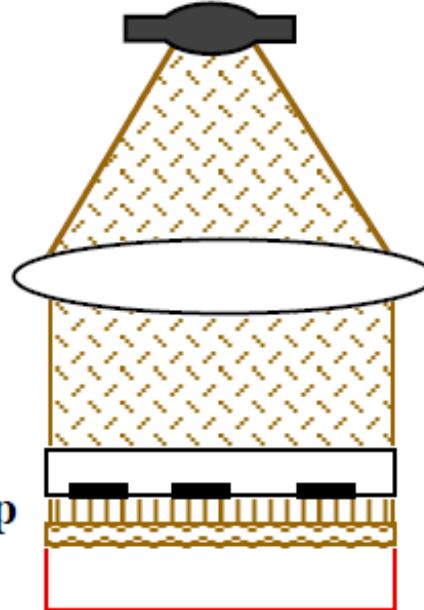
Usually 4X or 5X Reduction

Light Source



Optical System

Gap



Mask
Photoresist
Si Wafer

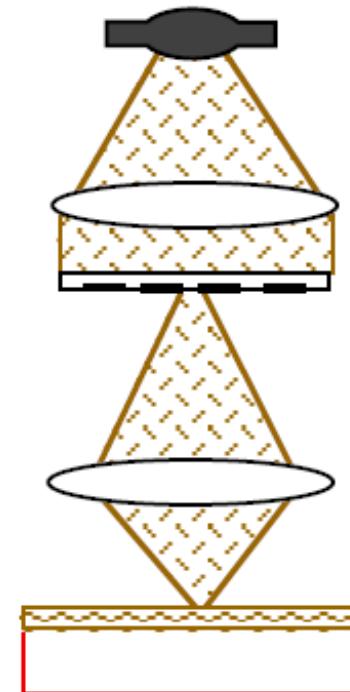
Figure 5.3 Contact Printing

High resolution. But mask wear, defect generation.

Fast, simple and inexpensive, choice for R&D.

Proximity Printing

Less mask wear /contamination, less resolution (depend on gap).



Projection Printing

No mask wear/contamination, mask de-magnified 4x (resist features 4x smaller than mask). Very expensive, mainly used for IC industry.

Contact/proximity exposure system (called mask aligner)

Hard to maintain contact or constant gap when wafer/mask is not even/flat.

Resolution (half-period for grating pattern) is given by:

$$R = \frac{3}{2} \sqrt{\lambda \left(g + \frac{t}{2} \right)}$$

g is gap (=0 for contact), t is resist thickness, and λ is wavelength.

4 objectives of optical exposure system

- Collect as much of radiation
- Uniform radiation over field of exposure
- Collimate and shape radiation
- Select exposure wavelength

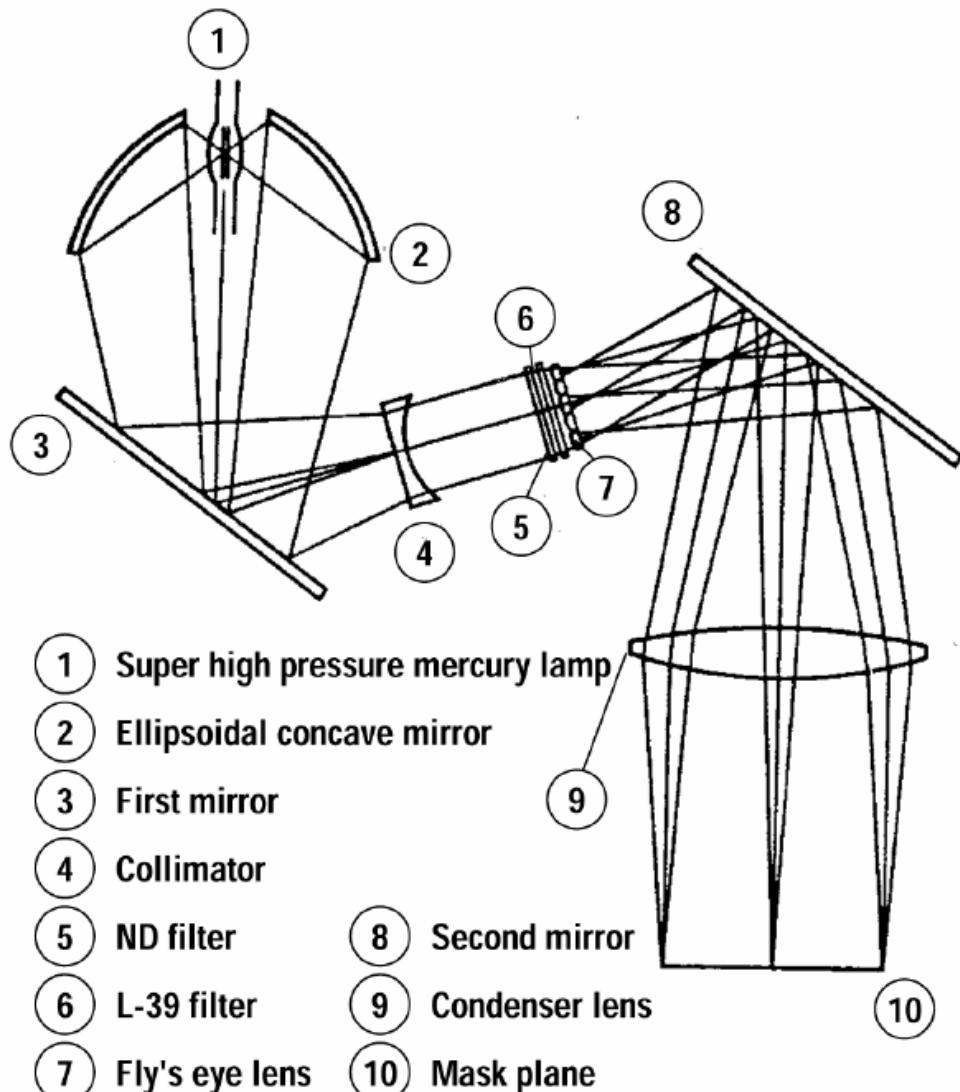


Figure 7.12 Schematic of a typical source assembly for a contact/proximity printer (after Jain).

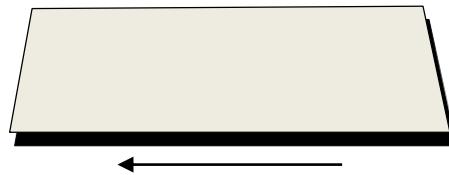
Stepper (step and repeat system)

Die-by-die exposure

Feature size (typically)

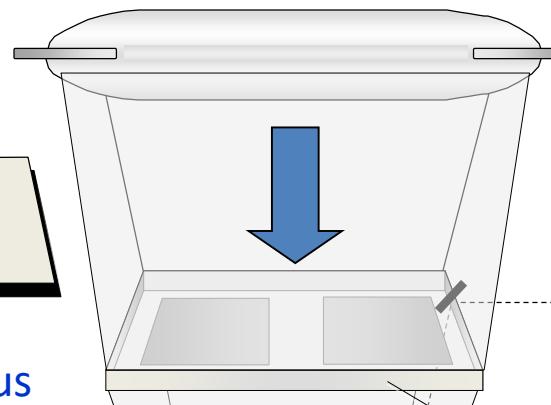
4x reduction

Shutter



Shutter is closed during focus
and alignment and removed
during wafer exposure

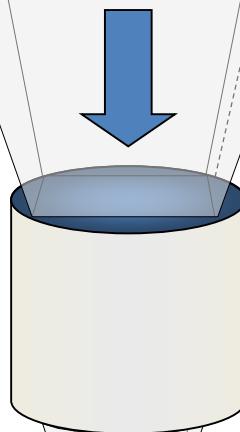
UV light source



Alignment laser

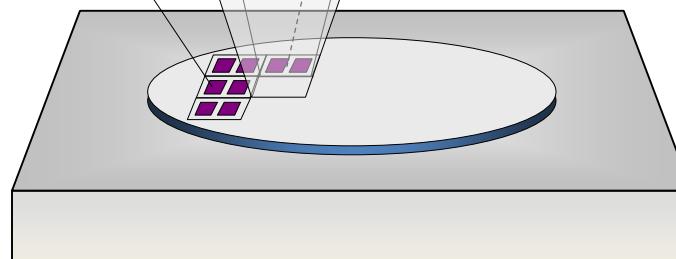


Reticle (may contain one or
more die in the reticle field)



Projection lens (reduces the size
of reticle field for presentation to
the wafer surface)

Single field exposure, includes:
focus, align, expose, step, and
repeat process

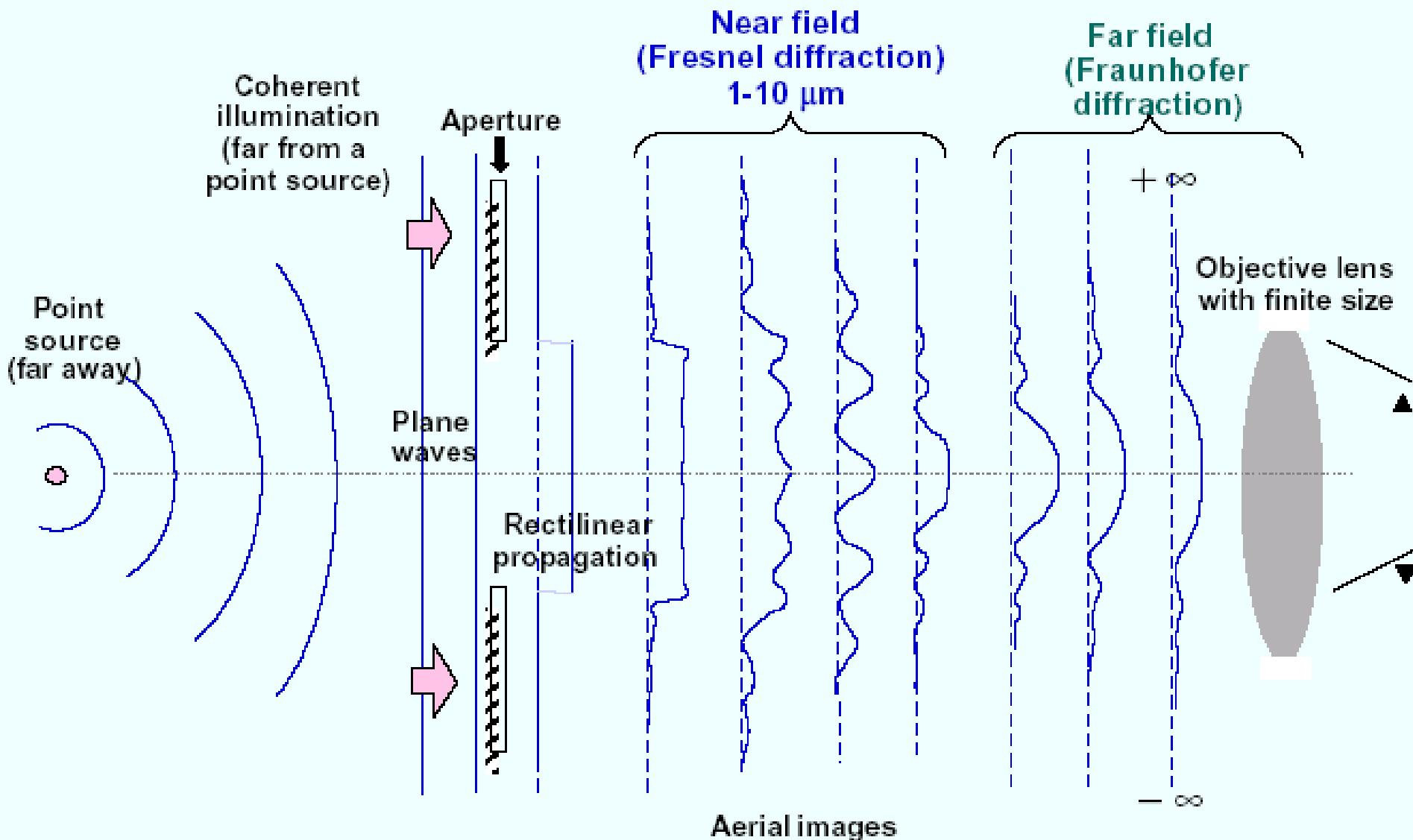


Wafer stage controls
position of wafer in
 X , Y , Z , ϕ

Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Light diffraction through an aperture on mask



Near field/Fresnel diffraction for contact/proximity exposure

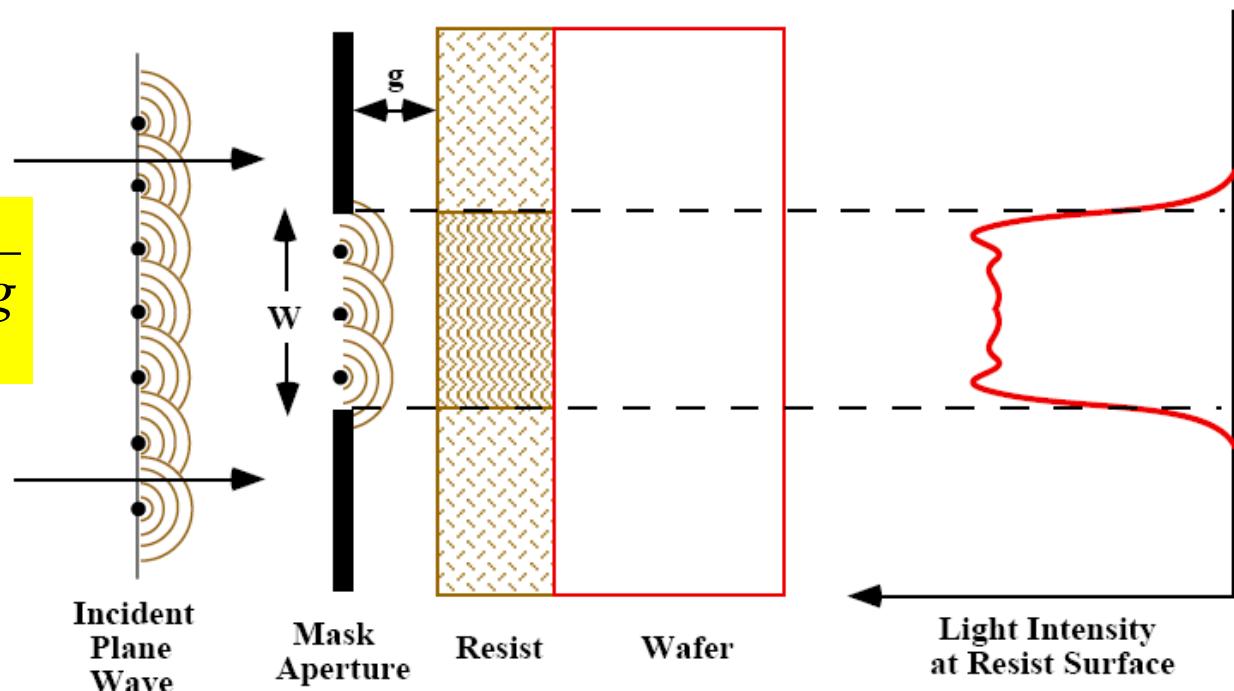
Near field: $\lambda < g < \frac{W^2}{\lambda}$
(g is gap)

$$W_{\min} \approx \frac{3}{2} \sqrt{\lambda \left(g + \frac{t}{2} \right)} \sim \sqrt{\lambda g}$$

(t is resist thickness)

For $g=10\mu\text{m}$, $\lambda=365\text{nm}$

$$W_{\min} \approx 2 \mu\text{m}$$



- Interference effects and diffraction result in “ringing” and spreading outside the aperture.
- Edges of image rise gradually (not abrupt) from zero.
- Intensity of image oscillates about the expected intensity.
- Oscillations decay as one approaches the center of the image.
- The oscillations are due to constructive and destructive interference of Huygen’s wavelets from the aperture in the mask.
- When aperture width is small, the oscillations are large
- When aperture width is large, the oscillations rapidly die out, and one approaches simple ray tracing when aperture $\gg \lambda$.

Far field/Fraunhofer diffraction for projection exposure

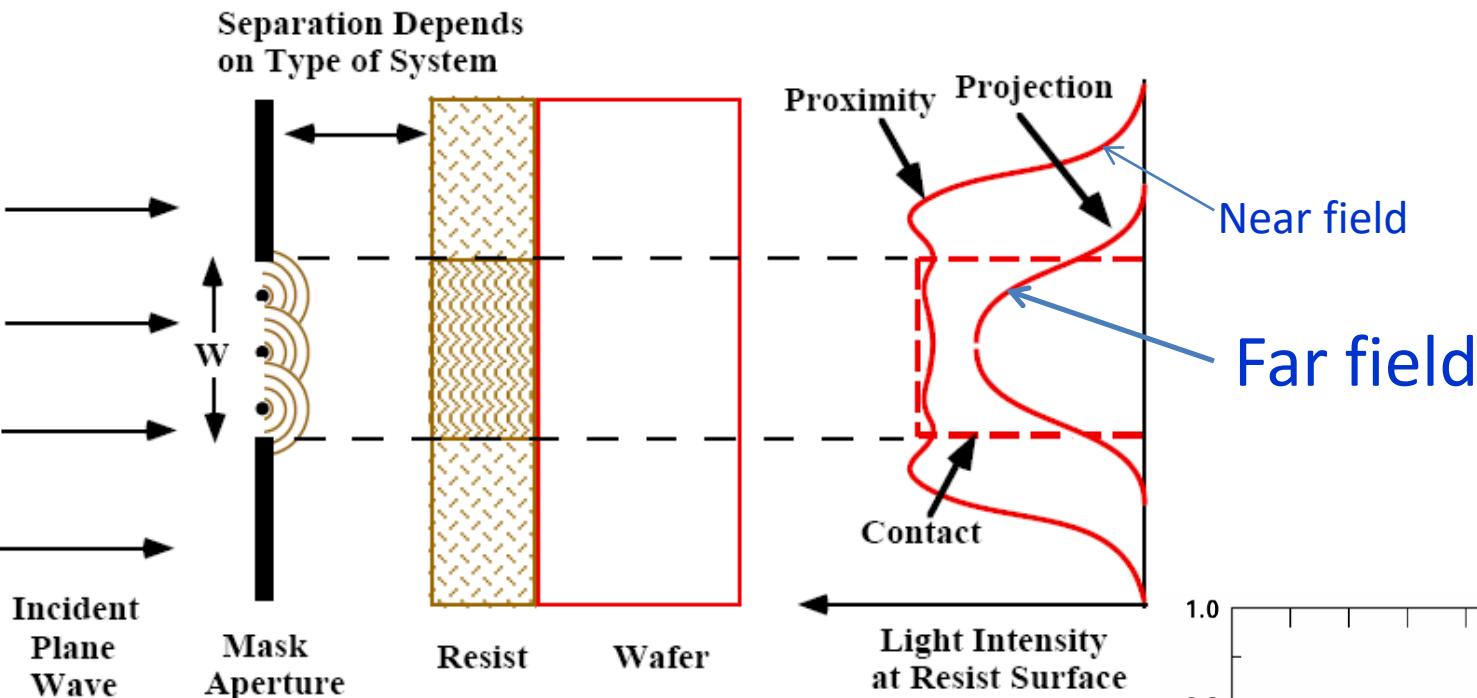
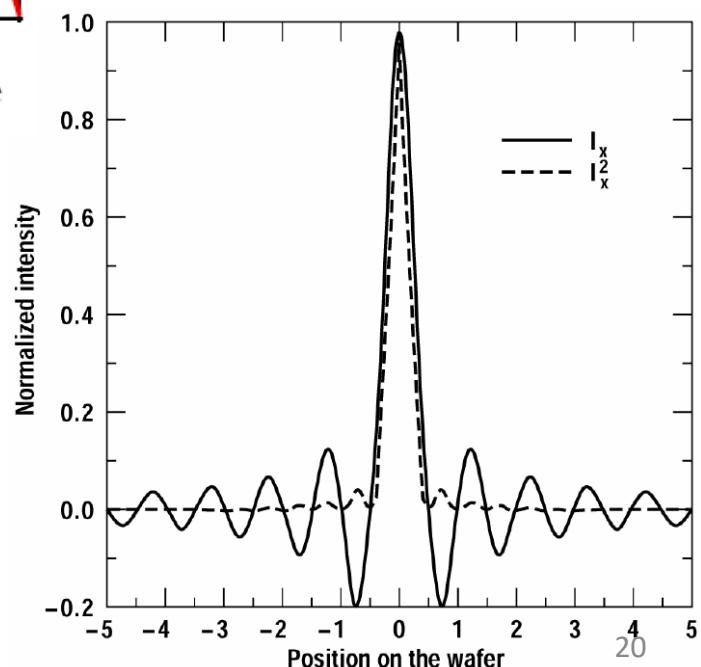


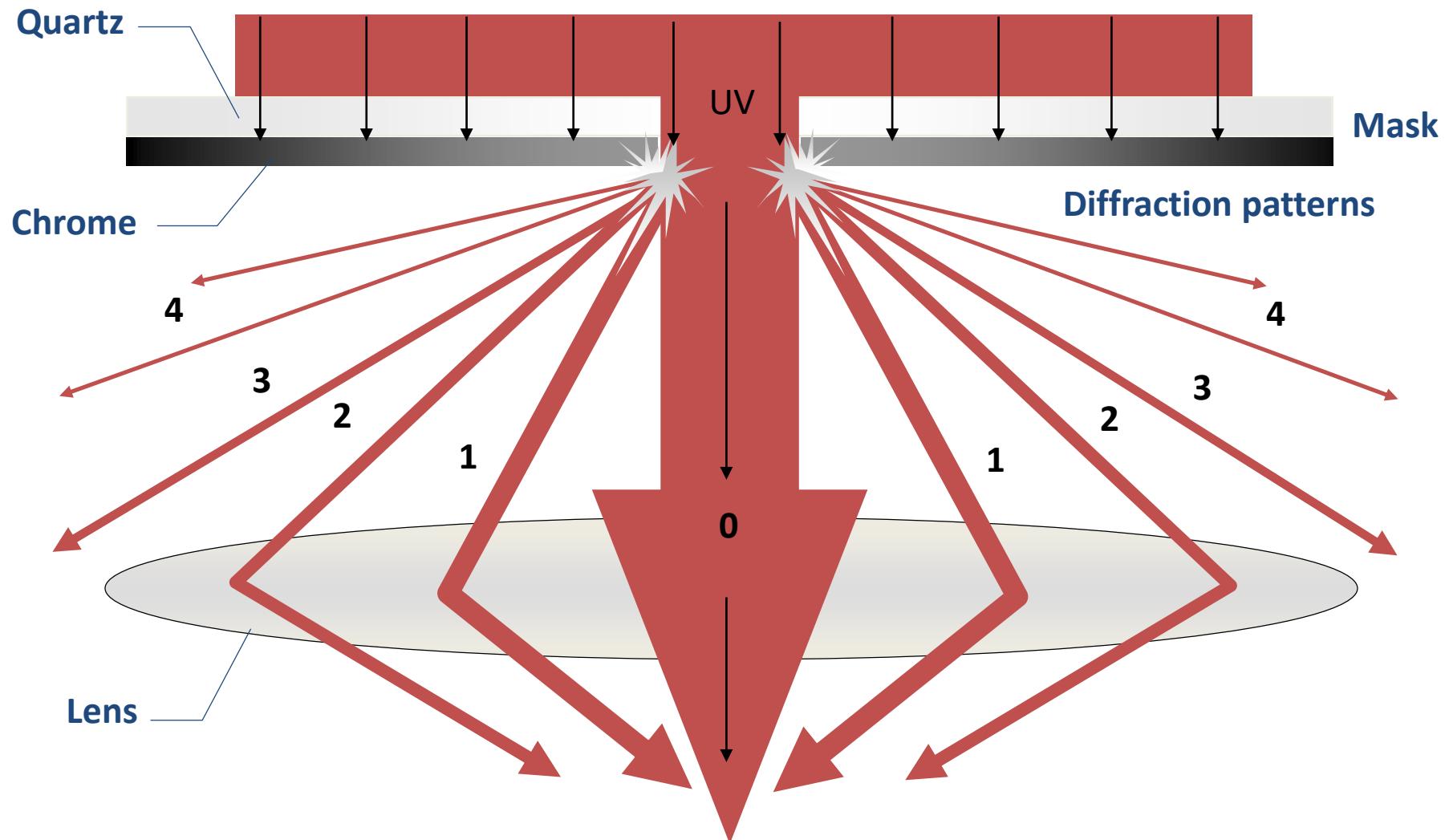
Figure 5.15

Far field: $W^2 \ll \lambda(g^2+r^2)^{1/2}$, r is position on the wafer.

Sharp maximum intensity at $x=0$, and intensity goes through 0 at integer multiples of one-half number.

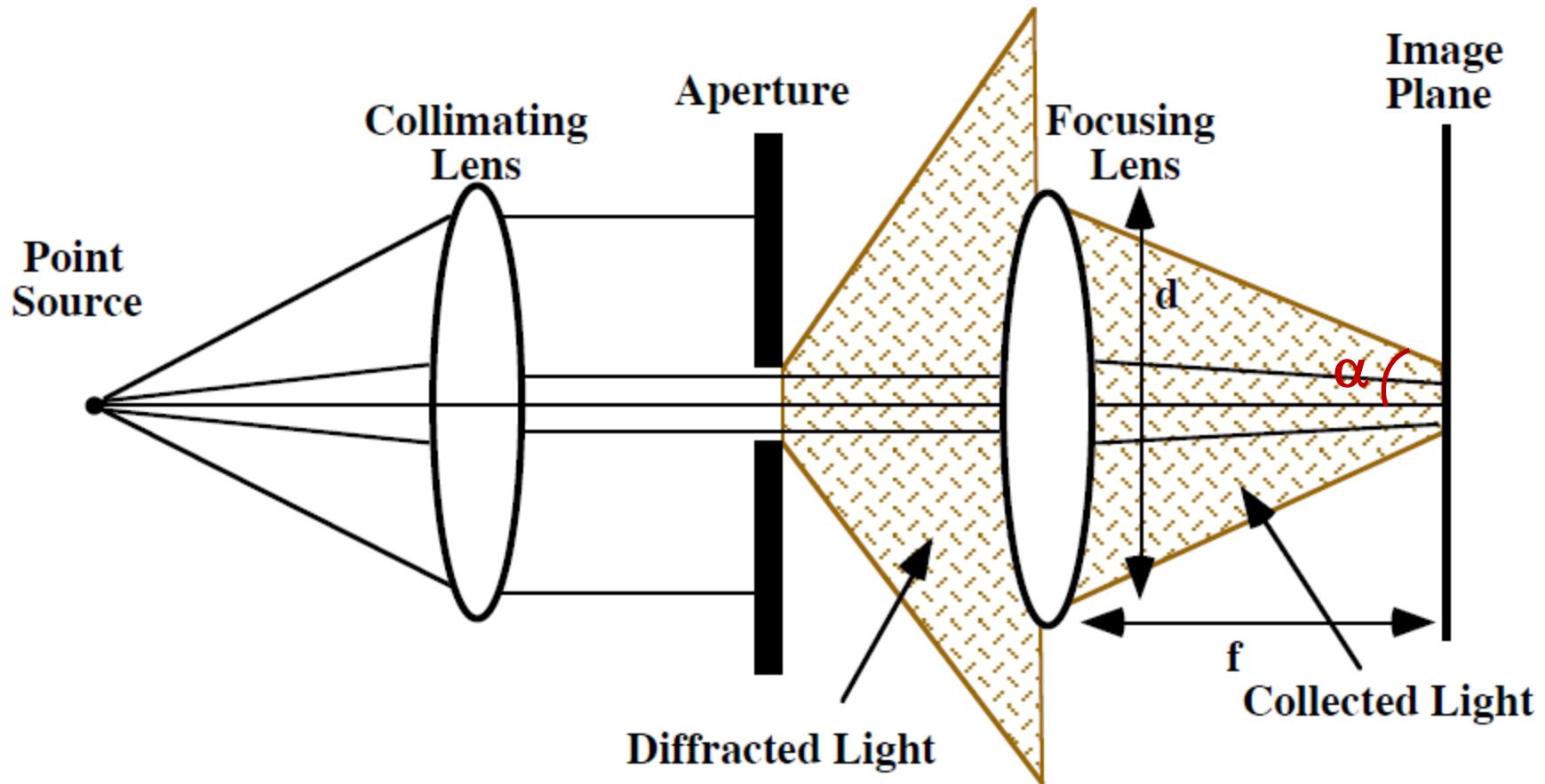


Lens capturing diffracted light



Large lens captures more diffracted light, and those higher order diffracted light carries high frequency (detail of fine features on mask) information.

Numerical aperture of a lens

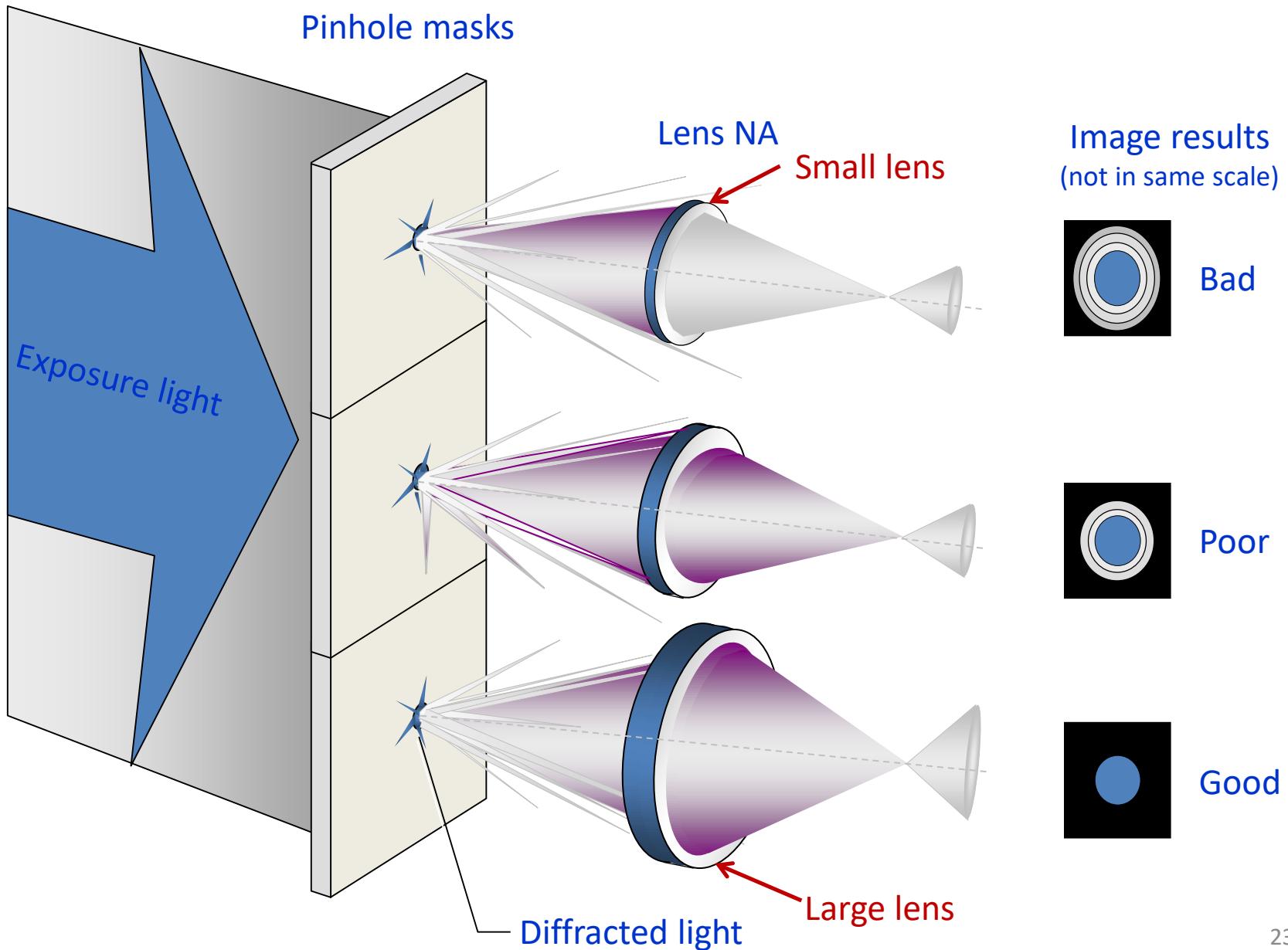


Numerical aperture (NA) of an optical system is a measure of the ability of the lens to collect light.

$NA \equiv n \sin \alpha$, n is refractive index for the medium at the resist surface (air, oil, water).

For air, refractive index $n=1$, $NA = \sin \alpha \approx (d/2)/f \propto d$ for small α .

Effect of numerical aperture on imaging



Light diffraction through a small circular aperture

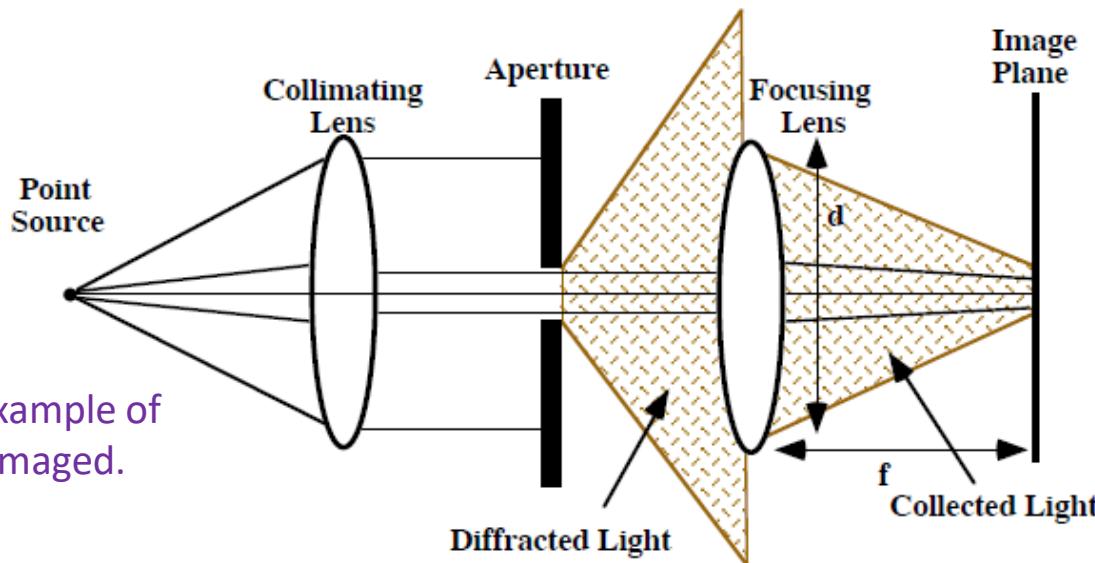


Figure 5.6 Qualitative example of a small aperture being imaged.

Light intensity on image plate

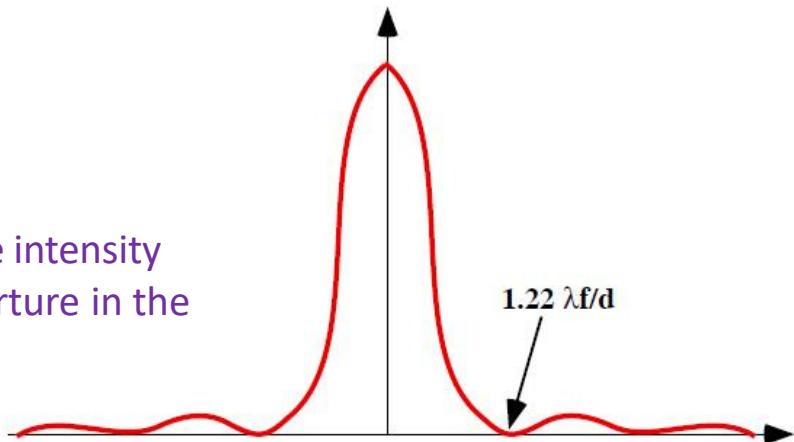
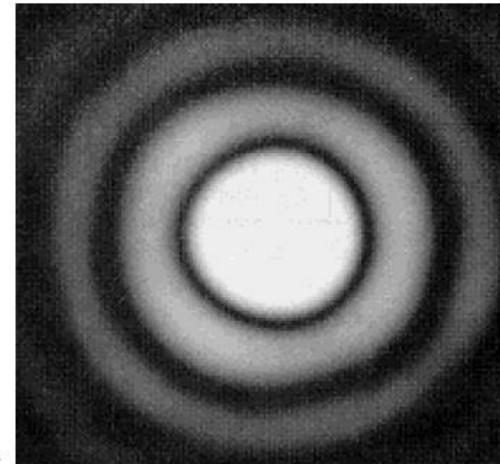


Figure 5.7 Image intensity of a circular aperture in the image plane.

"Airy disk"

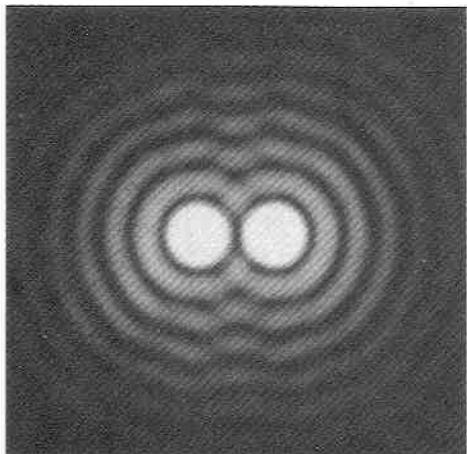


A point image is formed only if $\lambda \rightarrow 0$, $f \rightarrow 0$ or $d \rightarrow \infty$.

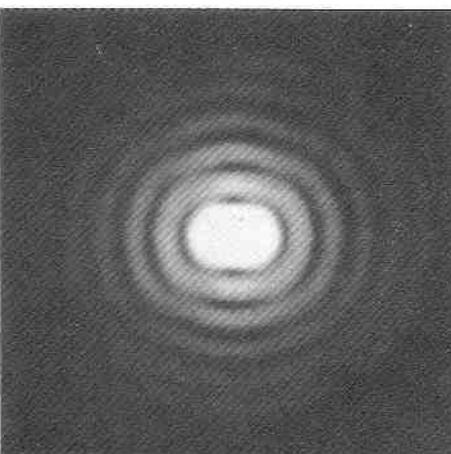
Rayleigh criteria for resolution



Lord Rayleigh



Resolved images



Unresolved images

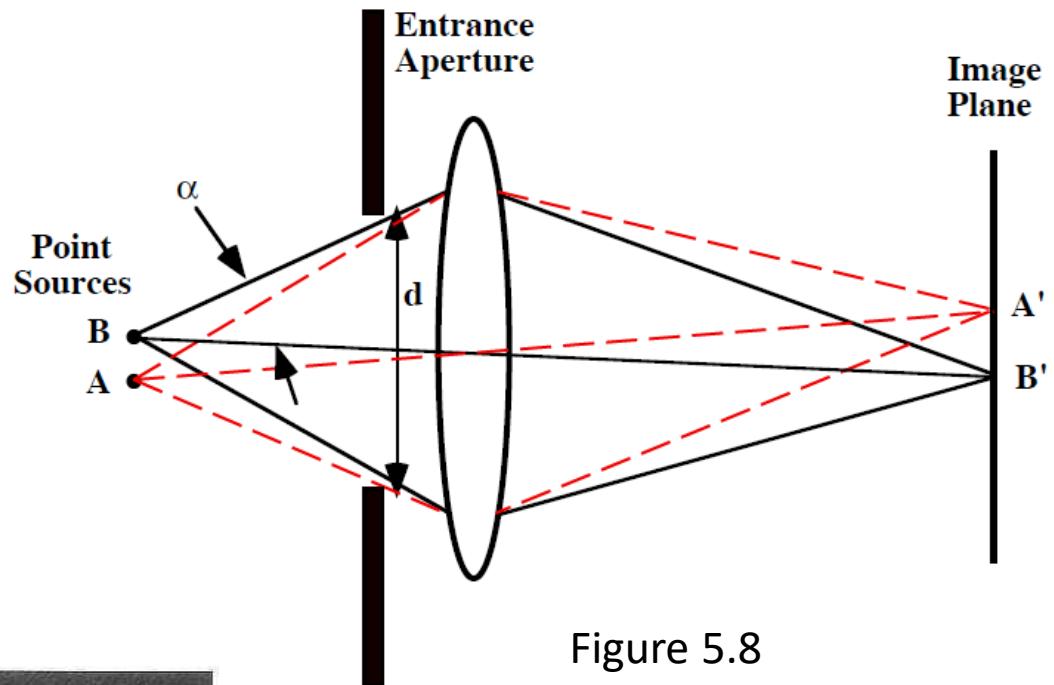


Figure 5.8

Rayleigh suggested that a reasonable criterion for resolution is that the central maximum of each point source lie at the first minimum of the Airy pattern.

Strictly speaking, this and next slides make sense only for infinitely far ($\gg f$) objects, like eye. Fortunately, 4x reduction means far object, and near (near focal plane) image.

Rayleigh criteria for resolution R

$$R = \frac{1.22\lambda f}{d} = \frac{1.22\lambda f}{n(2f \sin \alpha)} = \frac{0.61\lambda}{n \sin \alpha} = 0.61 \frac{\lambda}{NA} = k_1 \frac{\lambda}{NA}$$

K_1 factor has no well-defined physical meaning.

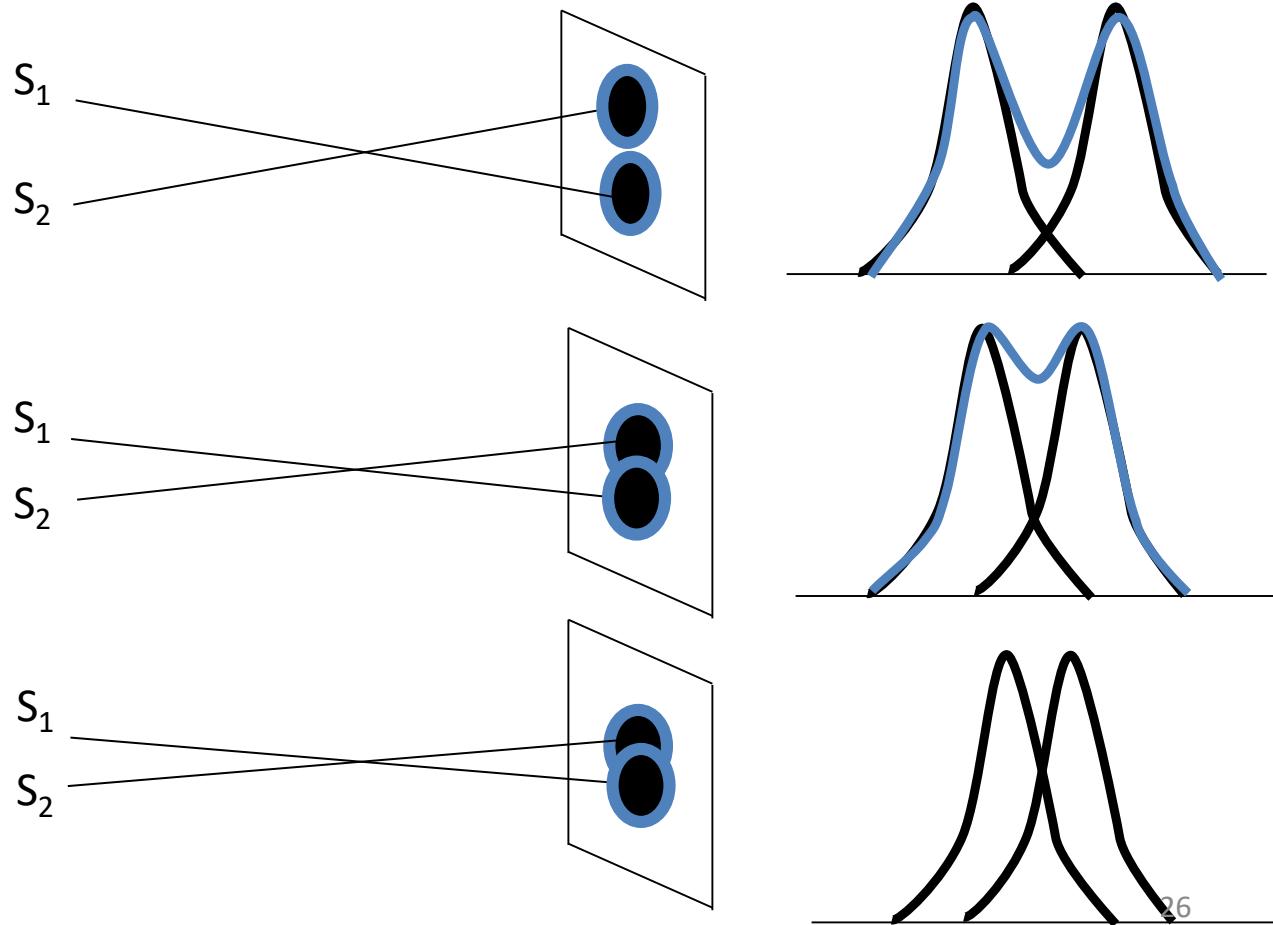
It is an experimental parameter, depends on the lithography system and resist properties.

To increase resolution,
one can:

Increase NA by using large
lens and/or immersion in
a liquid ($n > 1$).

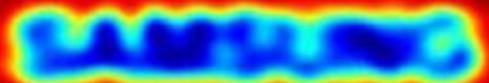
Decrease k_1 factor (many
tricks to do so).

Decrease λ (not easy,
industry still insists on
193nm).



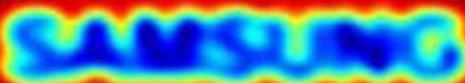
Effect of imaging/printing conditions

KrF, Dry, NA = 0.7, Annular

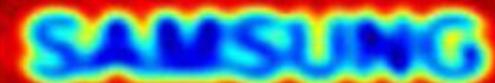


248 nm

KrF, Dry, NA = 0.8, Annular

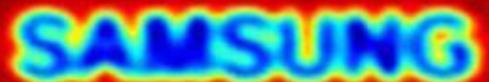


ArF, Dry, NA = 0.68, Annular

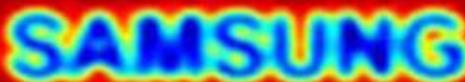


193 nm

ArF, Dry, NA = 0.75, Annular



ArF, Dry, NA = 0.85, Annular



ArF, Dry, NA = 0.93, Annular



F2, Dry, NA = 0.85, Annular



157 nm

ArF, Immersion, NA = 1.33, Annular



EUV, NA = 0.25, Annular

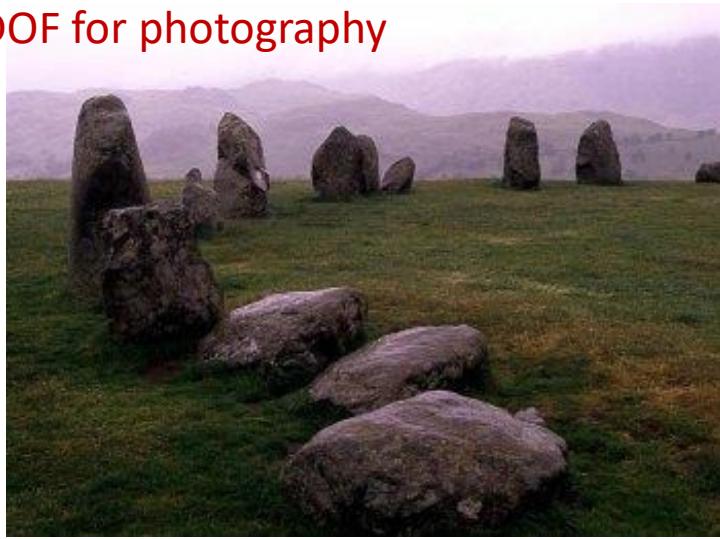


13.5 nm

Annular means an “off-axis illumination” method, which is one trick to decrease k_1 .
EUV: extreme UV, here wavelength 13.5nm. Immersion means exposure in water.

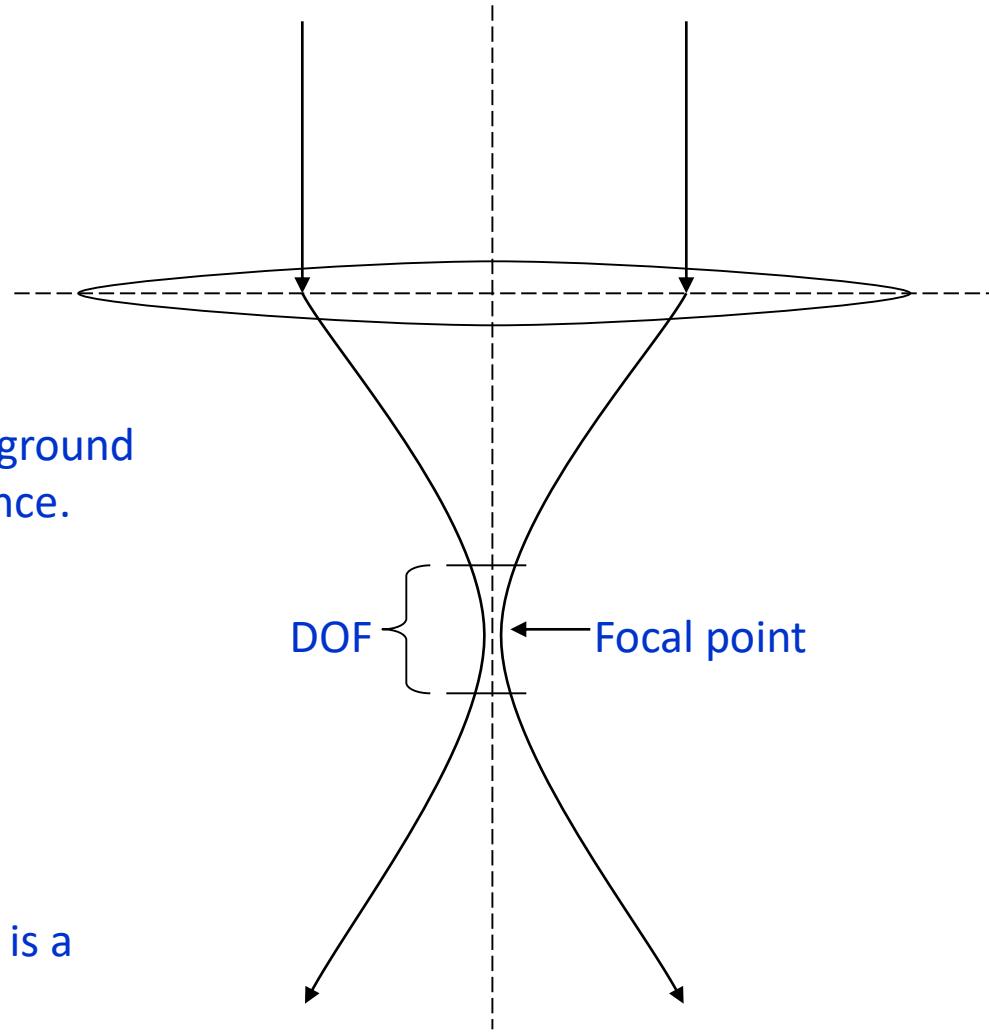
Depth of focus (DOF)

DOF for photography



Large DOF

DOF is the range in which the image is in focus and clearly resolved.



A small aperture was used to ensure the foreground stones were as sharp as the ones in the distance.



Small DOF
(background blurred)

What one need here is a telephoto lens at its widest aperture.

Rayleigh criteria for depth of focus (DOF)

Rayleigh criteria: the length of two optical paths, one on-axis, one from lens edge or limiting aperture, not differ by more than $\lambda/4$.

$$\lambda/4 = \delta - \delta \cos \theta$$

For small θ $\lambda/4 = \delta[1 - (1 - \theta^2/2)] \approx \delta\theta^2/2$

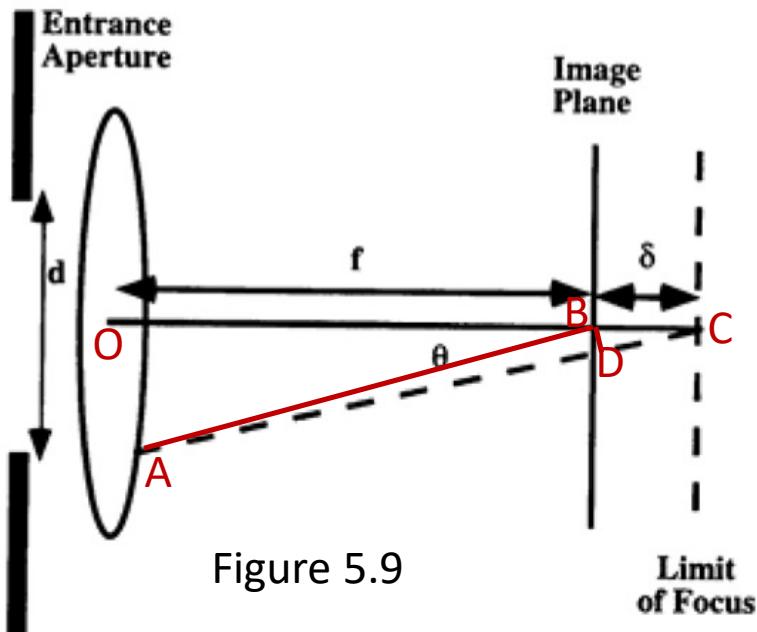


Figure 5.9

On axis, optical path increased by $OC-OB=\delta$.

From edge, increased by $AC-AB=DC=\delta\cos\theta$.

At point B (focal point), two branches have equal path.

$$\theta \approx \sin \theta = \frac{d}{2f} = NA$$

$$DOF = \delta = \pm k_2 \frac{\lambda}{(NA)^2}$$

Again, like the case of resolution, we used k_2 factor as an experimental parameter. It has no well-defined physical meaning.

Modulation transfer function (MTF)

Modulation transfer function is another useful concept.
It is a measure of image contrast on resist.

$$MTF = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

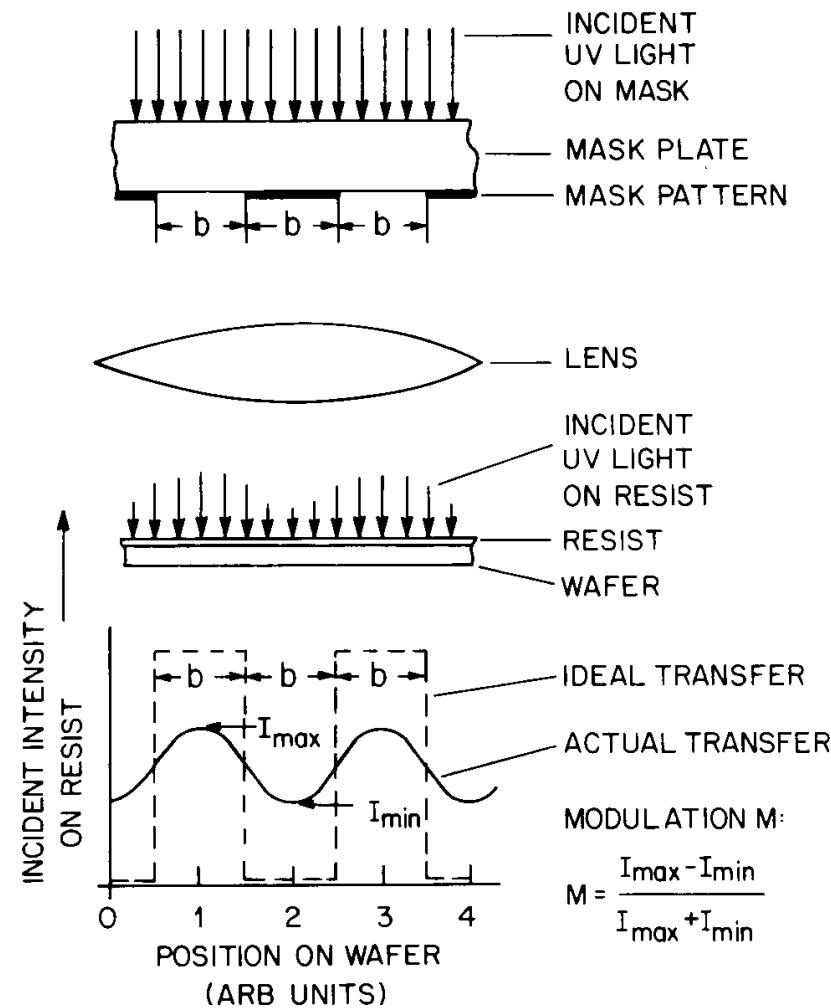
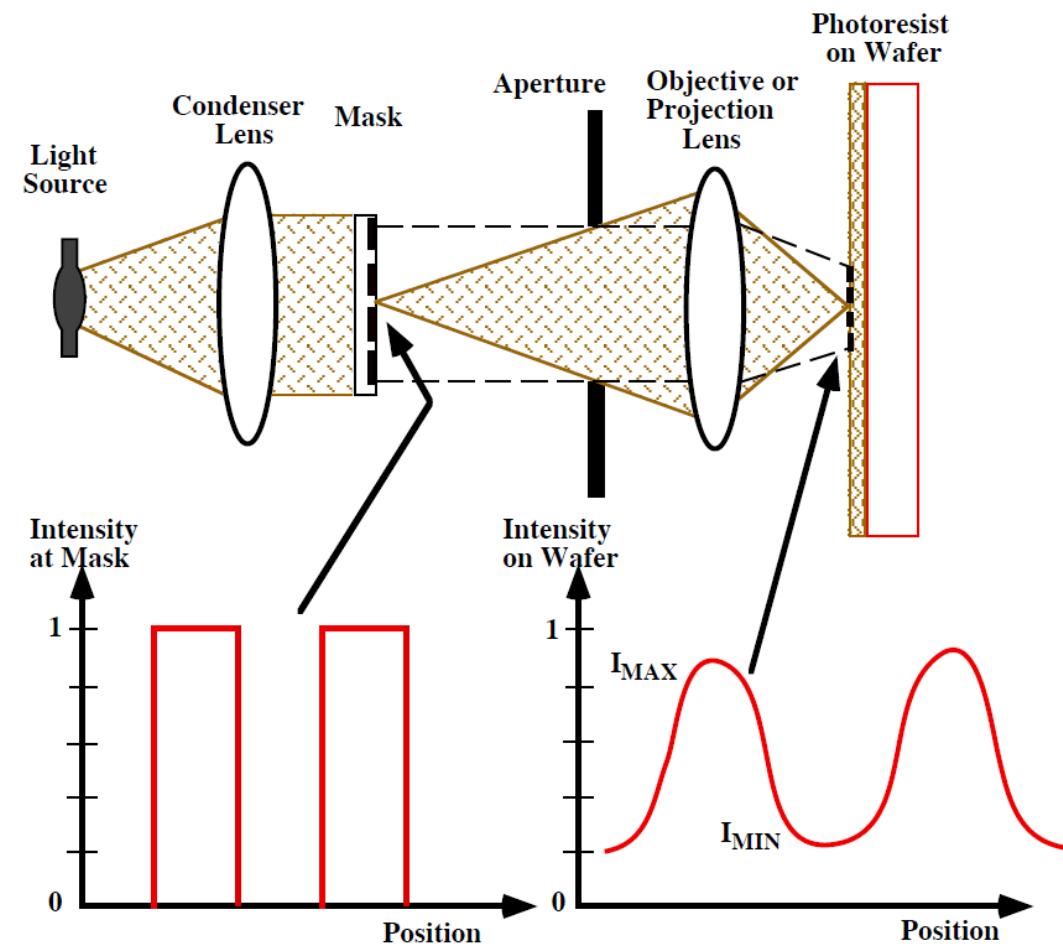
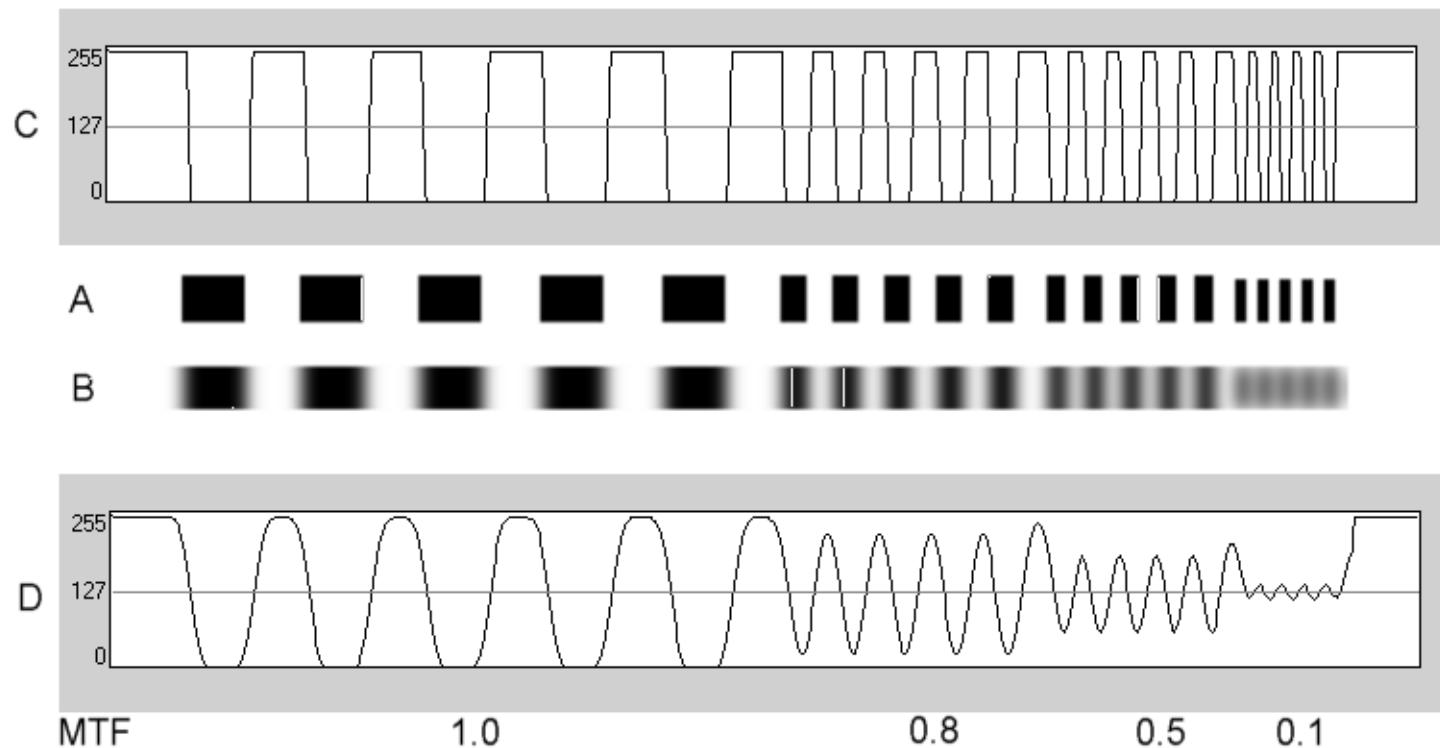


Figure 5.10

Modulation transfer function (MTF) for photography

Photography is similar to projection optical lithography: one forms image of the object (mask) on photo-resist, the other on a negative film (or image sensor for digital camera)



- A. the original test pattern
- B. the image of the test pattern
- C. the line profile of the original test pattern where 255=white and 0=black
- D. the line profile of the image of the test pattern where 255=white and 0= black

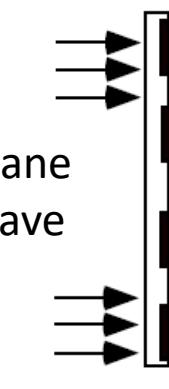
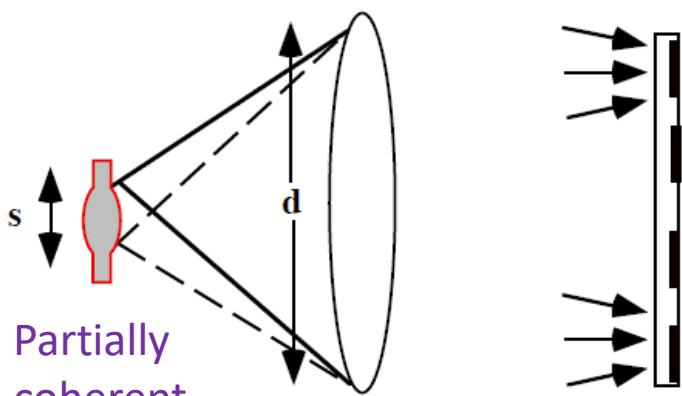
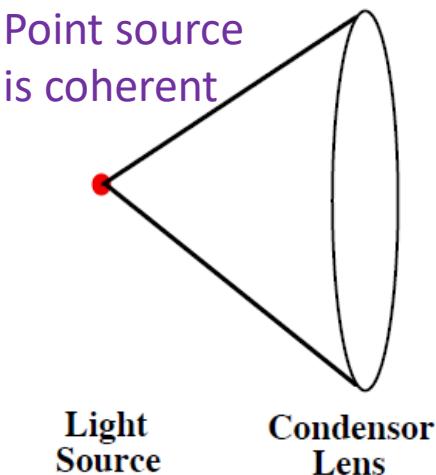
MTF and spatial coherence

Usually MTF > 0.5 is preferred.

It depends on λ , light source size (coherency), and optical system.

It certainly also depends on feature size (or period for a grating pattern).

Spatial coherence of light source



- Coherent light will have a phase to space relationship.
- Incoherent light or light with only partial coherence will have wave-fronts that are only partially correlated.
- Spatial coherence S is an indication of the angular range of light waves incident on mask, or degree to which light from source are in phase.
- Small S is not always good (see next slide).

$$S = \frac{\text{source diameter}}{\text{aperture diameter}} = \frac{s}{d}$$

Figure 5.12

MTF and spatial coherence

MTF vs. diffraction grating period on mask.

W = line width = space width of the grating.

Grating
photomask

X-axis of the plot: spatial frequency

$v=1/(2W)$, normalized to Rayleigh criterion

cutoff frequency $v_0=1/R=NA/(0.61\lambda)$.

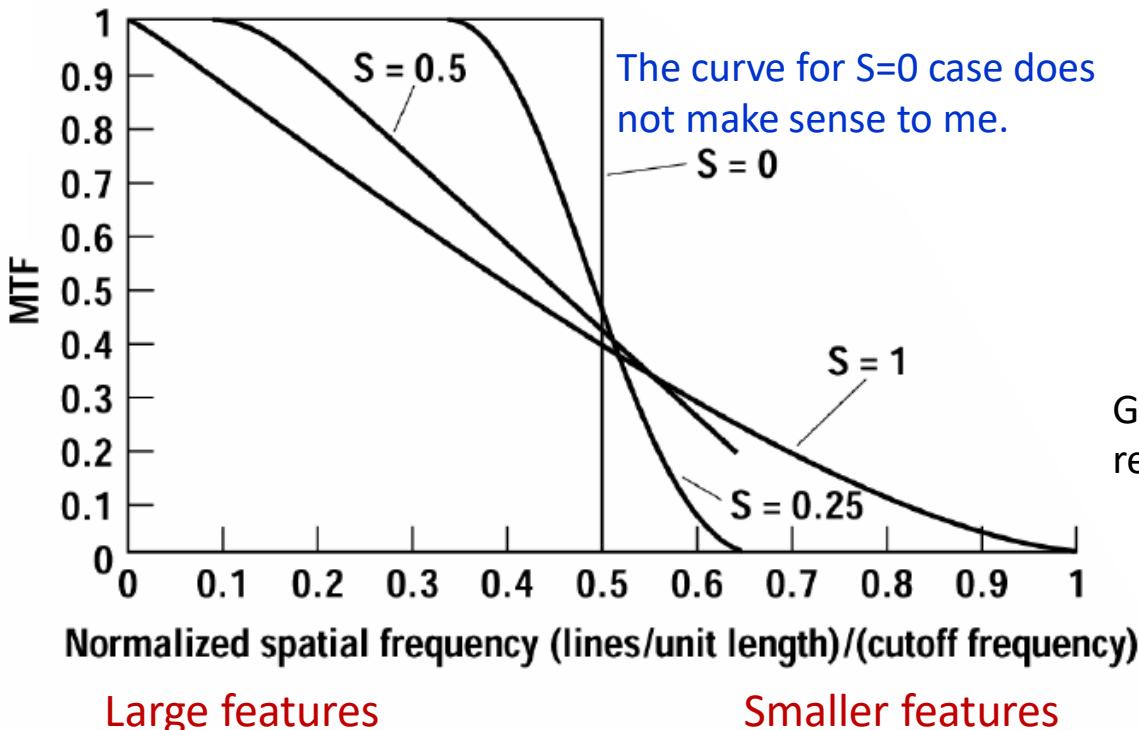
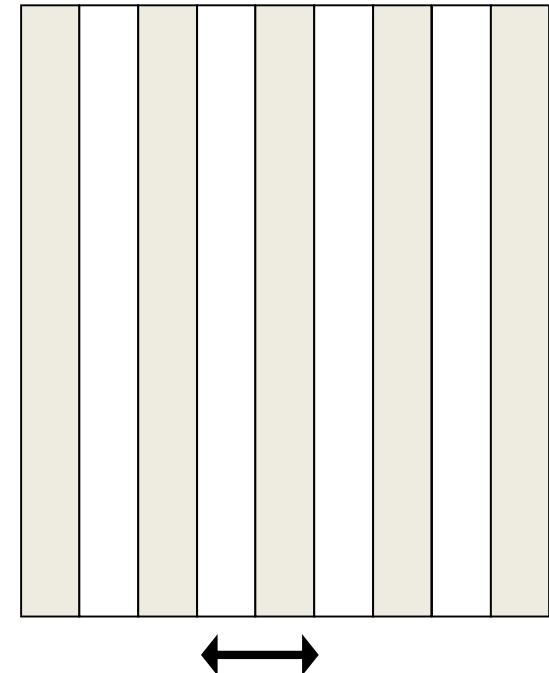


Figure 7.18 Modulation transfer function as a function of the normalized spatial frequency for a projection lithography system with spatial coherence as a parameter.

(similar to Figure 5.13)



Grating period on mask is actually $8W$ (4 \times reduction by lens to make $2W$ on resist)

Large S is good for smaller features, but bad for larger ones.

Trade-off is made, and industry chooses $S=0.5-0.7$ as optimal.

Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Surface reflection and standing wave

- Resist is partially reflective, so some light reaches resist bottom and is reflected.
- Constructive and destructive interference between incident and reflected light results in a periodic intensity distribution across the resist thickness.
- With change in exposure (light intensity) comes change in resist dissolution rate, leading to zigzag resist profile after development.
- Use of anti-reflecting coating (ARC) eliminates such standing wave patterns.
- Post exposure bake also helps by smoothing out the zigzag due to resist thermal reflow.
- (Also due to reflection, a metal layer on the surface will require a shorter exposure than exposure over less reflective film.)

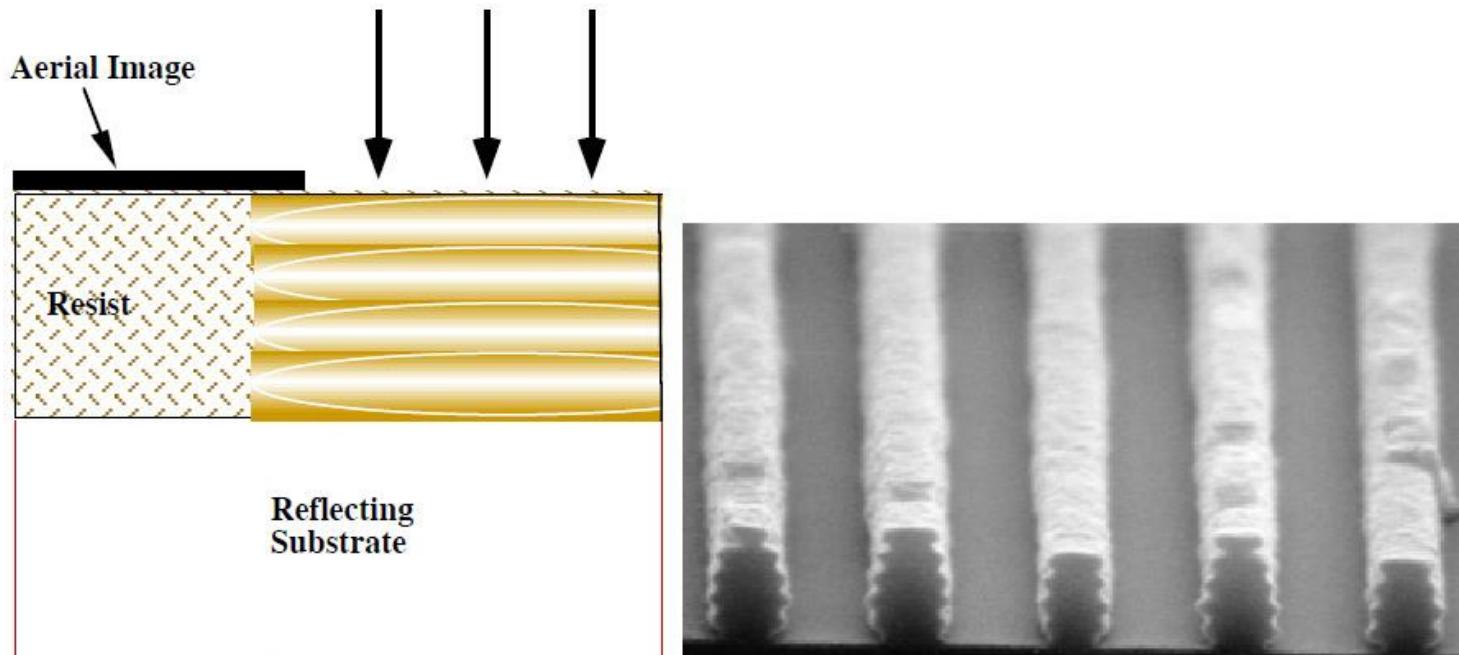
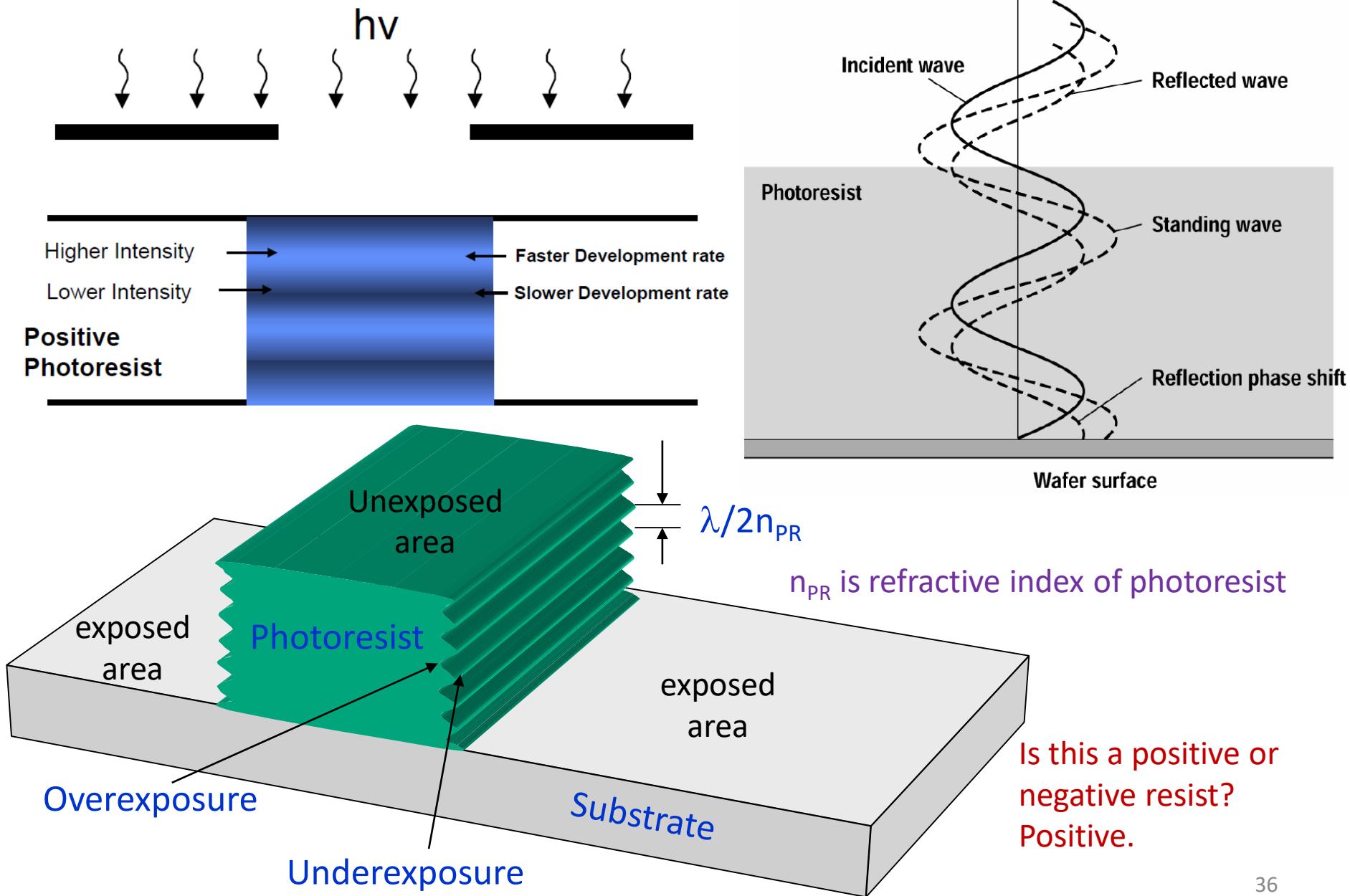


Figure 5.24

(Photo courtesy of A. Vladar and P. Rissman, Hewlett Packard.)

Standing wave effect on photoresist



Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Photoresist overview

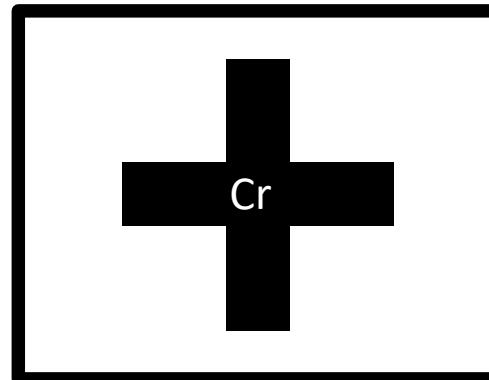
Photoresist is a liquid mixture that can be spun onto a substrate, exposed and developed into a pattern for subsequent processing.

Typically consists of 3 components:

- Resin - a binder that provides mechanical properties (adhesion, chemical resistance...).
- Sensitizer - photoactive compound.
- Solvent – e.g. n-butyl acetate, xylene, keep the resist in a liquid form for spin coating. Its content determines viscosity and hence resist thickness.

There are two types of photoresist:

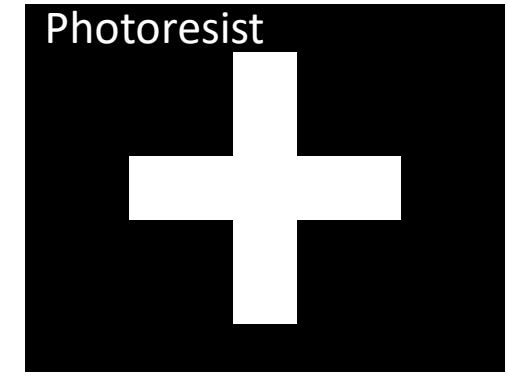
- Positive: exposed area removed by developer.
- Negative: unexposed area removed by developer.



Mask



Positive Resist



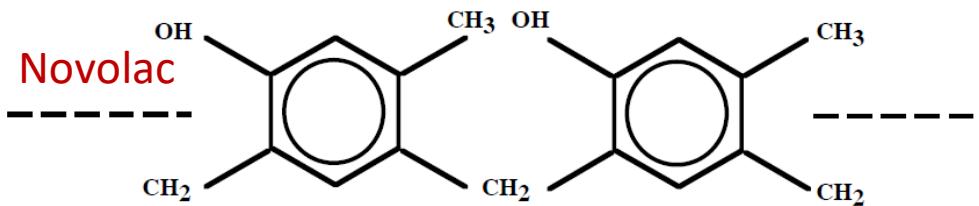
Negative Resist

Positive resist: DNQ

- It is the most popular positive resists for i-line (365nm) and g-line (436nm) exposure, but cannot be used for very short λ .
- It consists of diazonaphthoquinone (DNQ), which is the photoactive compound (PAC); and novolac, a matrix material called resin.
- After spinning and baking, resists contains roughly 1:1 PAC and resin.

Novolac

A polymer whose monomer is an aromatic ring with 2 methyl groups and an OH group. It dissolves easily in a base developer solution. Solvents are added to adjust the viscosity.



DNQ (diazo-naphto-quinone)

It is the PACs in these resists, and it acts as an inhibitor, reducing the dissolution rate of the resist in the developer.

This occurs by a chemical bonding of the PAC and the novolac at the surface of the resist where it is exposed to the developer.

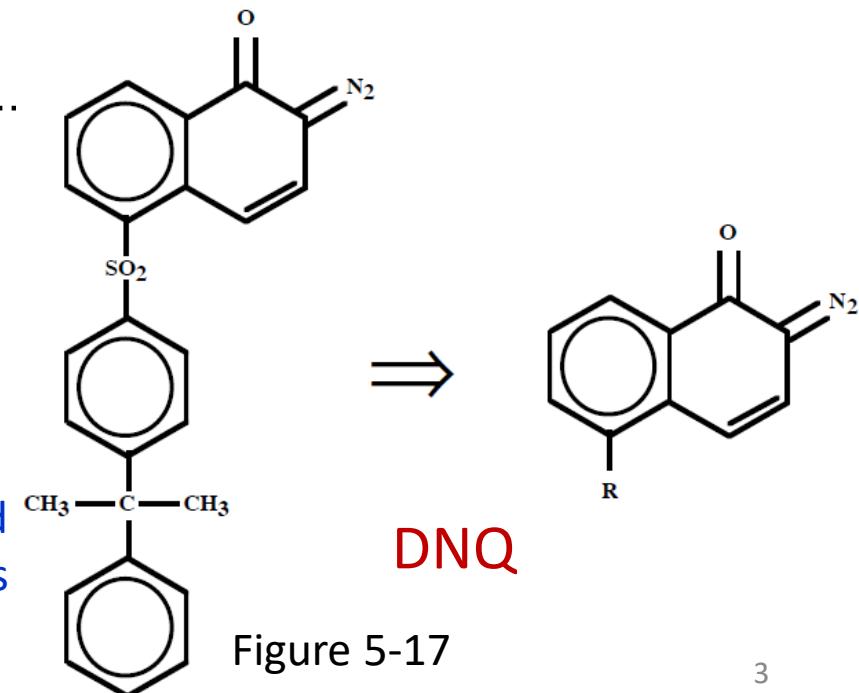


Figure 5-17

DNQ upon UV exposure

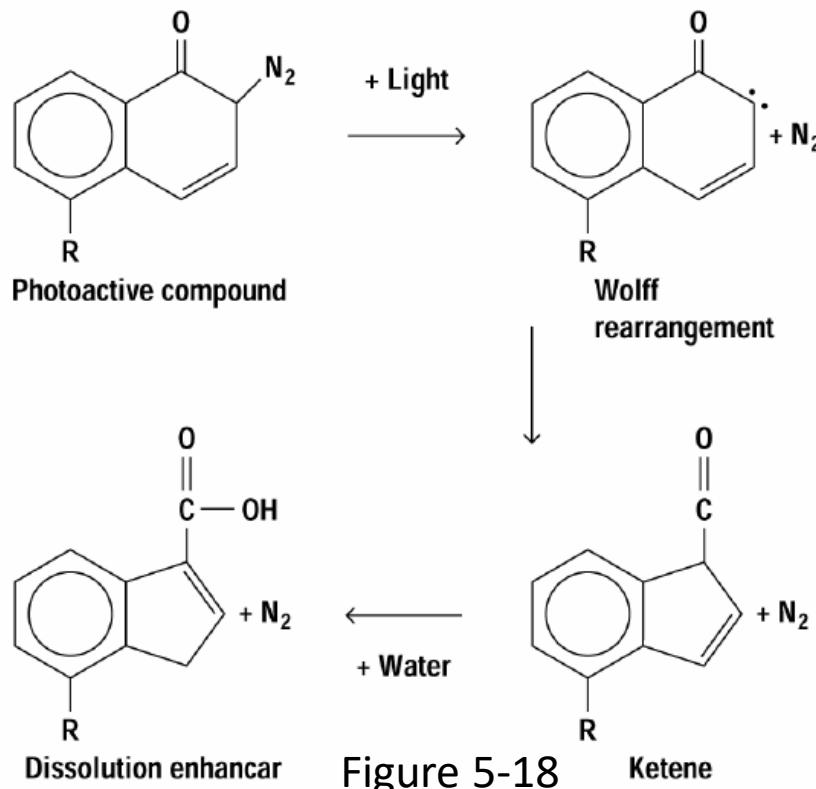


Figure 5-18

Sensitizers: (here it is DNQ)

- It is also called photoactive compounds (PAC).
- It absorb radiation and undergo chemical reactions to change their chemical dissolution properties in developer.
- The net result is Differential Dissolution rate (100:1) between areas that absorbed radiation and areas that did not absorb radiation.
- Sensitizers are developer resistant before they absorb radiation.

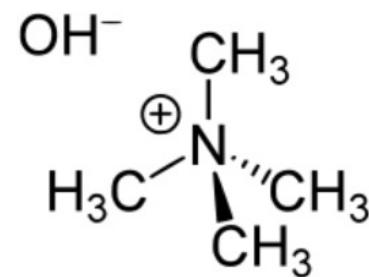
- Addition of UV light will free nitrogen molecule from the carbon ring leaving behind a highly reactive carbon site.
- One way to stabilize the structure is to move one of the carbons outside the ring, and the oxygen atom is covalently bonded to this external carbon atom.
- This process known as Wolff rearrangement.
- In presence of water, the resulting ketene molecule finally transforms into carboxylic acid, which is readily soluble in basic developer (KOH, NaOH, TMAH etc).

Developing of DNQ resist and its advantage

- Novolac resin is water soluble.
- But due to the addition of the DNQ PAC (to the novolac matrix in) at about a 1:1 ratio, the resist is almost insoluble in a base solution ($\text{pH} > 7$).
- Whereas after exposure, the generated carboxylic acid readily dissolve in base solutions.
- The chemical reaction that occurs during this dissolution is the breakdown of the carboxylic acid into water-soluble amines such as aniline (phenylamine, one H in NH_3 replaced by a benzene ring, $\text{C}_6\text{H}_7\text{N}$).
- This process continues until all of the exposed resist is removed.
- Typical developer solutions are KOH or NaOH diluted with water, yet in recent years the so-called MIF (metal ion free) developer based on TMAH dominates, because K^+ and Na^+ ions are very bad for device.

Dissolution rate in developer

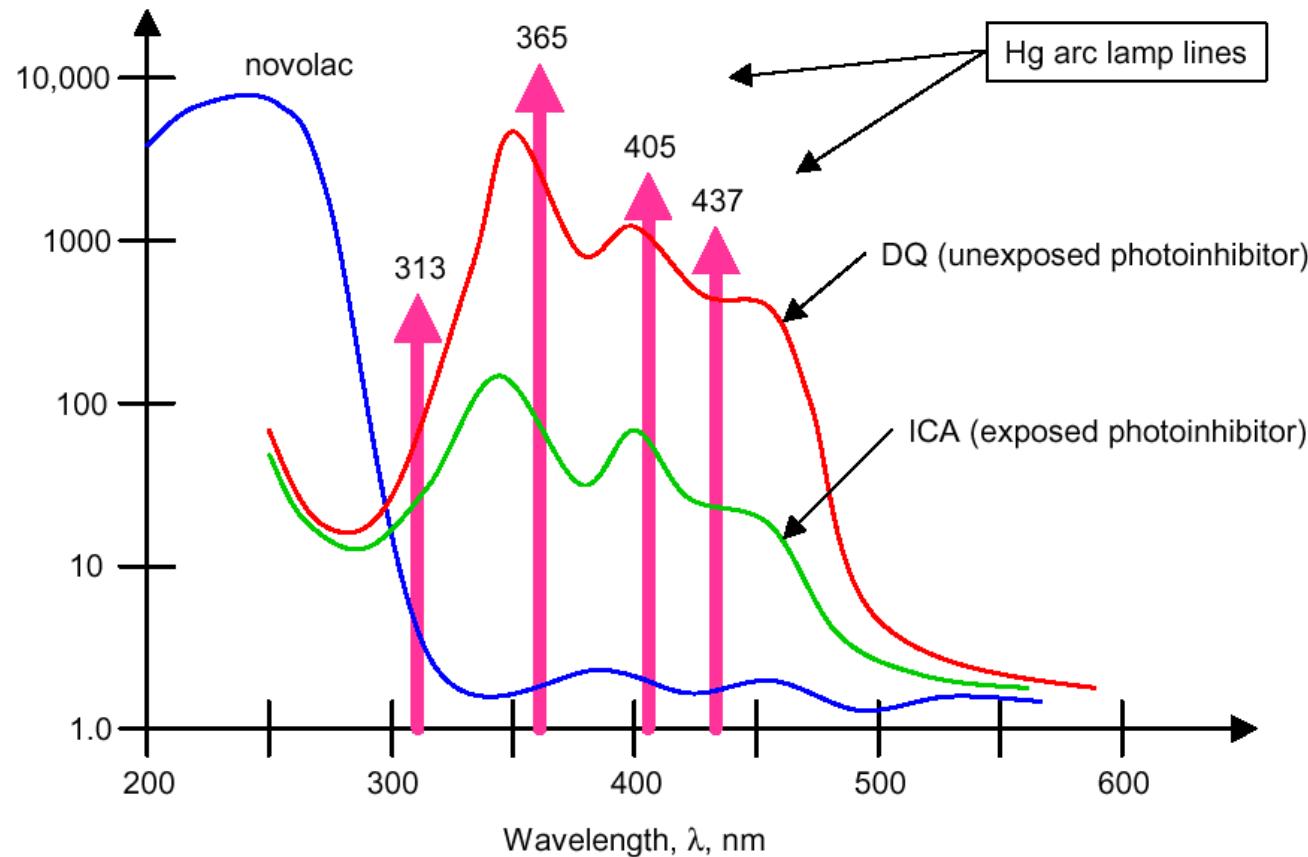
Without sensitizer	150 Å/s
With sensitizer	10-20 Å/sec
After exposure	1000-2000 Å/s



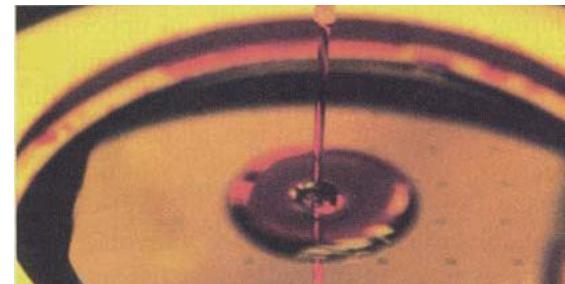
TMAH: tetra-methyl-ammonium hydroxide
H in NH_4OH replaced by CH_3 group.

Color of DNQ photoresist

Spectral Absorption of Novolac, DQ, and ICA



What should be the color of photoresist?



Photoresist properties

- **Resolution:** how fine a line the resist can reproduce from an areal image. It is determined by contrast, thickness, proximity effects, swelling and contraction after development.
- **Contrast:** ability of resist to distinguish between light and dark regions, measured by exposing resist of given thickness to varying radiation dose and measuring dissolution rate.
- **Sensitivity:** incident energy necessary to produce the photochemical reactions required for defining patterns. It is related to quantum yield ($=\# \text{ of photon-induced events}/\# \text{ of photons absorbed}$).
- **Etch resistance:** Novolac is a long-chain aromatic ring polymer that is fairly resistant to chemical attack. Therefore, the resist is a good mask for wet or dry plasma etching.
- **Spectral response curve:** should match the exposure light source.

Example: resist sensitivity

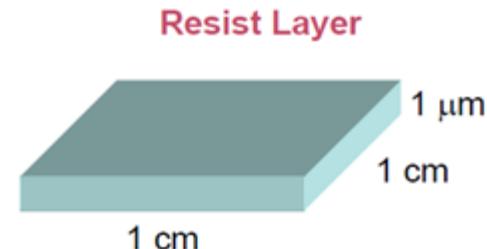
Photon energy $E=hf=hc/\lambda=4.54\times10^{-19}\text{J}$.

Number of photons: $(150\text{mJ/cm}^2)/4.54\times10^{-19}\text{J}=3.3\times10^{17}/\text{cm}^2$.

Volume/photon= $3\times10^{-22}/\text{cm}^3$. (volume= $1\mu\text{m}\times1\text{cm}\times1\text{cm}=10^{-4}\text{cm}^3$)

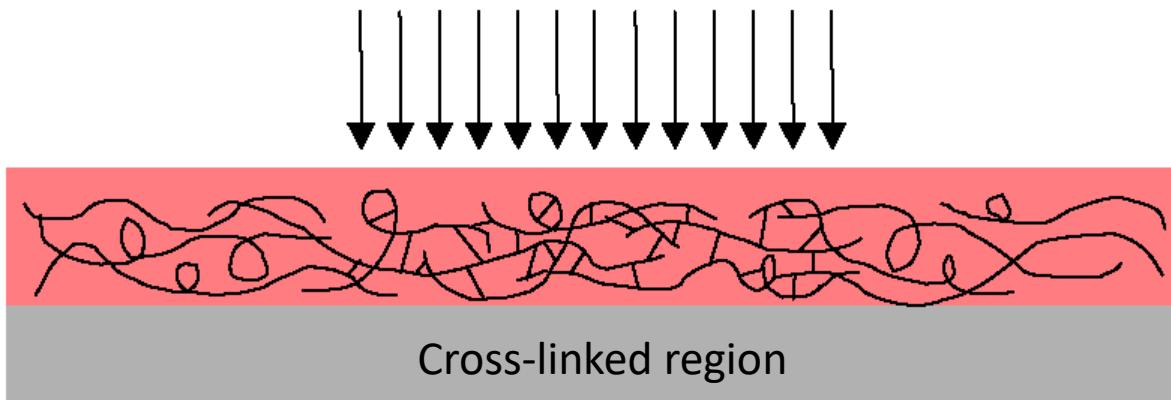
Mean photon separation: $(3\times10^{-22}/\text{cm}^3)^{1/3}=0.67\text{nm}$.

Resist	Kodak 809 UV Positive Resist
Sensitivity	$S = 150 \text{ mJ/cm}^2$
Exposure	G-Line (436 nm)
Thickness	$t = 1 \mu\text{m}$



Negative photoresist

- Negative photoresist becomes insoluble in regions exposed to light.
- It is a polymer with long chains. Molecular weight $10^4\text{-}10^6\text{ g/mol}$, about one order higher than that of DNQ positive resist.
- Irradiation results in bonding or cross-linking (form 3D molecular network) of adjacent polymer chains and increases of molecular weight.
- Unexposed resists dissolve in aromatic solvents such as benzene, toluene and xylene.



Component of *one* negative photoresist

- Resin: cyclized synthetic rubber resin, not sensitive to exposure, fast dissolution in organic solvent such as toluene and xylene.
- Sensitizer PAC: bis-arylzide
- Solvent: aromatic solvent
- Developer: organic solvents (we know for positive resist, it is inorganic)

Deep UV (DUV) resists

- Traditional g-line and i-line resists have maximum quantum efficiencies $\approx 30\%$ that limits its sensitivity. In addition, it absorbs too strongly for $\lambda < 1$ line ($=365\text{nm}$).
- Chemical amplification can improve sensitivity significantly, with *effective* quantum efficiency $>> 100\%$.
- DUV resists are all chemically amplified resist (again, this is true only for industrial application; for R&D, even a simple polymer like PMMA can be used as DUV resist).
- Photo-acid generator (PAG) is converted to an acid by photon exposure. Later, in a post exposure bake (PEB), the acid molecule reacts with a “blocking” molecule on a polymer chain, making it soluble in developer and regenerating the acid molecule.
- It is basically a **catalytic chain action**. In principle, only one photon is needed to generate one “seed” (acid catalyst), and all the rest reaction takes place during PEB. (this also means that PEB temperature needs to be tightly controlled for reproducible result)

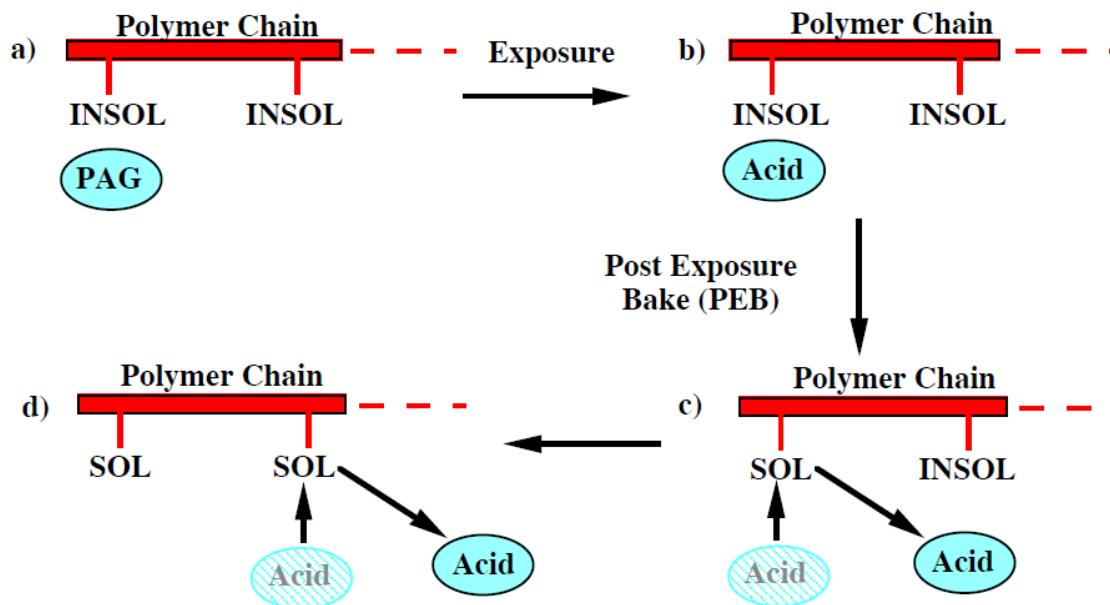


Figure 5-19 Basic operation of a chemically amplified resist. PAG is photo-acid generator, INSOL and SOL are the insoluble and soluble portions of the polymer base.

Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

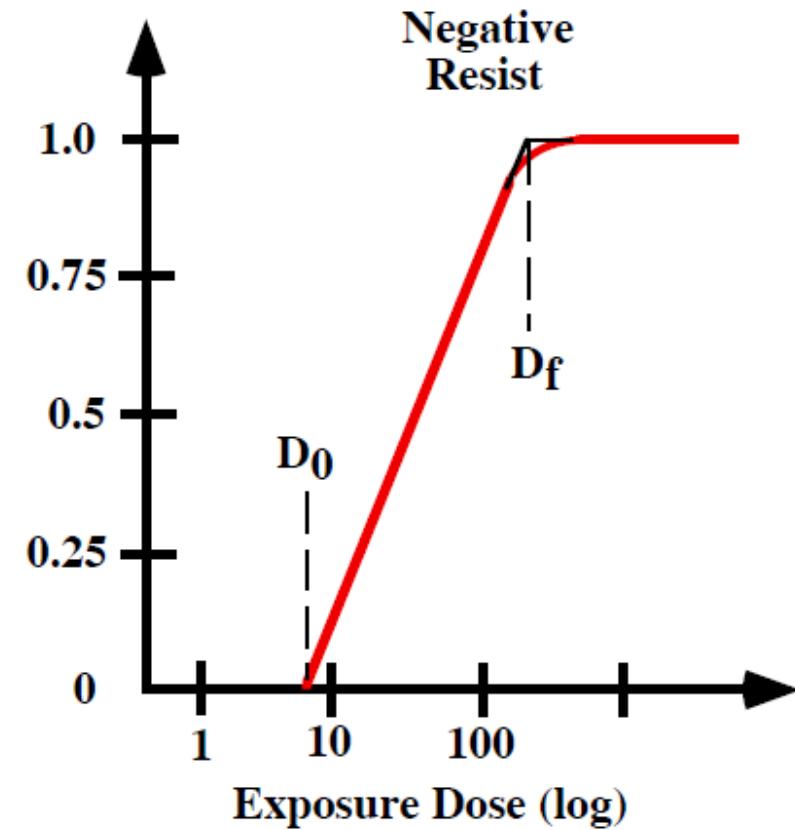
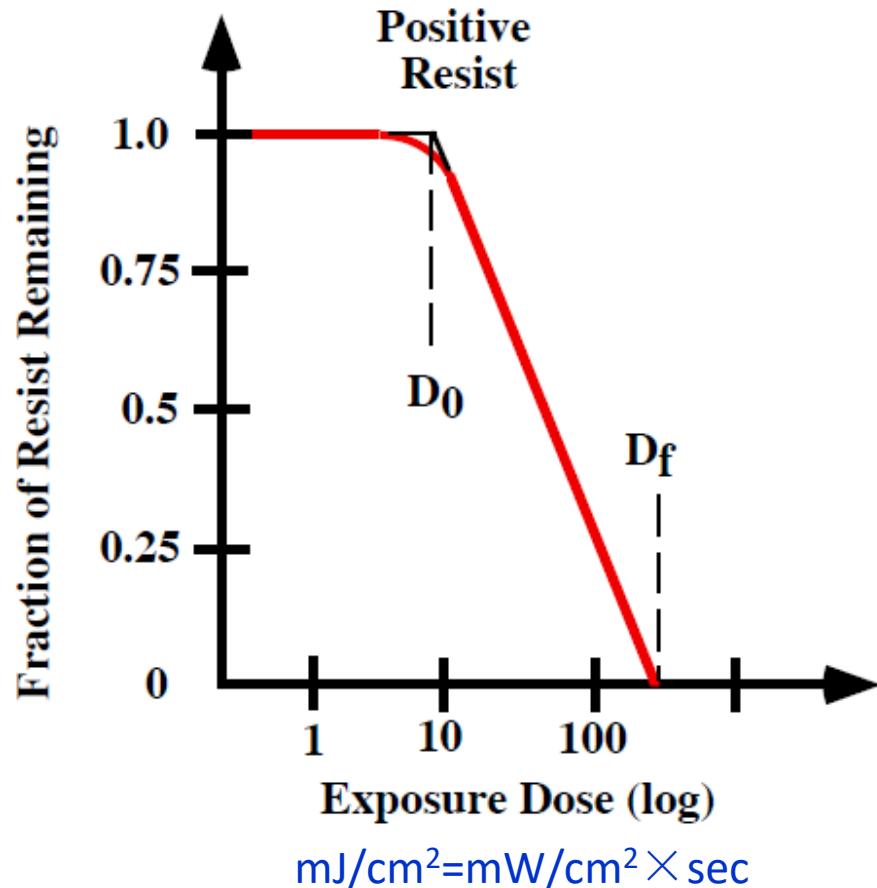
Contrast and sensitivity

Contrast γ is defined as:

$$\gamma = \frac{1}{\log_{10} \frac{D_f}{D_0}}$$

D_f is Sensitivity.

- Typical g-line and i-line resists: $\gamma=2-3$, $D_f \sim 100 \text{ mJ/cm}^2$.
- DUV resists: $\gamma=5-10$, $D_f \sim 20 - 40 \text{ mJ/cm}^2$. (chemically amplified)
- γ and D_f are not intrinsic properties of the resist - they depend on process conditions (developer, development time, baking time, λ , substrate...).

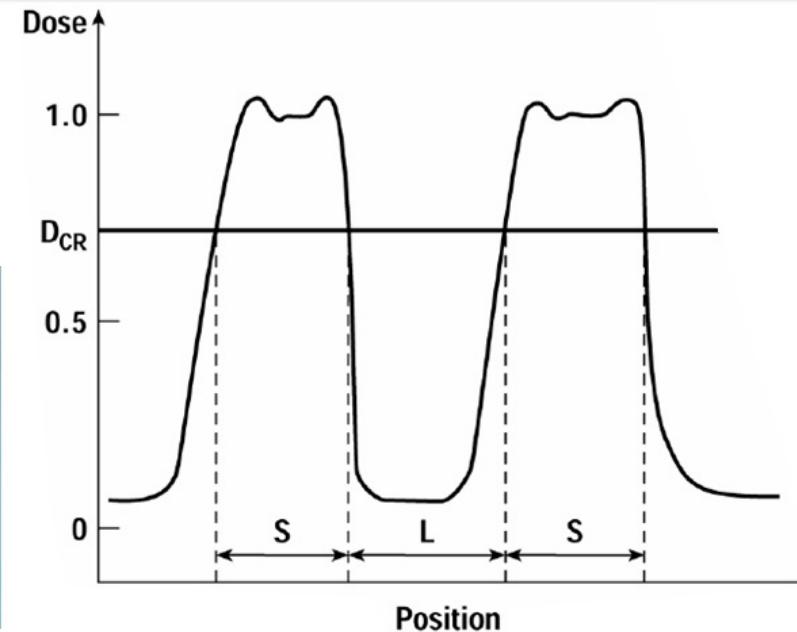
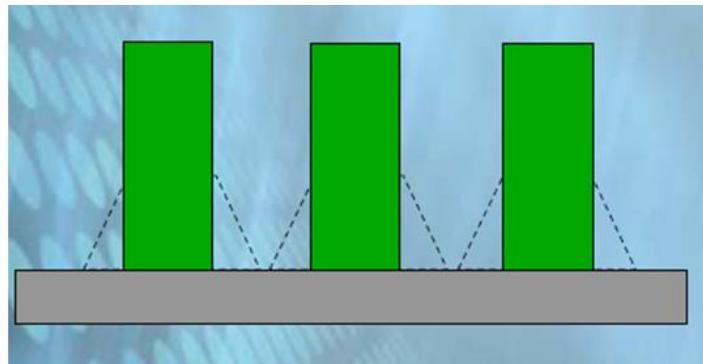


Ideal resist response: $\gamma \rightarrow \infty$, $D_0 = D_f = D_{cr}$

D_{cr} : critical exposure dose.

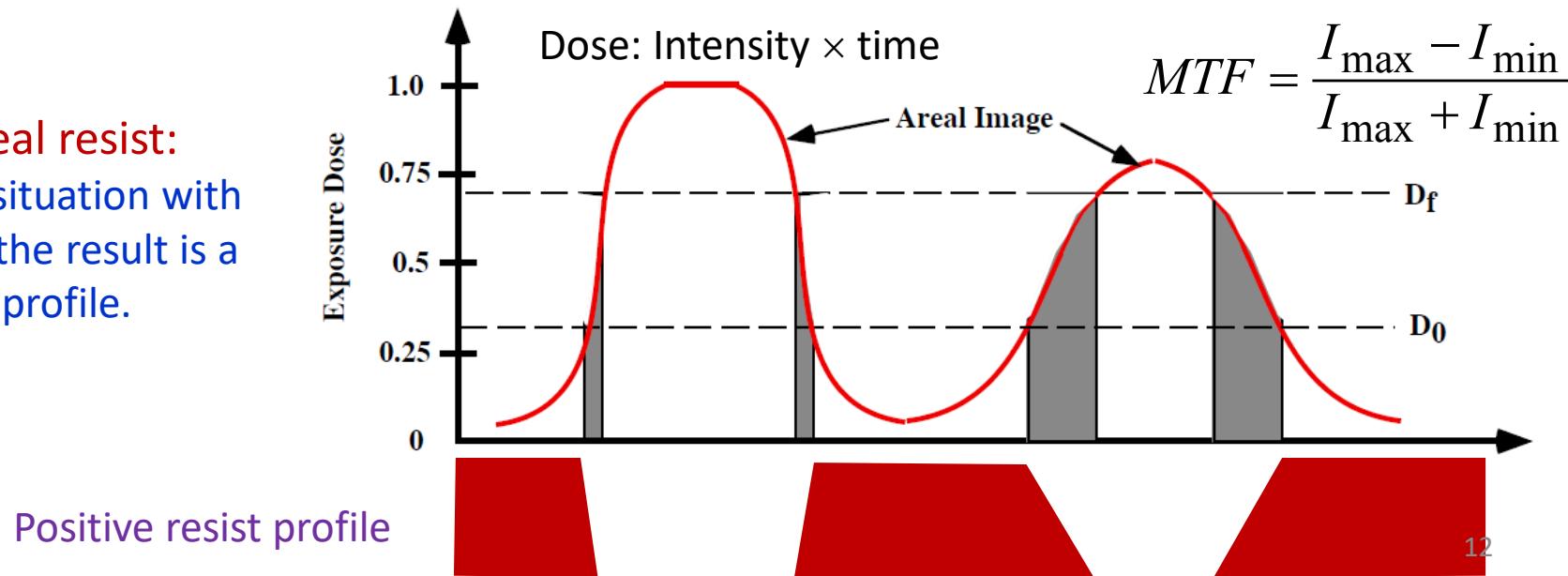
Resist receives exposure dose $> D_{cr}$ will completely dissolve during developing.

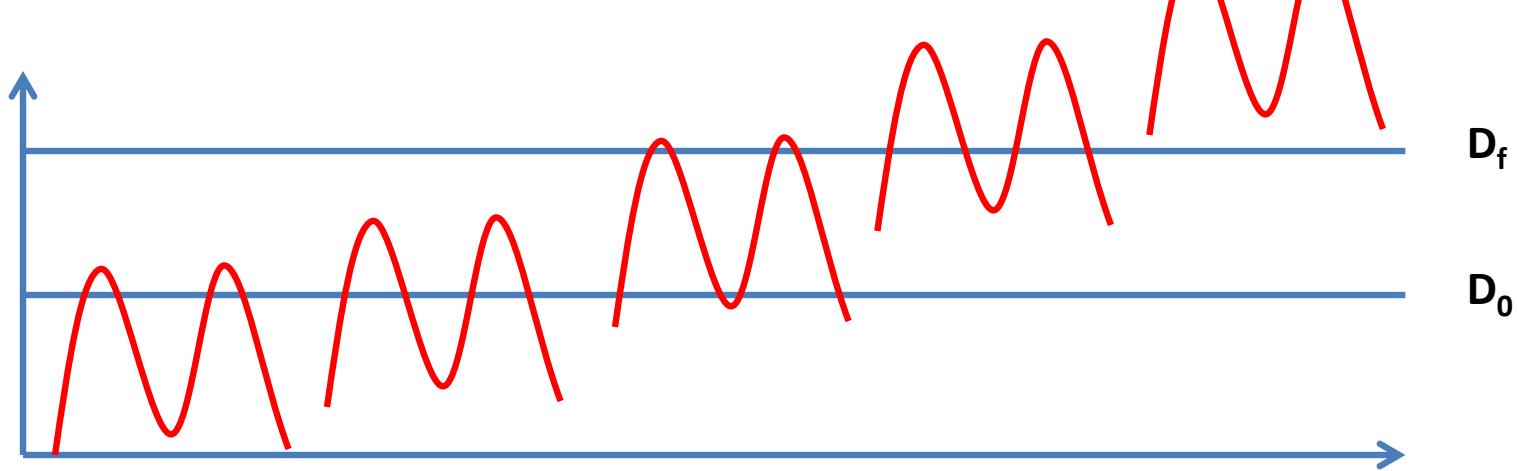
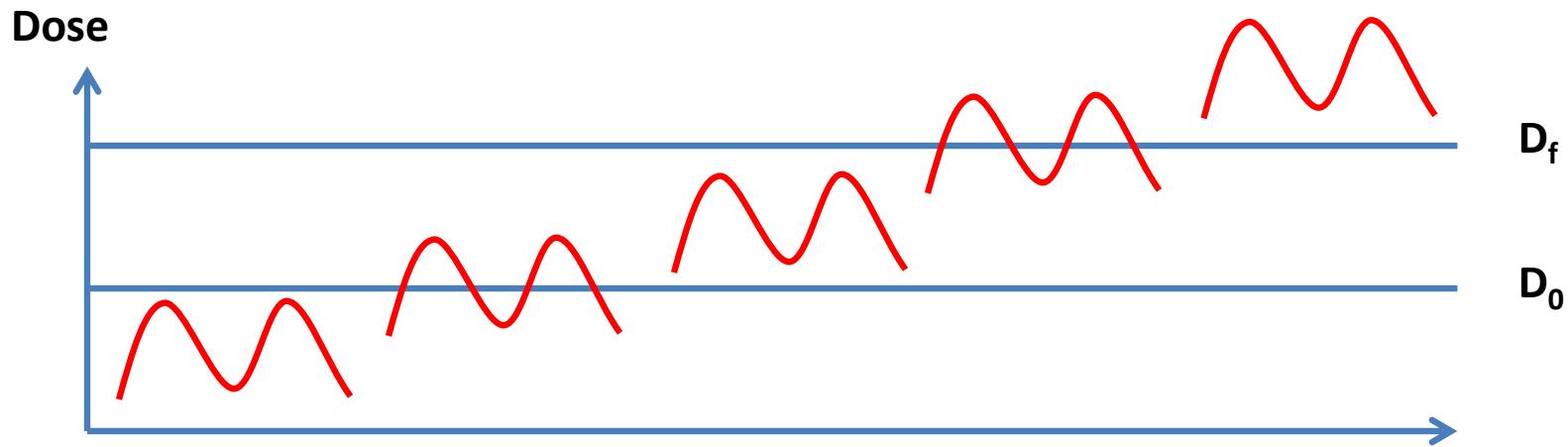
Dose $< D_{cr}$ will not be attacked during developing.



Ideal resist:
vertical resist
profile.

Non-ideal resist:
for real situation with
finite γ , the result is a
tapered profile.





Resist critical modulation transfer function (CMTF)

By analogy to the MTF for optical systems, the CMTF for resists is defined as:

$$CMTF = \frac{D_f - D_0}{D_f + D_0} = \frac{10^{1/\gamma} - 1}{10^{1/\gamma} + 1}$$

For MTF, *higher* is better

For CMTF, *lower* is better

Typical CMTF values for g and i-line resists are about 0.4. Chemically amplified DUV resists achieve CMTF values of 0.1 - 0.2.

In general $CMTF < MTF$ is required for the resist to resolve the aerial image.

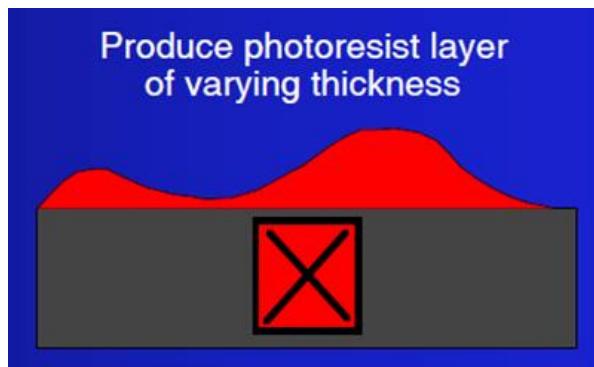
(e.g. if $MTF=1$, then any $D_f - D_0$ is OK.)

(e.g. if $CMTF=0$, then any MTF is OK)

Gray-scale photolithography

Conventional photolithography cannot produce arbitrary 3D structures with varying thicknesses.

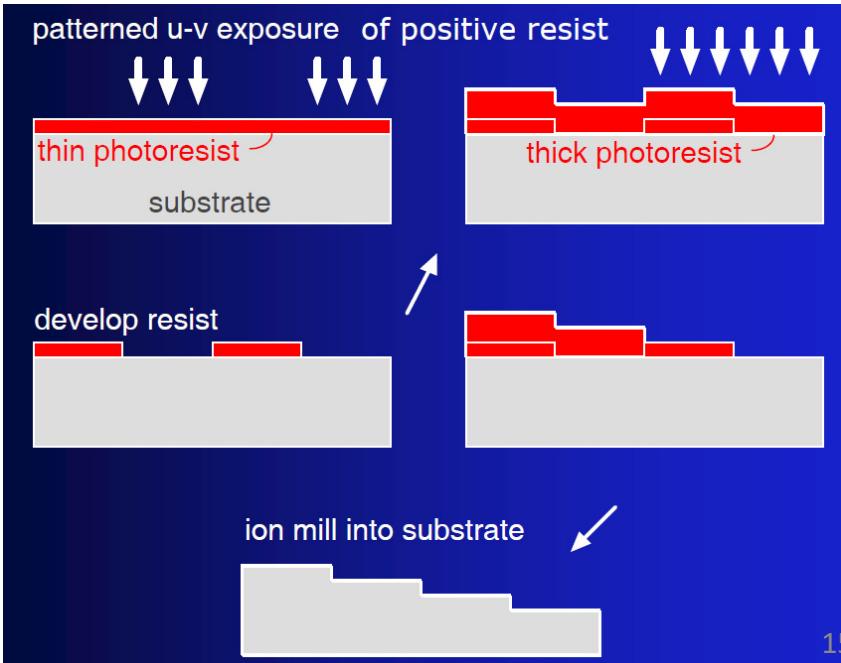
3D structures are useful for micro-optics: lens arrays, integrated optics, micro-opto-electromechanical systems on a chip (MOEMS), grayscale diffractive elements, beam shaping, and wave-front analysis.



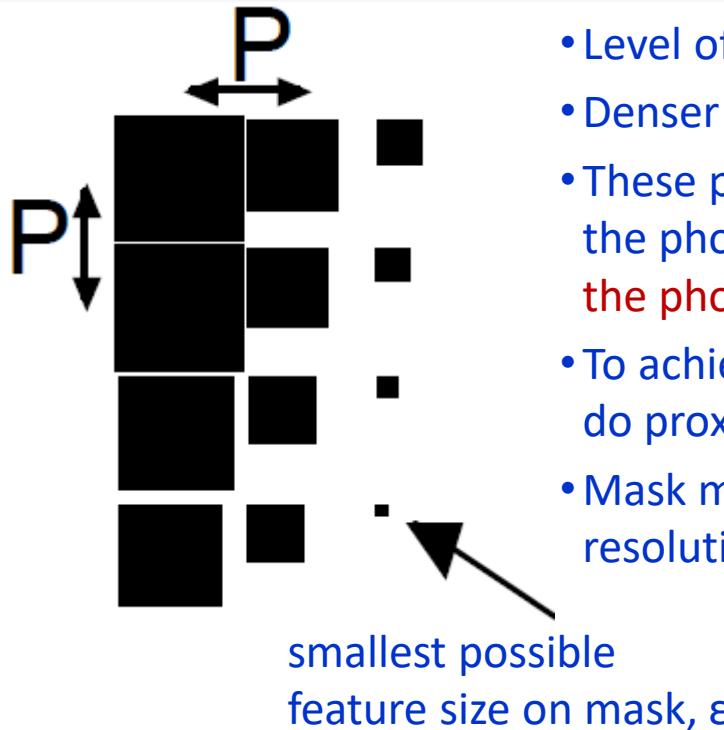
3D lithography needs 3D photomask (i.e. Cr thickness varies across the mask).

This process will produce a 3D (though not real 3D) mask if the substrate in the graph is replaced by Cr deposited on quartz.

Micro-lens array



Typical pseudo grayscale mask (actually a binary black and white mask)



- Level of “gray” is decided by pattern density.
- Denser pattern looks darker.
- These patterns have feature size < resolvable feature size of the photolithography (i.e. $P <$ resolution R), so **what is seen by the photolithography is the *averaged pattern*.**
- To achieve low resolution such that $R > P$, one can, for example, do proximity lithography with a large gap.
- Mask must be produced by e-beam exposure or high resolution laser beam writing.

P must be below resolution limit of mask aligner.

Range of gray tones:

$0 \rightarrow 1$, in steps of $(\epsilon/P)^2$

Example: $P = 1 \mu\text{m}$, $\epsilon = 0.1\mu\text{m}$

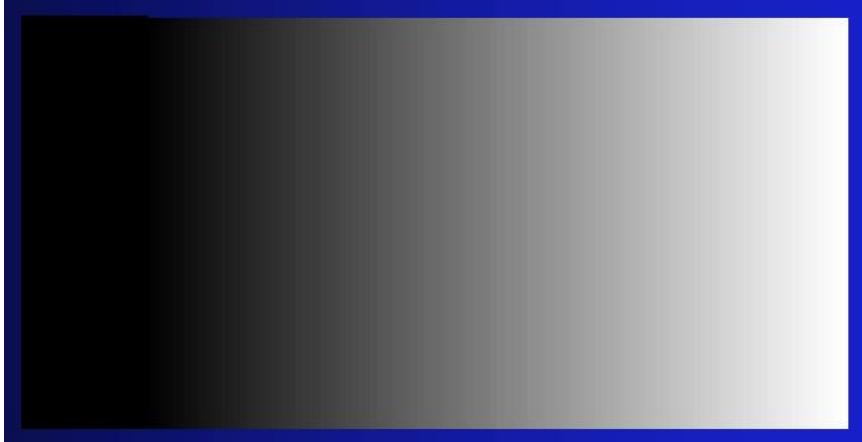
$\rightarrow 100$ grayscale steps

Design considerations:

- Mask design depends critically on a stable, consistent, reproducible lithographic process!
- Creation of a reliable depth vs gray-level calibration relationship is essential.
- Overall, the process is very tricky, need lots of trials.

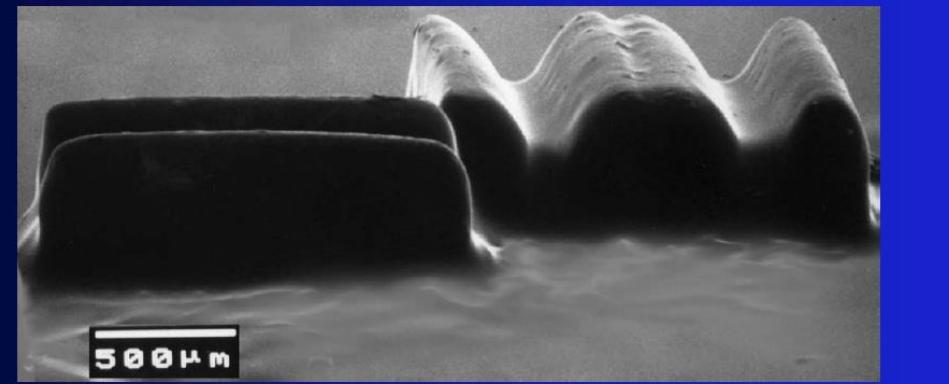
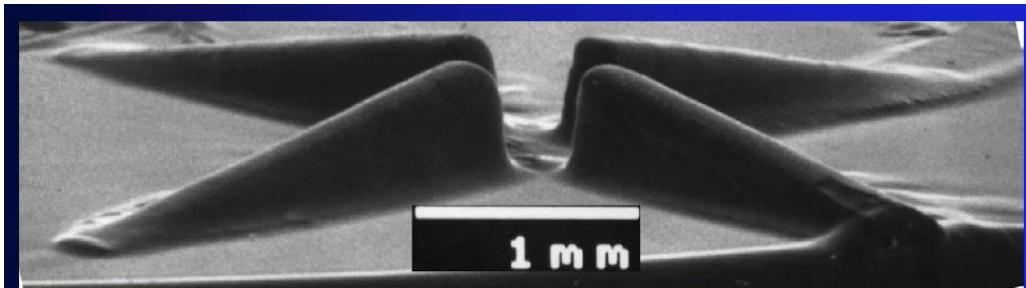
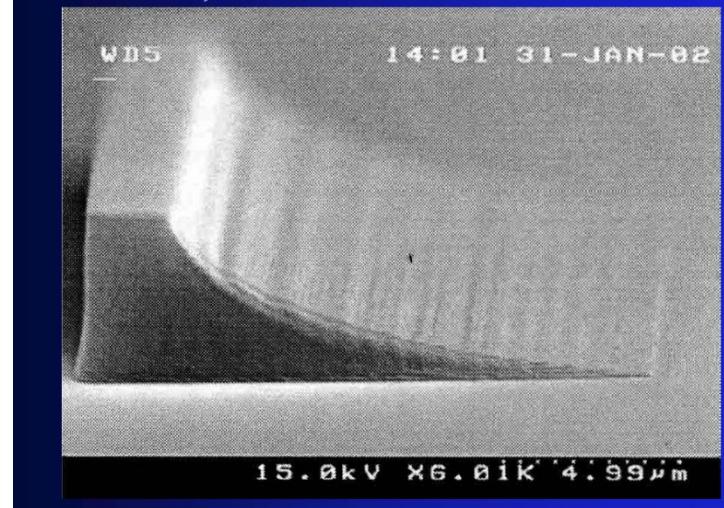
Examples

Linearly stepped grayscale mask, 64 levels



Exposure using linearly stepped grayscale mask

Positive resist, AZ6420

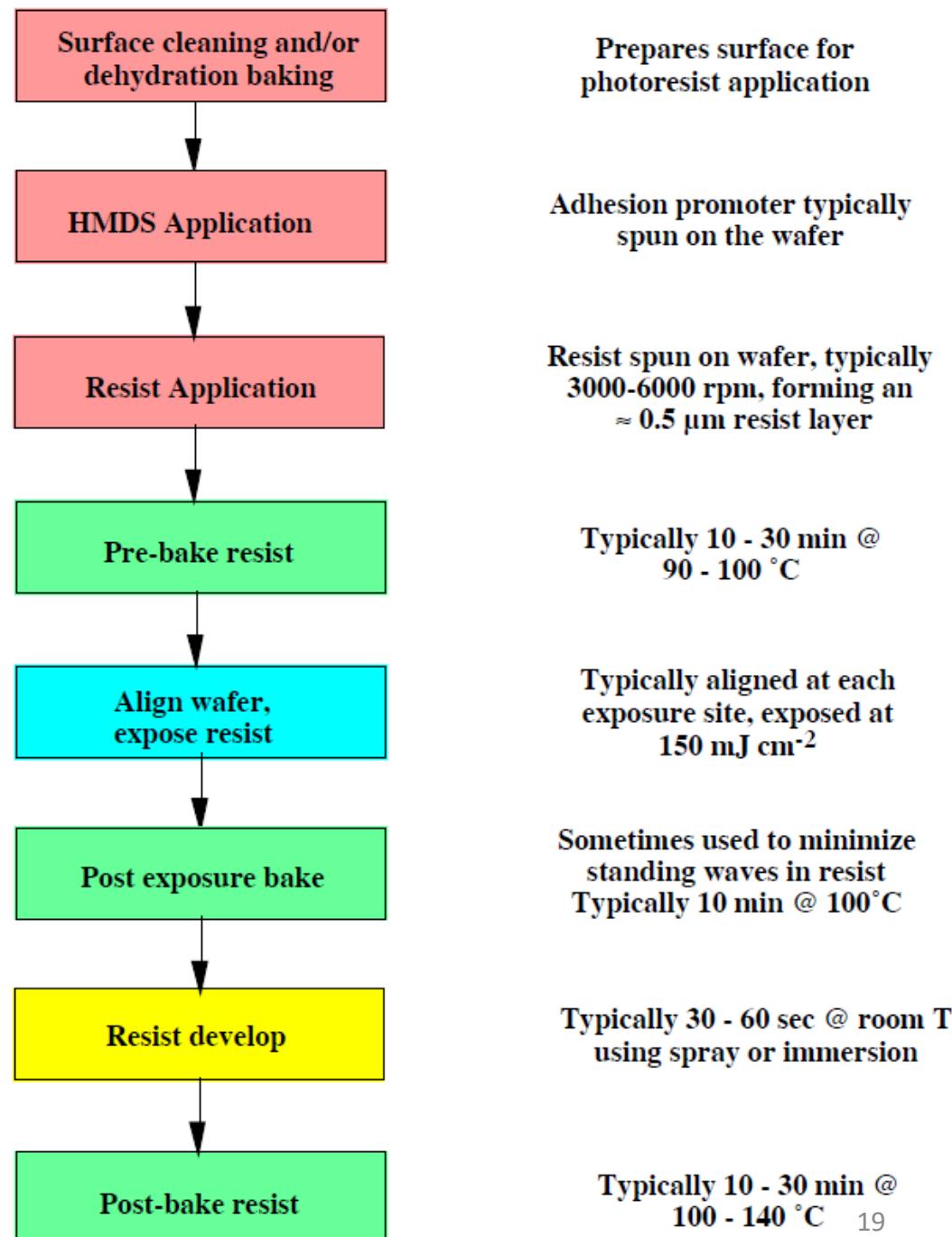


SU-8 grayscale structures

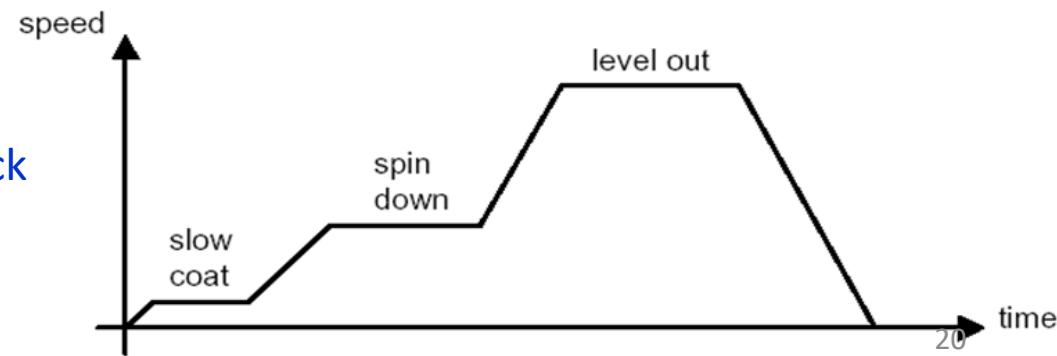
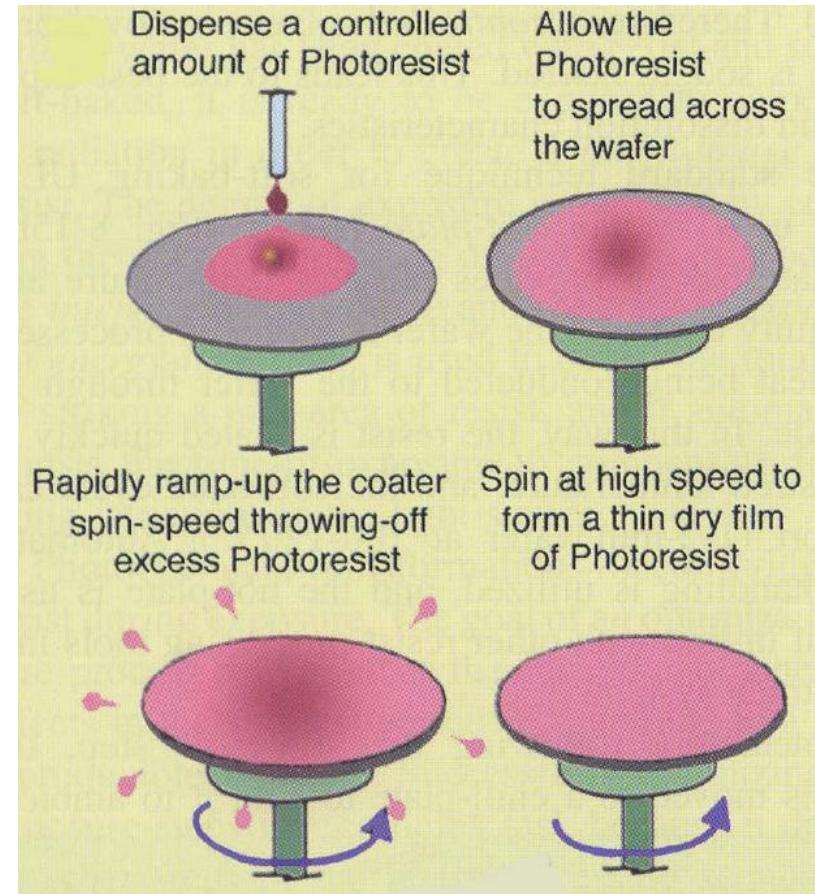
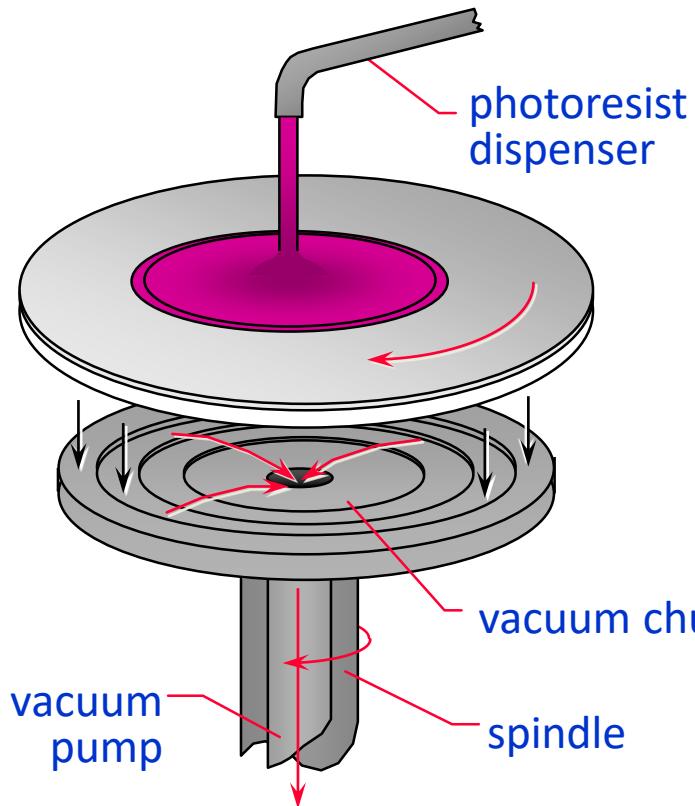
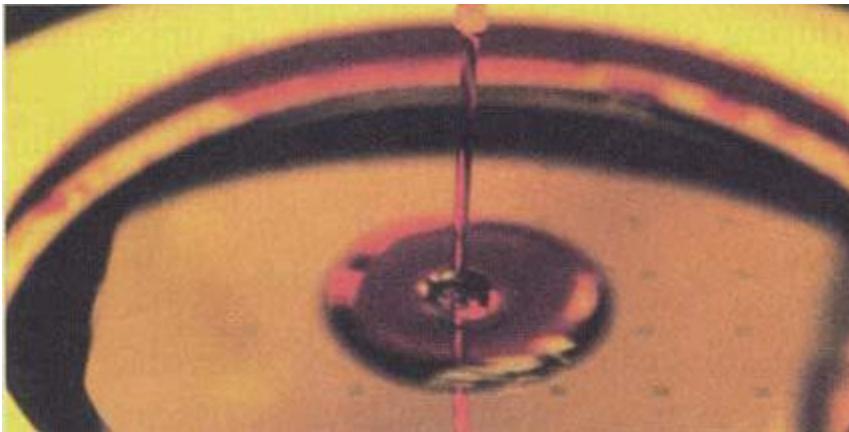
Chapter 5 Lithography

1. Introduction and application.
2. Light source and photomask, alignment.
3. Photolithography systems.
4. Resolution, depth of focus, modulation transfer function.
5. Other lithography issues: none-flat wafer, standing wave...
6. Photoresist.
7. Resist sensitivity, contrast and gray-scale photolithography.
8. Step-by-step process of photolithography.

Typical photoresist process flow for DNQ g-line and i-line positive resist



Photoresist spin coating

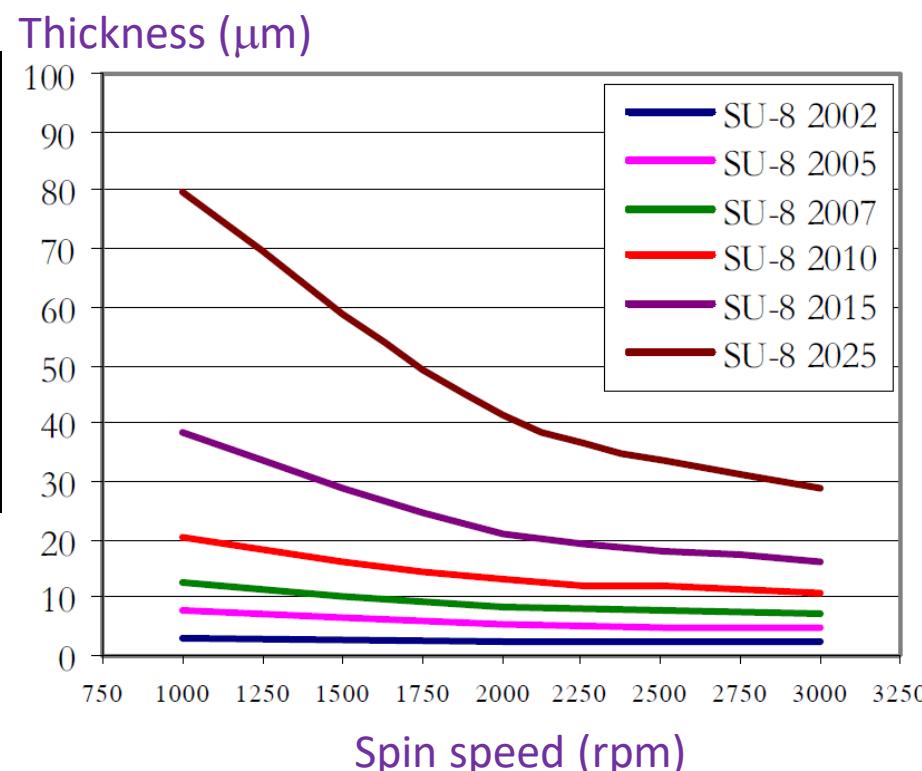


The physics of spin

- Resist thickness after the spin process is related to properties of the resist and the spin.
- The most important factor is resist viscosity and solvent evaporation rate (how volatile it is).
- Resist won't be uniform if the solvent evaporates too fast.
- Thickness will approach zero if solvent never evaporates (assume long-enough spin time).
- Typically after 20sec spinning at peak speed, film thickness becomes stable (i.e. almost no solvent left).

Photoresist properties	Spin properties
Viscosity	Spin Speed
Drying characteristics	Acceleration Rate
Dispense volume	Spin Time

SU-8 is very thick resist (typical resist only ~1μm thick). Here the series has different amount of solvent. More solvent, less viscosity, thinner film.



Two primary pattern transfer techniques

Direct etch:

Photoresist is applied on top of the layer to be patterned.

Unwanted material is etched away, using resist as mask.

Lift-off:

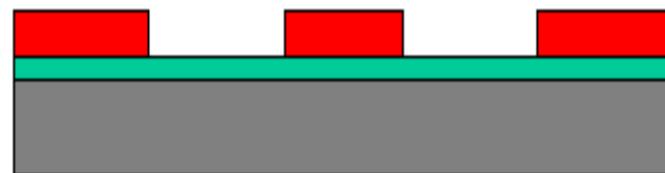
Patterned layer is deposited on top of the photoresist.

Unwanted material is lifted off when resist is removed.

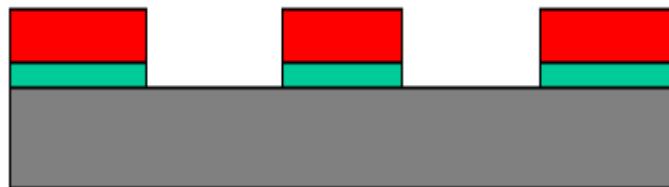
Direct etch: (using resist as etching mask)



deposit thin film of desired material



coat and pattern photoresist



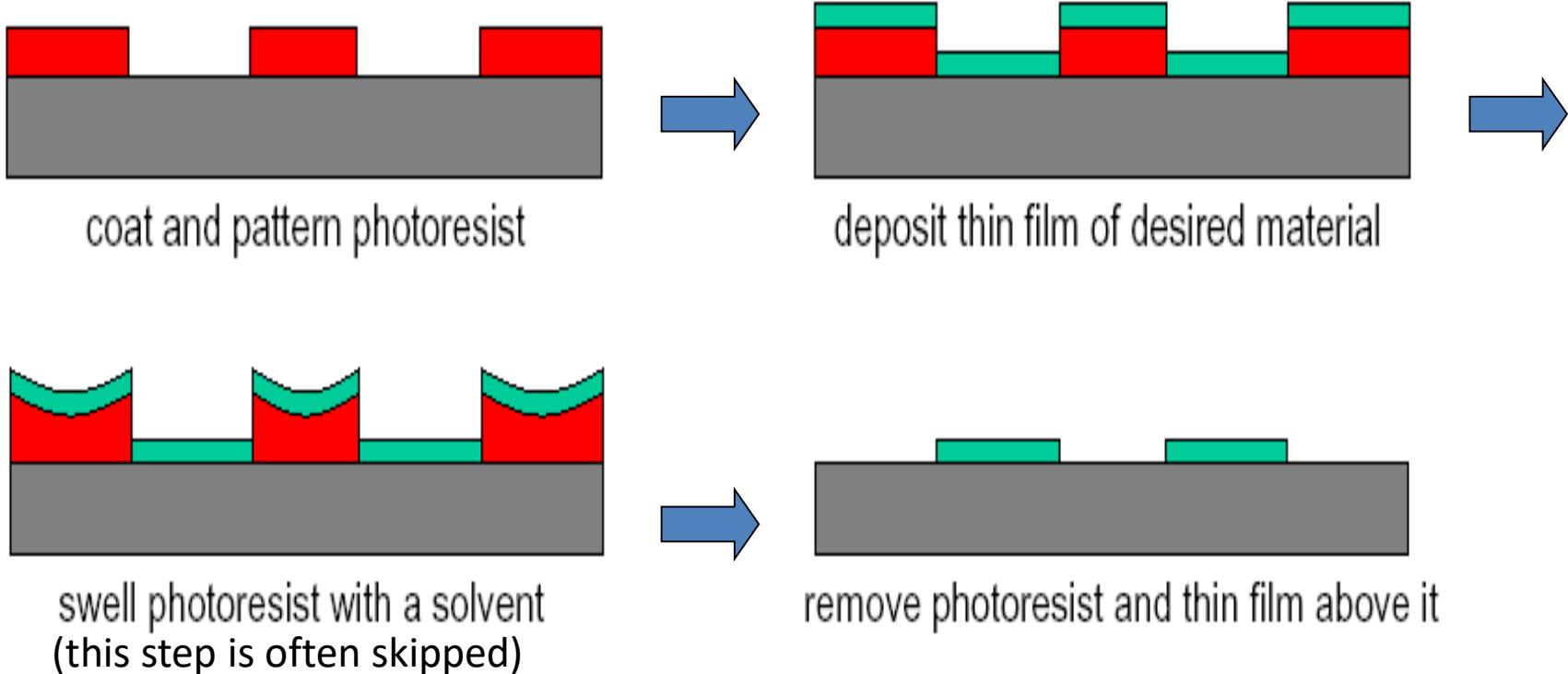
etch film using photoresist as mask



remove photoresist

Usually anisotropic dry etch, though wet etch
(isotropic) is also OK for low-resolution applications.

Lift-off



- For R&D, liftoff is very popular since it patterns metals easily.
- Most metals are hard to etch by dry plasma etch (reactive ion etching), then liftoff is the only method. (dry etch is anisotropic, thus maintaining feature size/resolution)
- But for low resolution application, direct etch method using wet etching is OK. (metal can be etched easily using acids..., but it is isotropic and pattern widens due to lateral etch)
- Liftoff is almost never used for industry, due to the low yield (pattern edge not “clean”, metal debris in liftoff solution that fall on other parts of the wafer).

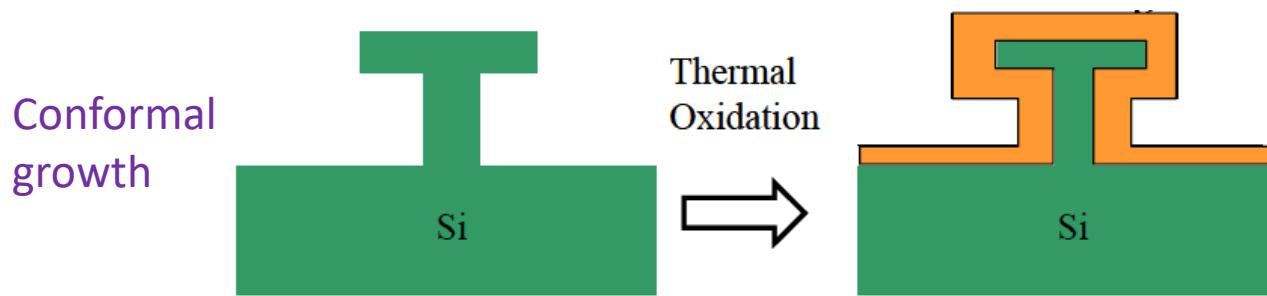
Chapter 6 Thermal oxidation and the Si/SiO₂ interface



1. SiO₂ properties and applications.
2. Manufacturing methods and equipment.
3. Measurement methods.
4. Deal-grove model (linear parabolic model).
5. Thin oxide growth, dependence on gas pressure and crystal orientation
6. Cl-containing gas, 2D growth, substrate doping effect .
7. Interface charges, dopant redistribution.

Properties of thermally grown SiO_2

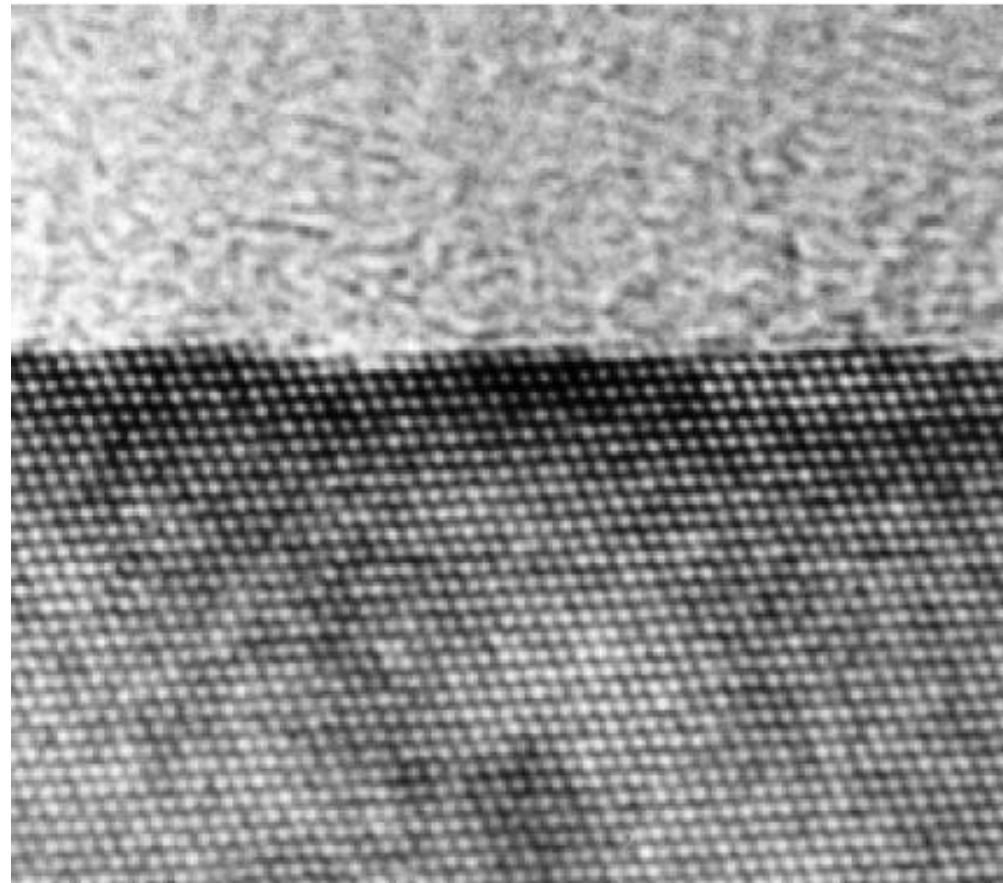
- It is amorphous.
- Stable, reproducible and conformal SiO_2 growth
- Melting point: 1700°C
- Density: 2.21 g/cm^3 (almost the same as Si that is 2.33 g/cm^3)
- Crystalline SiO_2 [Quartz] = 2.65 gm/cm^3
- Atomic density: $2.3 \times 10^{22} \text{ molecules/cm}^3$
(For Si, it is $5 \times 10^{22} \text{ atoms/cm}^3$)
- Refractive index: $n=1.46$
- Dielectric constant: $\epsilon=3.9$ (why not $=n^2$?)
- Excellent electrical insulator: resistivity $\rho > 10^{20} \Omega\text{cm}$, energy gap $E_g=8-9 \text{ eV}$.
- High breakdown electric field: $>10^7 \text{ V/cm}$



The Si/SiO₂ interface

Thermal oxide
(amorphous)

Si substrate
(single crystal)

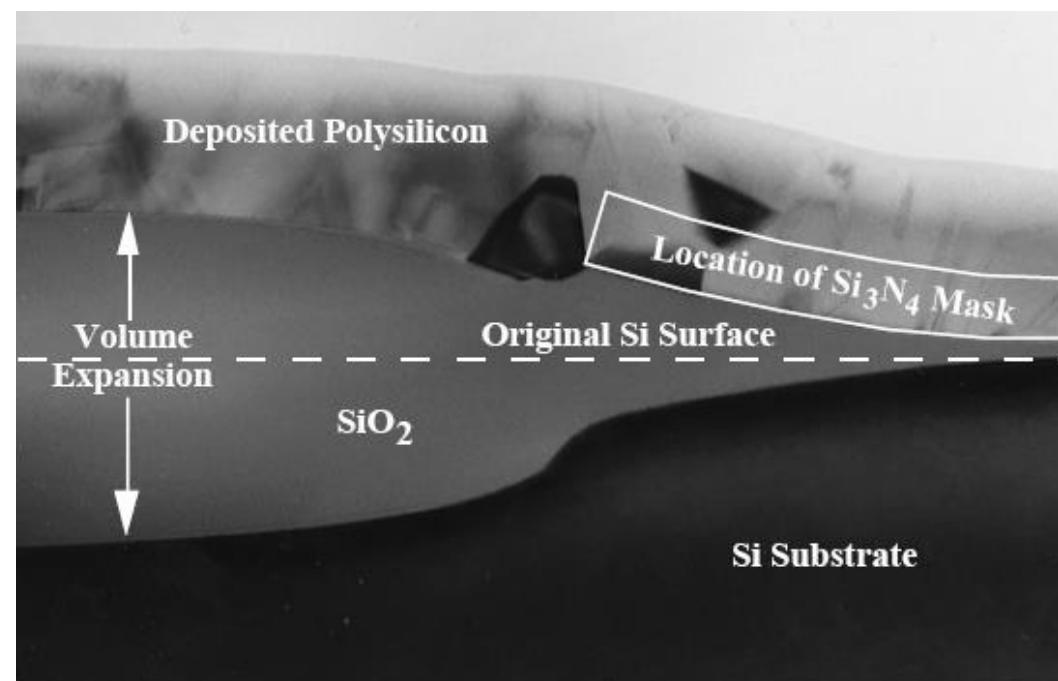
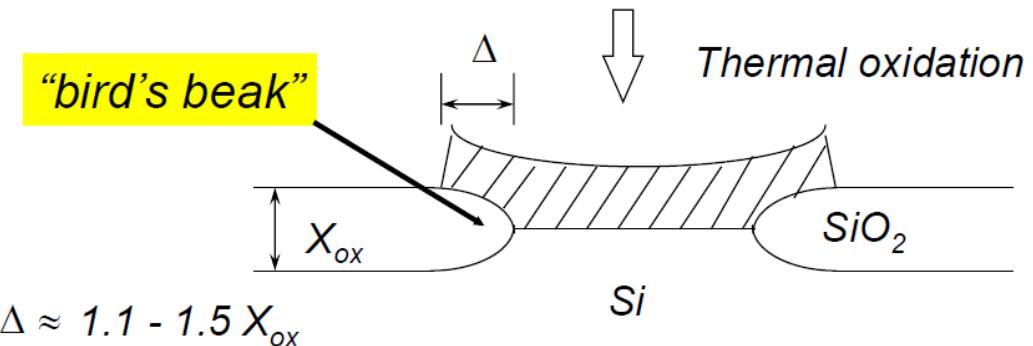
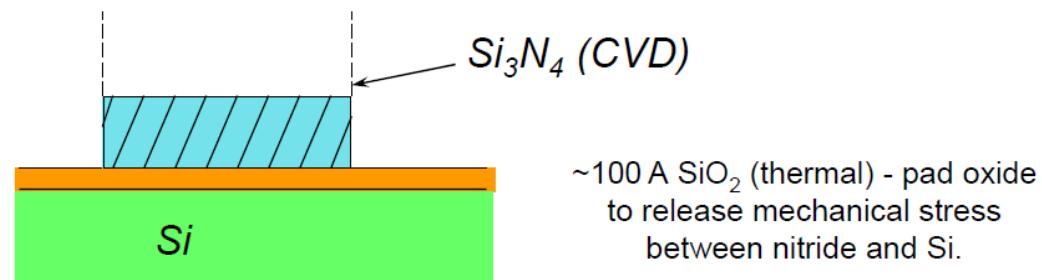
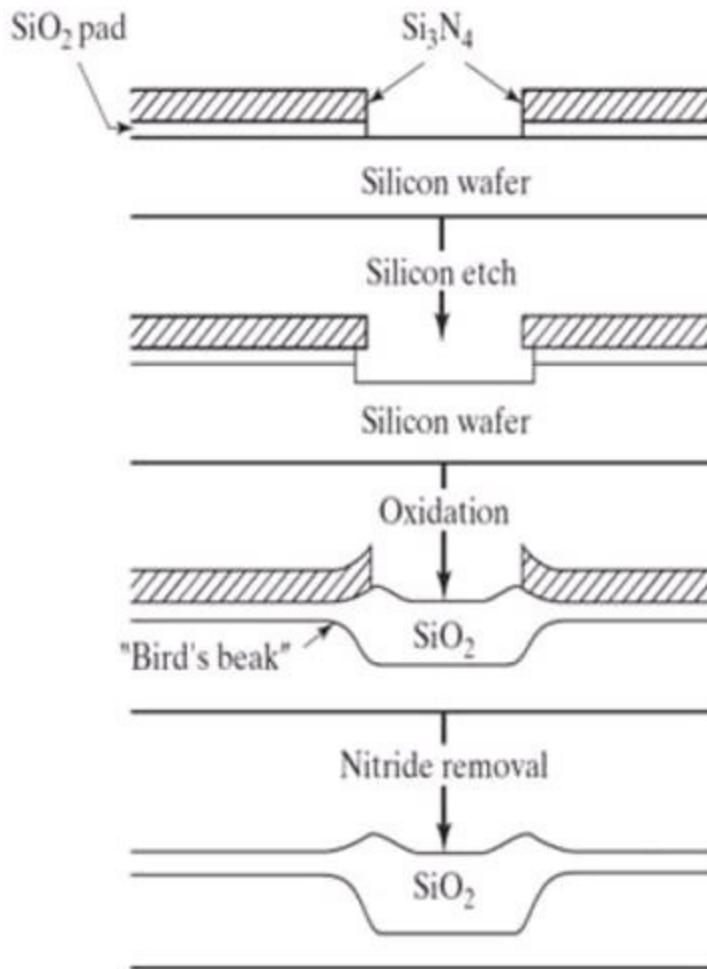


(Photo courtesy of J. Bravman.)

The perfect interface between Si and SiO₂ is one major reason why Si is used for semiconductor devices (instead of Ge...)

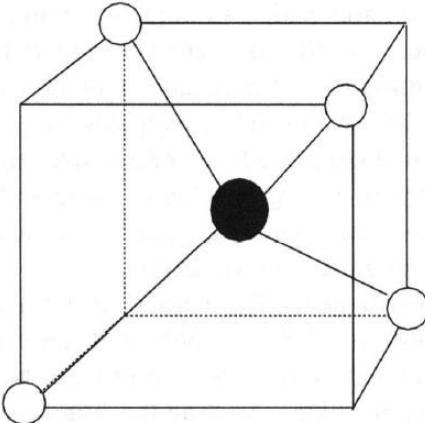
Local Oxidation of Si (LOCOS)

Fully recessed process attempts to minimize bird's peak.

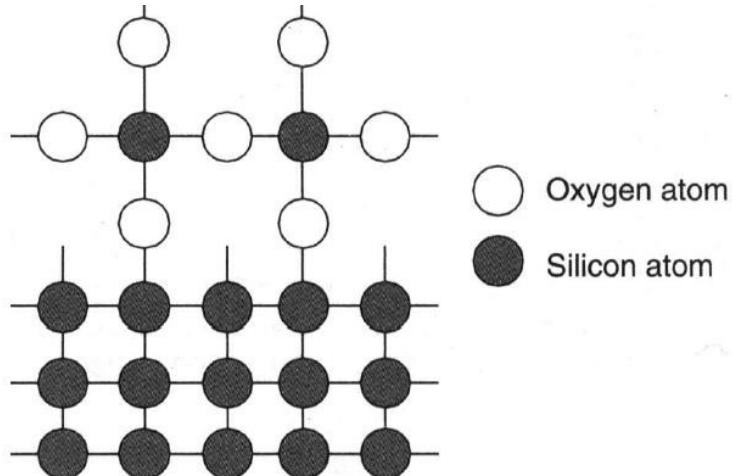


(Photo courtesy of J. Bravman.)

Oxide Structure

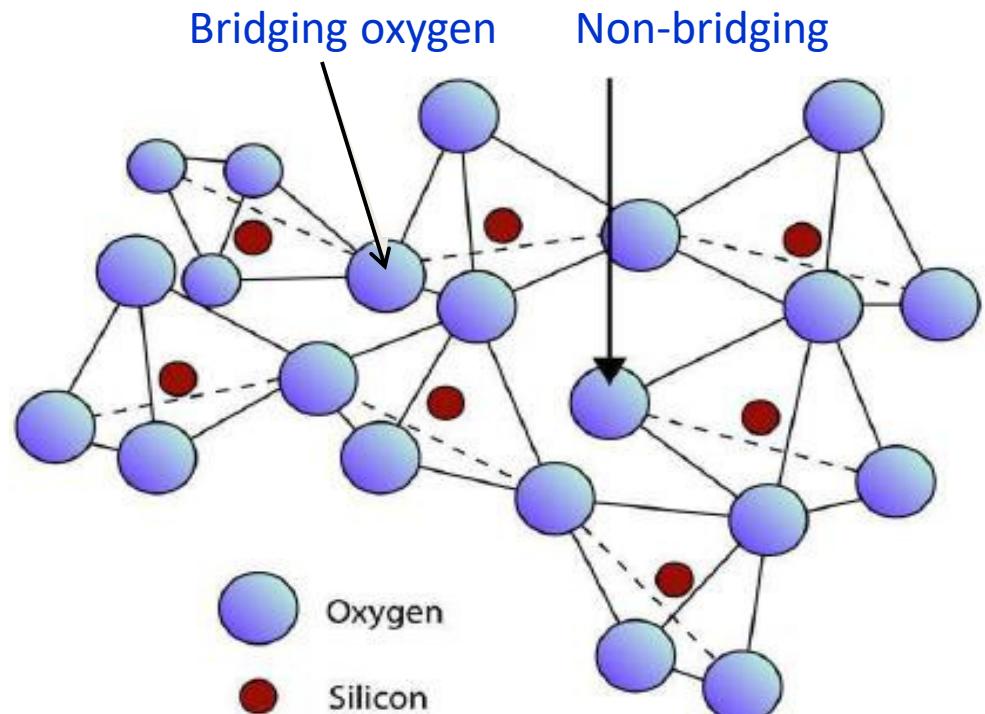


Basic structure of silica: a silicon atom tetrahedrally bonds to four oxygen atoms

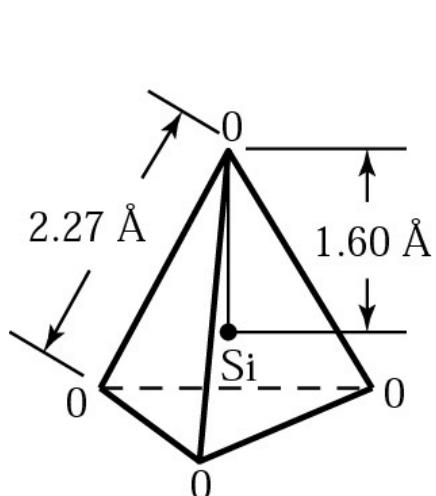


The structure of silicon-silicon dioxide interface: some silicon atoms have dangling bonds.

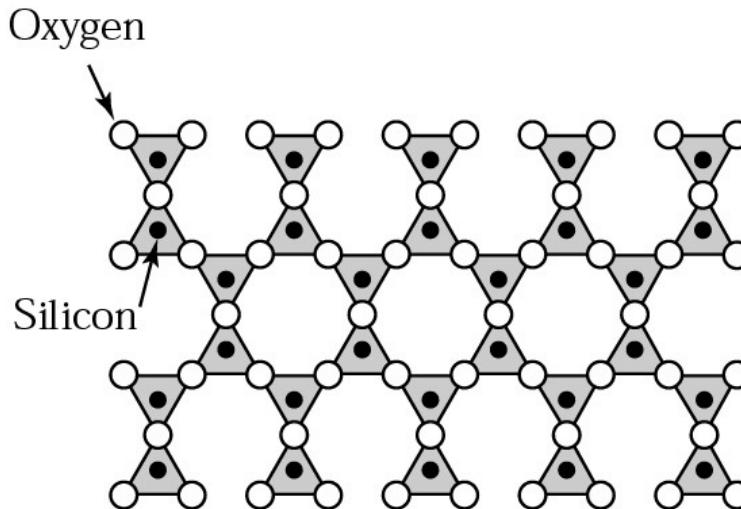
Amorphous tetrahedral network



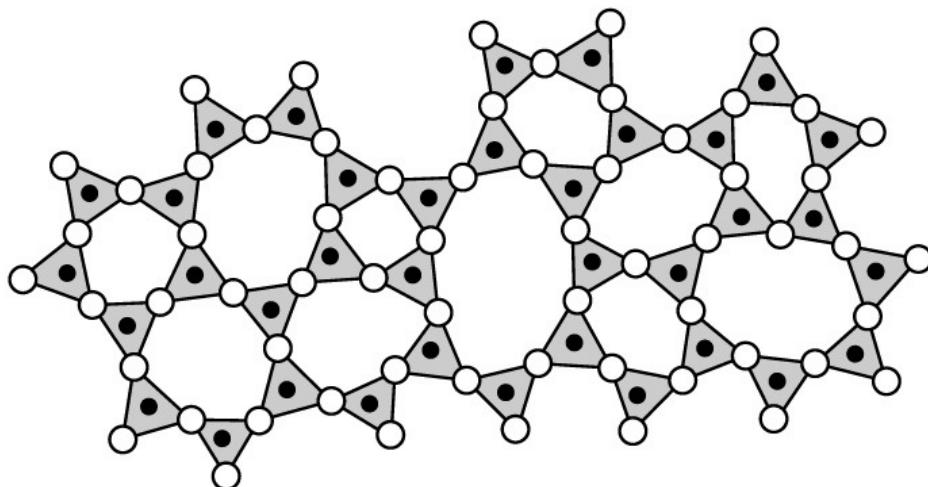
Oxide Structure



(a)



(b)

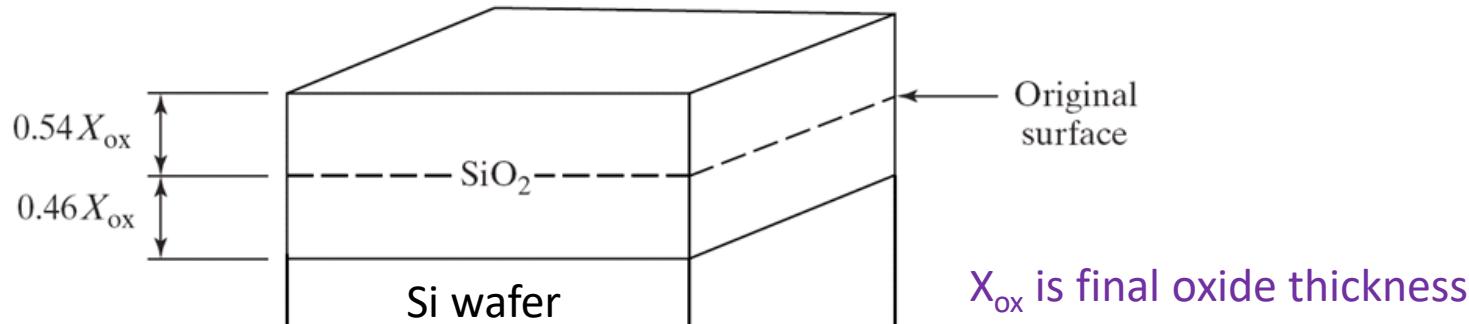


(c)

Dry and wet oxidation

Dry oxidation: $\text{Si(s)} + \text{O}_2\text{(g)} \rightarrow \text{SiO}_2\text{(s)}$; Wet/steam oxidation: $\text{Si(s)} + 2\text{H}_2\text{O(g)} \rightarrow \text{SiO}_2\text{(s)} + 2\text{H}_2\text{(g)}$

- Both typically 900-1200°C, wet oxidation is about 10x faster than dry oxidation.
- **Dry oxide:** thin 0.05-0.5μm, excellent insulator, for gate oxides; for very thin gate oxides, may add nitrogen to form oxynitrides.
- **Wet oxide:** thick <2.5 μm, good insulator, for field oxides or masking. Quality suffers due to the diffusion of the hydrogen gas out of the film, which creates paths that electrons can follow.
- Room temperature Si in air creates “native oxide”: very thin ~1-2nm, poor insulator, but can impede surface processing of Si.
- Volume expansion by 2.2x (=1/0.46), so SiO_2 film has compressive stress.

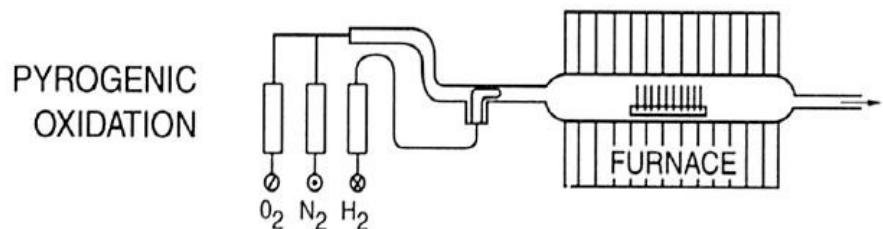
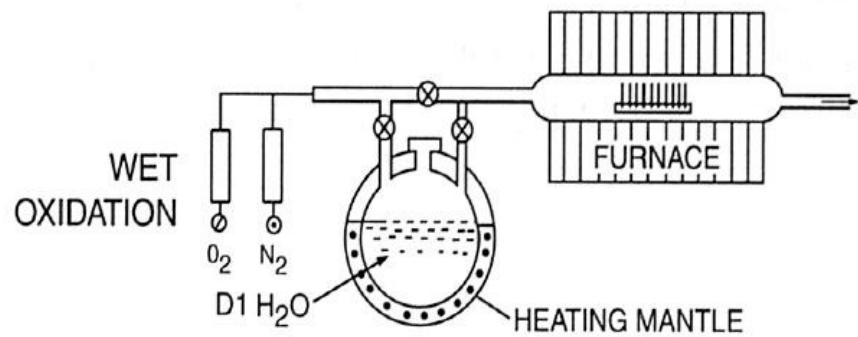
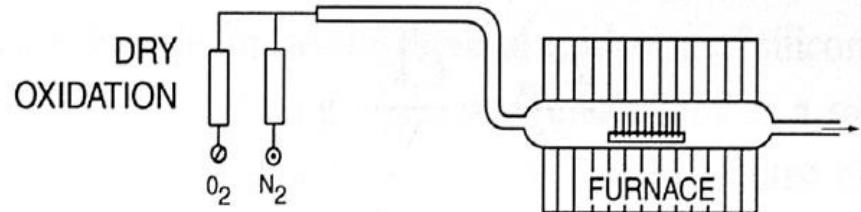


$$\frac{\text{Thickness of Si}}{\text{Thickness of } \text{SiO}_2} = \frac{\text{Molar volume (Si)}}{\text{Molar volume (SiO}_2)} = \frac{\frac{\text{Molecular weight (Si)}}{\text{Density (Si)}}}{\frac{\text{Molecular weight (SiO}_2)}{\text{Density (SiO}_2)}} = \frac{\frac{28.9 \text{ g/mol}}{2.33 \text{ g/cm}^3}}{\frac{60.08 \text{ g/mol}}{2.21 \text{ g/cm}^3}} = 0.46$$

Chapter 6 Thermal oxidation and the Si/SiO₂ interface

1. SiO₂ properties and applications.
2. Manufacturing methods and equipment.
3. Measurement methods.
4. Deal-grove model (linear parabolic model).
5. Thin oxide growth, dependence on gas pressure and crystal orientation
6. Cl-containing gas, 2D growth, substrate doping effect .
7. Interface charges, dopant redistribution.

Thermal silicon oxidation methods



Wet oxidation using H₂ and O₂ is more popular (cleaner) than using H₂O vapor.

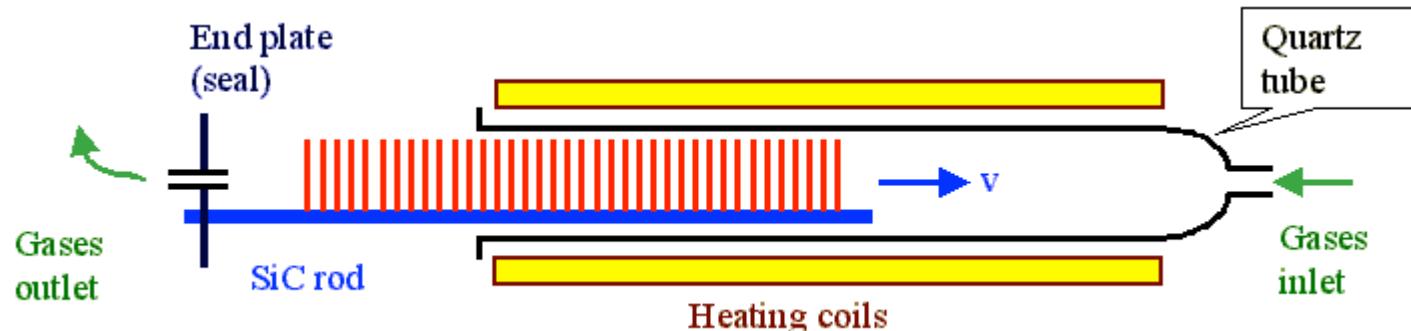
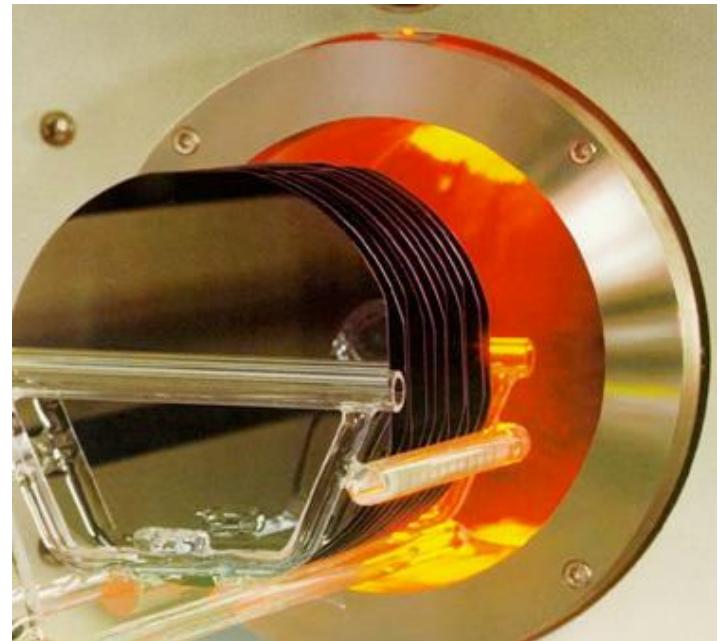


A three-tube horizontal furnace with multi-zone temperature control



Vertical furnace
(not popular)

Thermal oxidation equipment



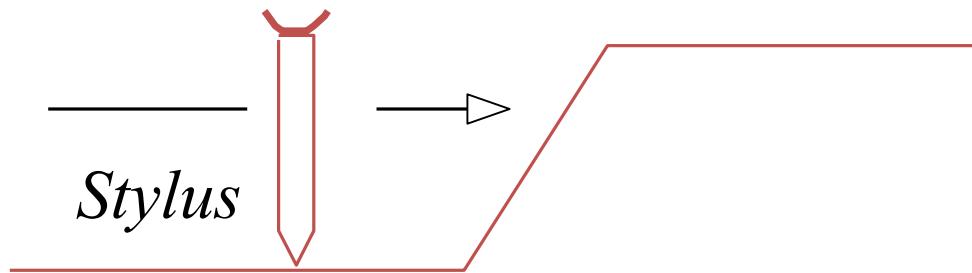
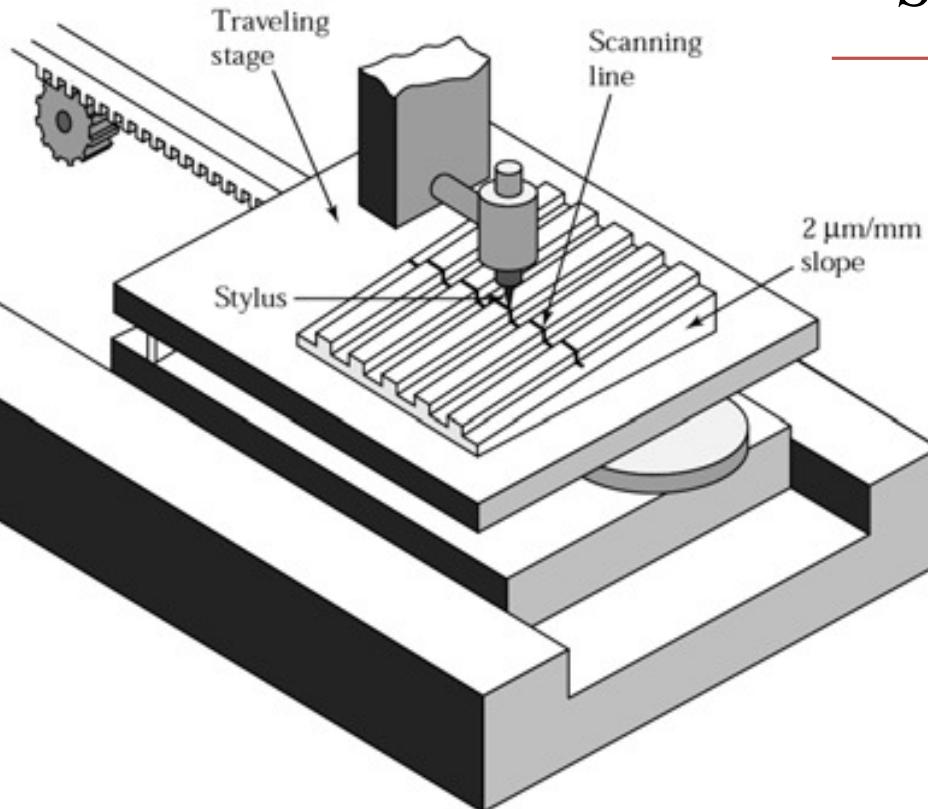
- The tubular reactor made of quartz or glass, heated by resistance.
- Oxygen or water vapor flows through the reactor and past the silicon wafers, with a typical velocity of order 1cm/s.

Chapter 6 Thermal oxidation and the Si/SiO₂ interface

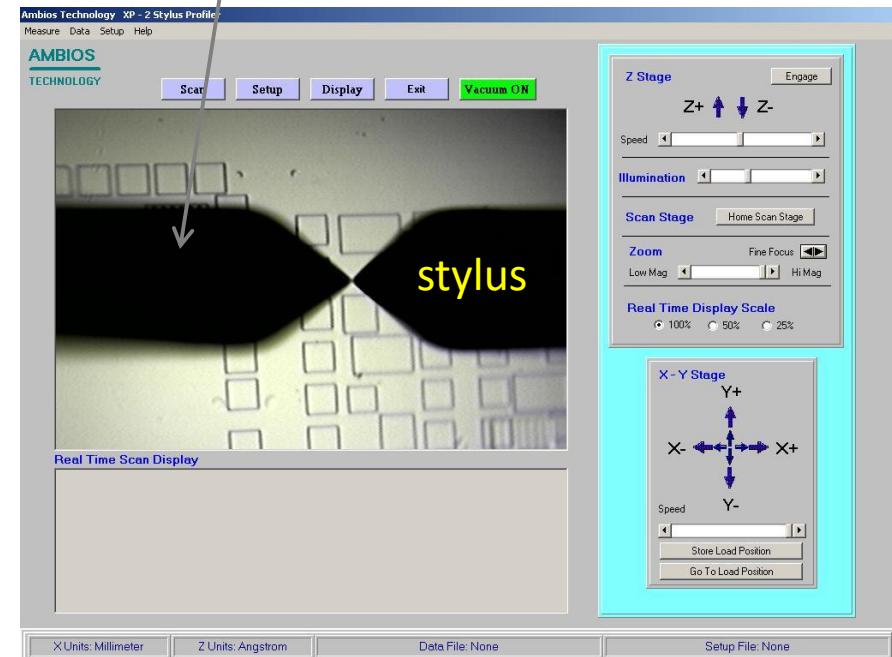
1. SiO₂ properties and applications.
2. Manufacturing methods and equipment.
3. Measurement methods (mechanical, optical, electrical).
4. Deal-grove model (linear parabolic model).
5. Thin oxide growth, dependence on gas pressure and crystal orientation
6. Cl-containing gas, 2D growth, substrate doping effect .
7. Interface charges, dopant redistribution.

Surface profilometry (Dektak): mechanical thickness measurement

Oxide etched away by HF over part of the wafer and a mechanical stylus is dragged over the resulting step.

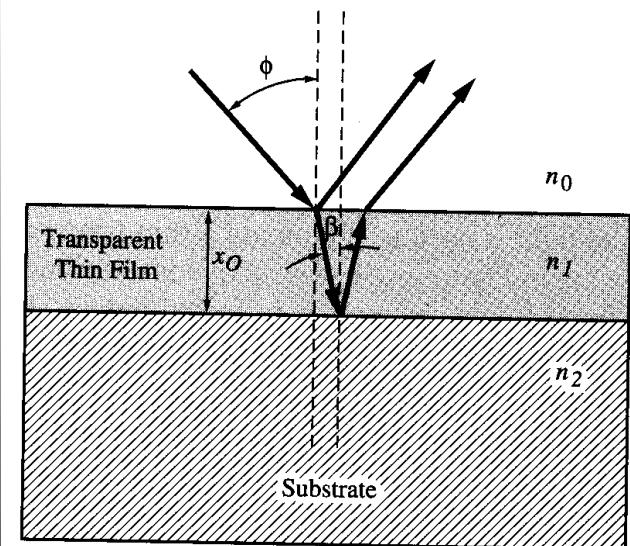
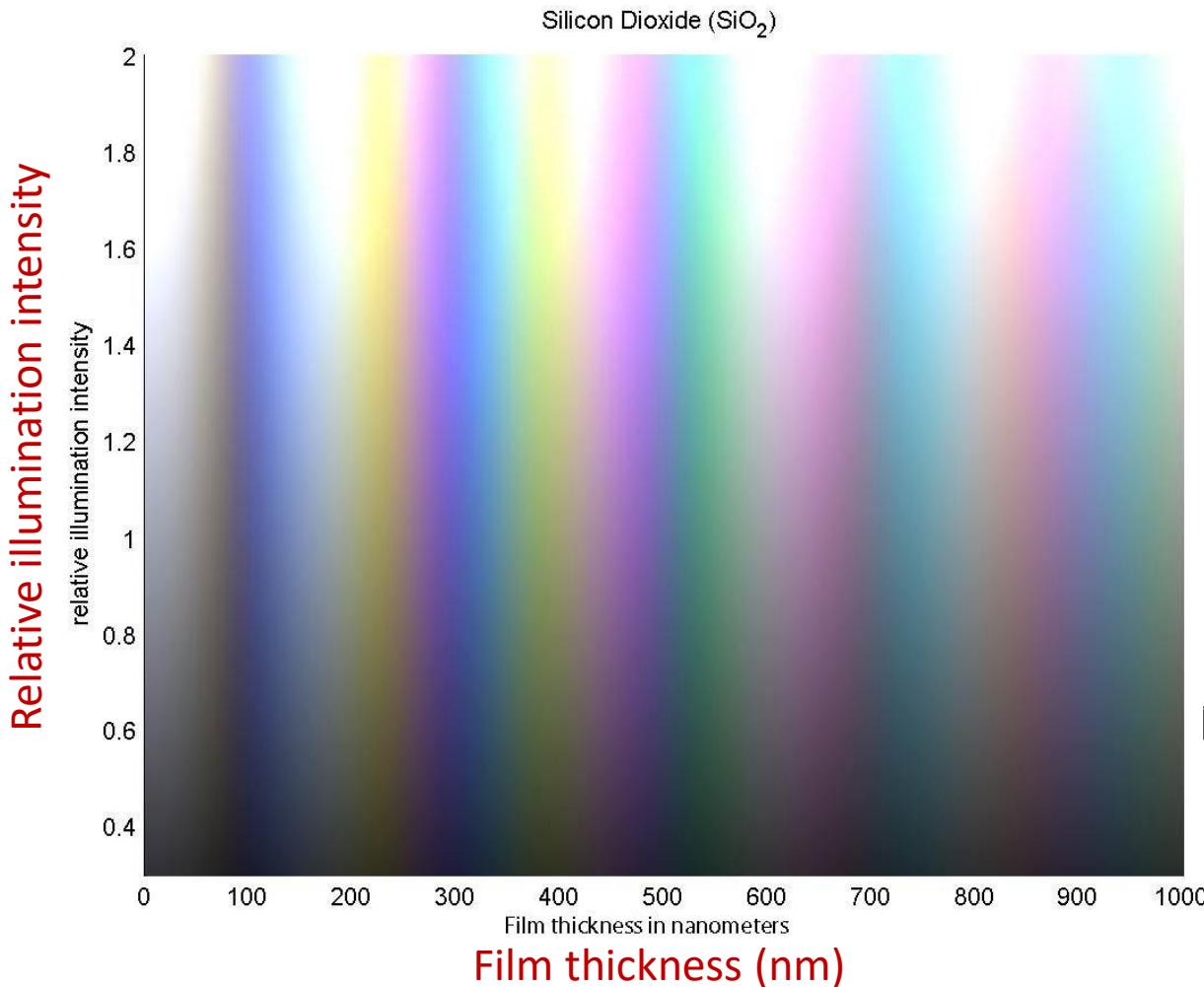


Mirror image of stylus



AFM can also be used for thickness measurement.
(AFM: atomic force microscopy)

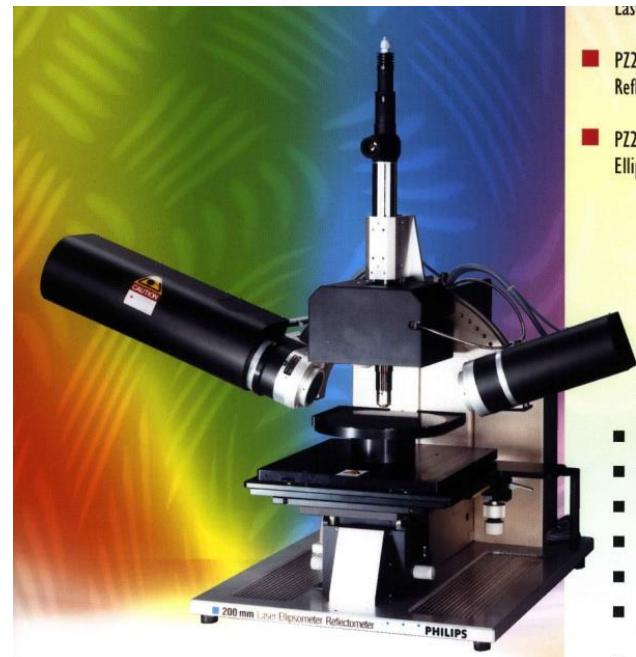
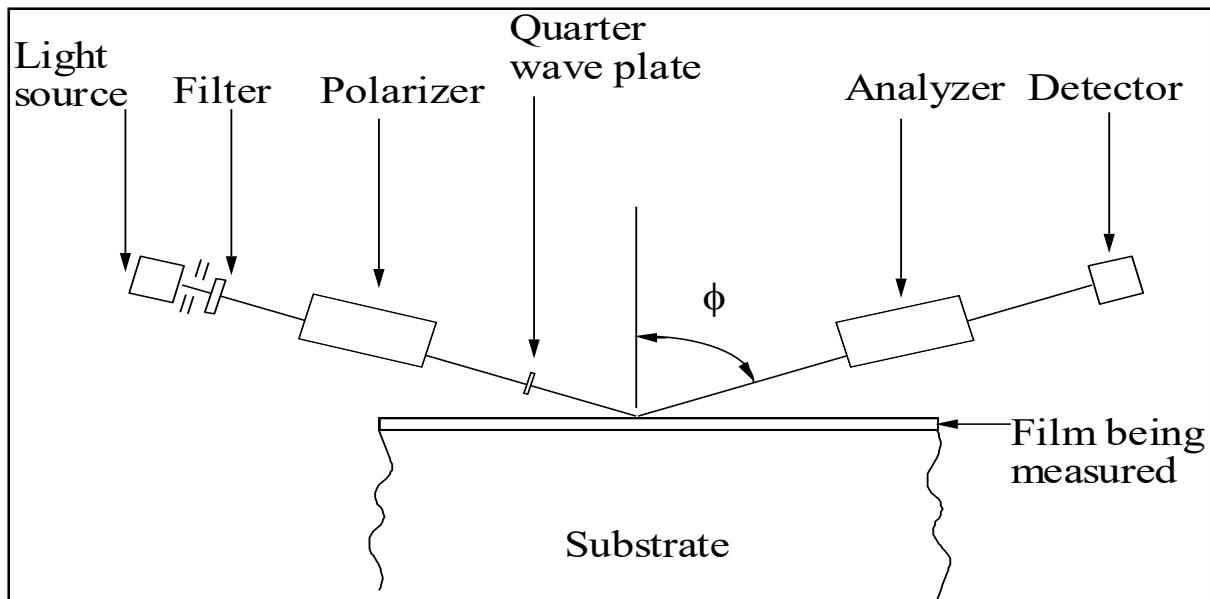
Thickness determination by looking the color



- Oxide thickness for constructive interference (viewed from above $\phi=0^\circ$) $x_o=k\lambda/2n$, $n=1.46$, $k=1, 2, 3\dots$
- Our eye can tell the color difference between two films having 10nm thickness difference.

Optical thickness measurement: ellipsometry

Very accurate (1nm accuracy)



- After quarter wave plate, the linear polarized light becomes circular polarized, which is incident on the oxide covered wafer.
- The polarization of the reflected light, which depends on the thickness and refractive index (usually known) of the oxide layer, is determined and used to calculate the oxide thickness.
- Multiple wavelengths/incident angles can be used to measure thickness/refractive index of each film in a multi-film stack.

Chapter 6 Thermal oxidation and the Si/SiO₂ interface

1. SiO₂ properties and applications.
2. Manufacturing methods and equipment.
3. Measurement methods (mechanical, optical, electrical).
4. Deal-Grove model (linear parabolic model).
5. Thin oxide growth, dependence on gas pressure and crystal orientation
6. Cl-containing gas, 2D growth, substrate doping effect .
7. Interface charges, dopant redistribution.

Bruce Deal and Andrew Grove



Time Magazine's Man of the Year (1997)

As a young researcher at Fairchild Semiconductor, he “wrote the book” on SiO₂ growth: the *Deal-Grove model*.

Grove worked at Fairchild Semiconductor before becoming the fourth employee at the nascent Intel Corporation. He became Intel's president in 1979, its CEO in 1987, and its Chairman and CEO in 1997.

Grove is credited with having transformed Intel from a manufacturer of memory chips into one of the world's dominant producers of microprocessors. During his tenure as CEO, Grove oversaw a 4,500% increase in Intel's market capitalization from \$4 billion to \$197 billion, making it, at the time, the world's most valuable company.

Overview

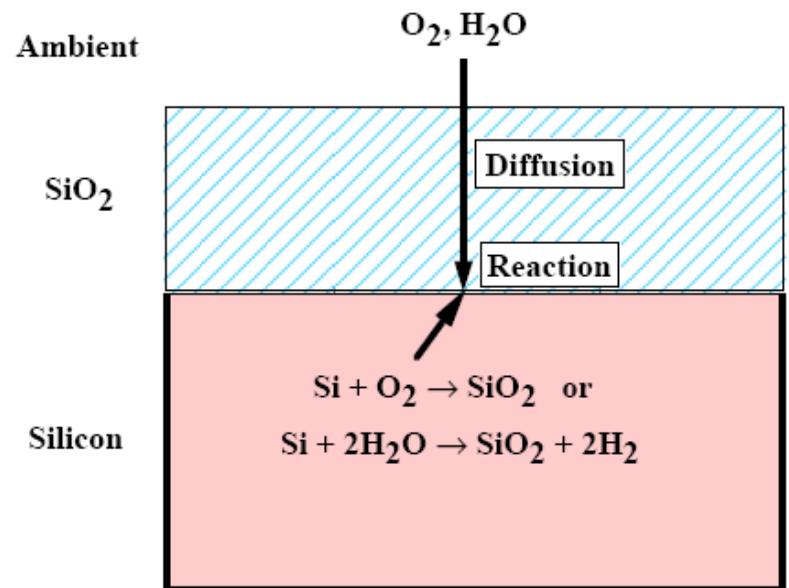
Deal-Grove model: linear-parabolic model for 1D oxide growth (along z-direction) from un-patterned wafer surface.

It is applicable to:

- Oxidation temperature 700-1200°C
- Gas pressure at wafer surface 0.1-25 atm.
- Oxide film thickness 20-2000nm.

O gets in or Si gets out for reaction?

H_2O is for *wet oxidation*
 O_2 for *dry oxidation*

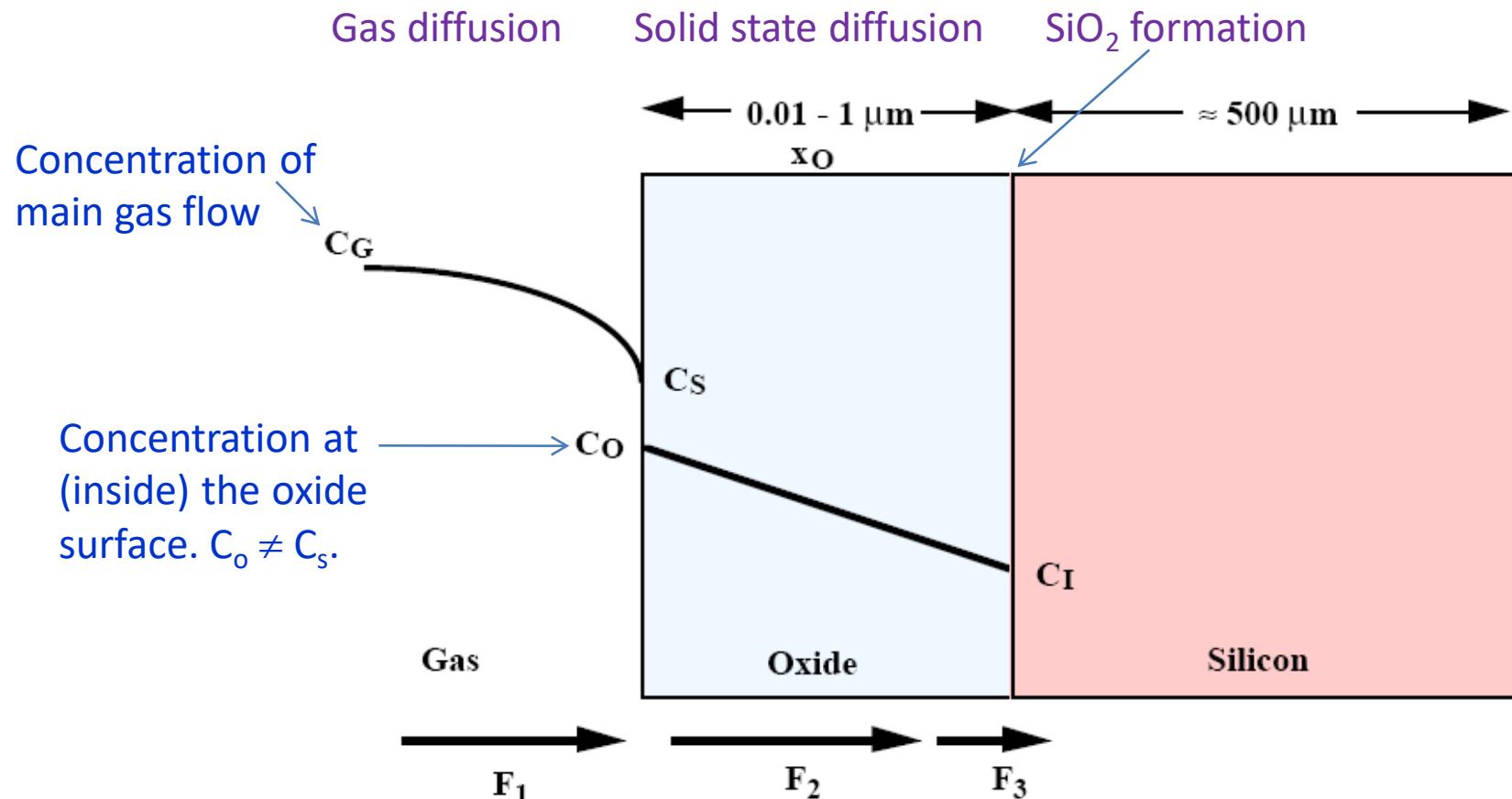


Considering dry oxygen molecules as the oxidant species, by radio active tracer, it has been shown that oxidation proceeds by inward movement of O_2 molecules through SiO_2 to the Si – SiO_2 interface where the reaction $\text{Si}(\text{solid}) + \text{O}(\text{gas}) \rightarrow \text{SiO}_2$ takes place.

This forms a contrast with the case of Copper whose oxidation proceeds by the outward motion of the metallic ion and also with the case of anodic oxidation of silicon, where silicon moves outward.

D-G model: three flux at equilibrium

F (oxidant flux): number/(cm²-s); C (oxidant concentration): number/cm³



F_1 : flux of oxidizing species transported from the gas phase to the gas-oxide interface.

F_2 : flux across the existing oxide toward the silicon substrate.

F_3 : flux reacting at the SiO_2 interface (this term is not strictly flux/flow).

In steady state, the three fluxes are equal.

D-G model: three flux at equilibrium

$$F_1 = h_g (C_G - C_s)$$

h_g is mass transfer coefficient in cm/sec.

For ideal gas, $P_s V = N K_B T$, so $C_s = N/V = P_s/KT$.

According to Henry's law, the concentration of a gas species dissolved in a solid is proportional to the partial pressure of that species at the solid surface. So $C_o = H P_s$, where H is Henry's constant.

(this is something similar to the law of segregation of dopant at an interface)

$$F_1 = h_g \left(C_G - \frac{P_s}{kT} \right) = h_g \left(C_G - \frac{C_o}{H kT} \right)$$

Define $h = h_g / H kT$, $C^* = H kT C_G = H P_G$, then

$$F_1 = h(C^* - C_o)$$

D-G model: three flux at equilibrium

Using Fick's law of diffusion, D (cm²/sec) is oxidant diffusivity in the oxide, x is oxide thickness, then

$$F_2 = D \frac{C_o - C_I}{x}$$

The effective diffusivities of both O₂ and H₂O are on the same order (about 5×10³μm²/hr at 1100°C).

The rate of reaction should be proportional to the oxidant concentration at the Si/SiO₂ interface C_I , K_s interface reaction rate constant (cm/sec), then

$$F_3 = k_s C_I$$

In steady state, the three fluxes are equal.

$$F_1 = F_2 = F_3$$

Now we have two equations to solve two unknown C_o and C_I .

$$F_1 = h(C^* - C_o)$$

$$F_2 = D \frac{C_o - C_I}{x}$$

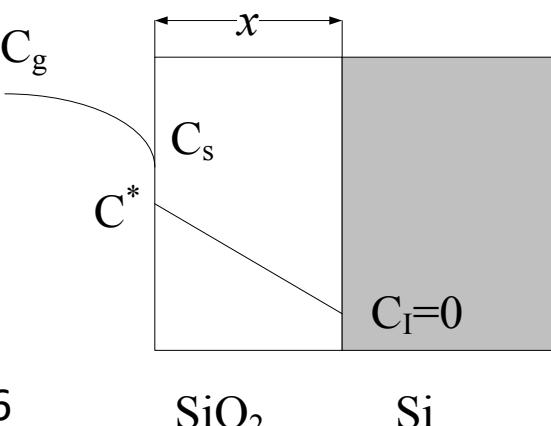
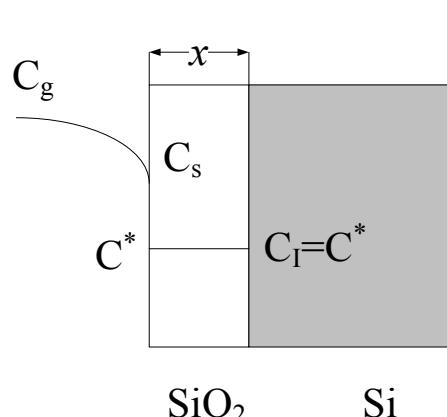
$$F_3 = k_s C_I$$

Deal-Grove model $F_1 = F_2 = F_3$

$$C_I = \frac{C^*}{1 + \frac{k_s}{h} + \frac{k_s x}{D}} \approx \frac{C^*}{1 + \frac{k_s x}{D}} \quad (h \gg k_s)$$

$$C_O = \frac{\left(1 + \frac{k_s x}{D}\right) C^*}{1 + \frac{k_s}{h} + \frac{k_s x}{D}} \approx C^*, \text{ since } h \gg k_s$$

This means the gas absorption rate at the oxide surface is much faster than chemistry occurring at the Si/SiO₂ interface.



$k_s x / D \ll 1$, reaction rate limited

$k_s x / D \gg 1$, diffusion limited

$k_s x / D \approx 1$
corresponds to
oxide thickness
50-200 nm.

Deal-Grove model

Assume N_1 as the number of oxidant molecules incorporated per unit volume of oxide grown. For O_2 (dry) oxidation, $N_1=2.2 \times 10^{22} \text{ cm}^{-3}$; for H_2O (wet) oxidation, $N_1=4.4 \times 10^{22} \text{ cm}^{-3}$ ($2\times$ that of dry, since H_2O has only one oxygen atom).

Then the oxidation rate R (F is flux, $F=F_1=F_2=F_3=K_s C_l$, unit is number/ $(\text{cm}^2 \cdot \text{s})$; unit for R is cm/s).

$$R = \frac{F}{N_1} = \frac{dx}{dt} = \frac{k_s C^*}{N_1 \left[1 + \frac{k_s}{h} + \frac{k_s x}{D} \right]}$$

$$N_1 \int_{x_i}^{x_0} \left[1 + \frac{k_s}{h} + \frac{k_s x}{D} \right] dx = \int_0^t k_s C^* dt$$

Let $B=2DC^*/N_1$, $A=2D(1/k_s+1/h)$,
 $B/A \approx C^*k_s/N_1$, ($h \gg k_s$)

$$\frac{x_o^2 - x_i^2}{B} + \frac{x_o - x_i}{B/A} = t$$

X_i account for any oxide present at the start of the oxidation.

Deal-Grove model

Rewrite the above equation,

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

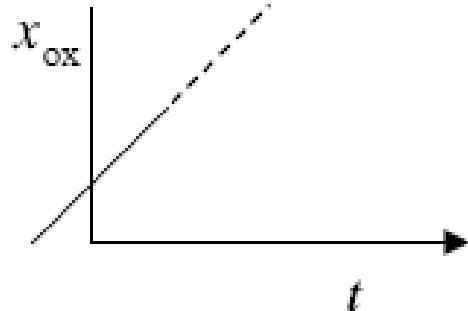
$$\tau = \frac{x_i^2 + Ax_i}{B}$$

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

$B = 2DC^*/N_1$ —parabolic rate constant, contribution of flux F_2 .
 $B/A \approx C^*k_s/N_1$ —linear rate constant, contribution of flux F_3 .

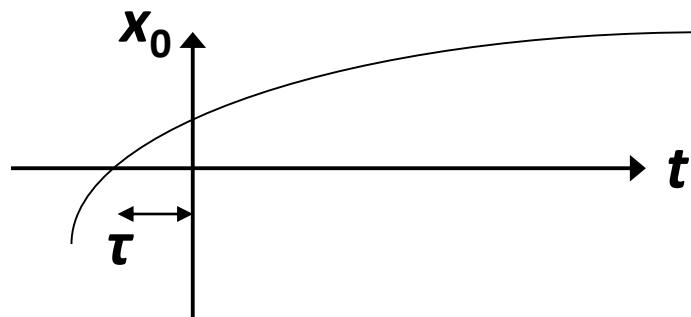
Two limiting forms, when one of the two terms dominates.

For short time/thin oxide, linear.



$$x_o \approx \frac{B}{A}(t + \tau)$$

For long time/thick oxide, parabolic.



$$x_o^2 \approx B(t + \tau)$$

Arrhenius expression for B and B/A

D-G model is applicable for: oxidation on flat un-patterned surface, lightly doped substrates, using simple O₂ or H₂O ambient, and when the oxide thickness is larger than about 20nm.

Experimentally it is found:

$$B = C_1 \exp(-E_1 / kT)$$

$$\frac{B}{A} = C_2 \exp(-E_2 / kT)$$

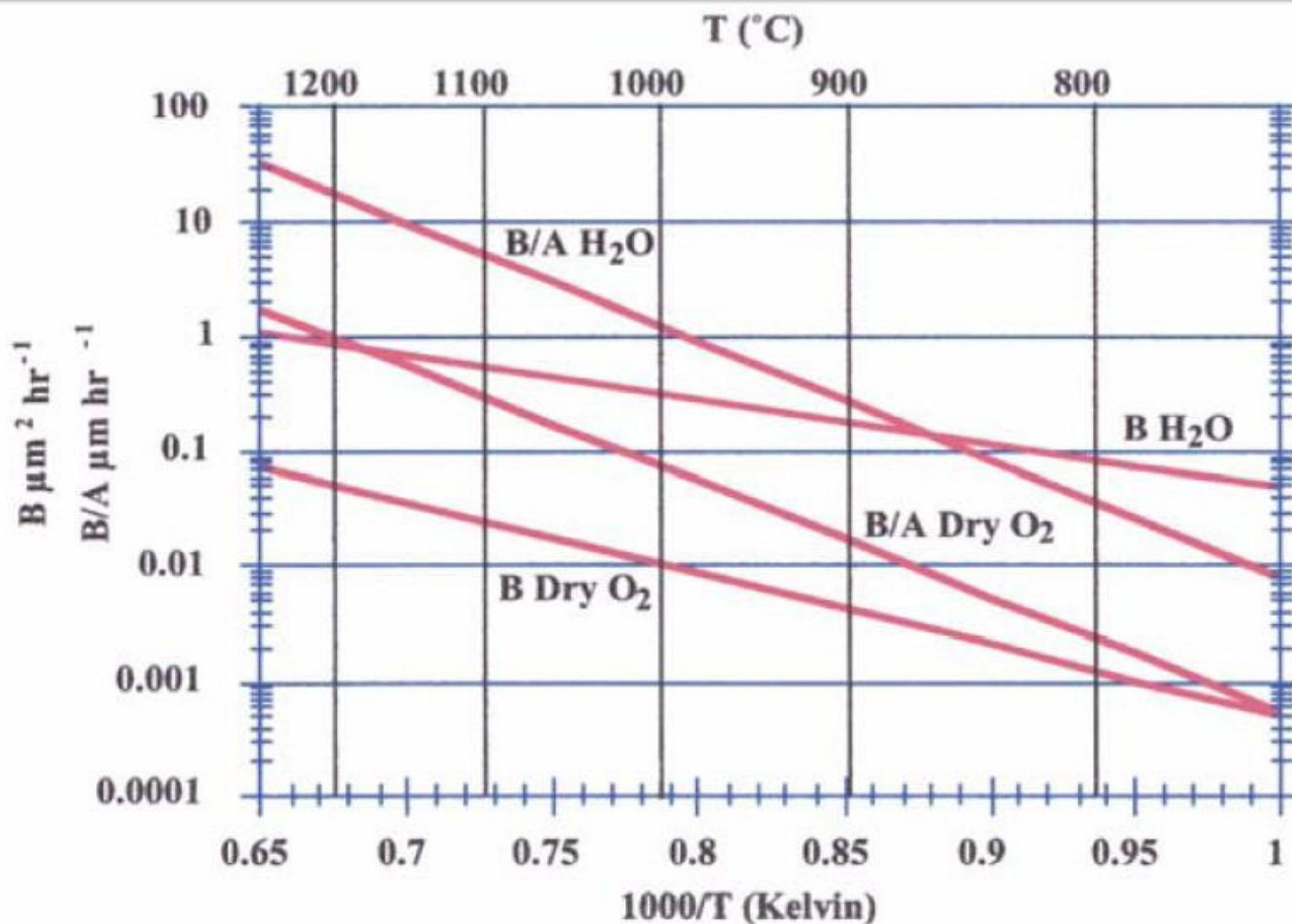
B=2DC*/N₁ \propto D, D depends on T exponentially.

B/A \approx C*/N₁ \propto k_s. From table below, E₂ is independent of whether O₂ or H₂O oxidation, as well as crystalline direction – E₂ represents a fundamental process only related to the substrate (e.g. Si-Si bond breaking).

Table 6.2 Rate constants describing (111) silicon oxidation kinetics at 1 atm total pressure. For the corresponding values for (100) silicon, all C₂ values should be divided by 1.68.

Ambient	B	B/A
Dry O ₂	$C_1 = 7.72 \times 10^2 \mu^2 \text{ hr}^{-1}$ $E_1 = 1.23 \text{ eV}$	$C_2 = 6.23 \times 10^6 \mu \text{ hr}^{-1}$ $E_2 = 2.0 \text{ eV}$
Mostly H ₂ O, but some O ₂ , no longer popular Wet O ₂	$C_1 = 2.14 \times 10^2 \mu^2 \text{ hr}^{-1}$ $E_1 = 0.71 \text{ eV}$	$C_2 = 8.95 \times 10^7 \mu \text{ hr}^{-1}$ $E_2 = 2.05 \text{ eV}$
H ₂ +O ₂ reacts to form H ₂ O H ₂ O	H ₂ O only, more popular now $C_1 = 3.86 \times 10^2 \mu^2 \text{ hr}^{-1}$ $E_1 = 0.78 \text{ eV}$	$C_2 = 1.63 \times 10^8 \mu \text{ hr}^{-1}$ $E_2 = 2.05 \text{ eV}$

Arrhenius expression for B and B/A



B and B/A for O_2 and H_2O oxidation of (111) Si. Values taken from above table.

Calculated dry oxidation rate from D-G model

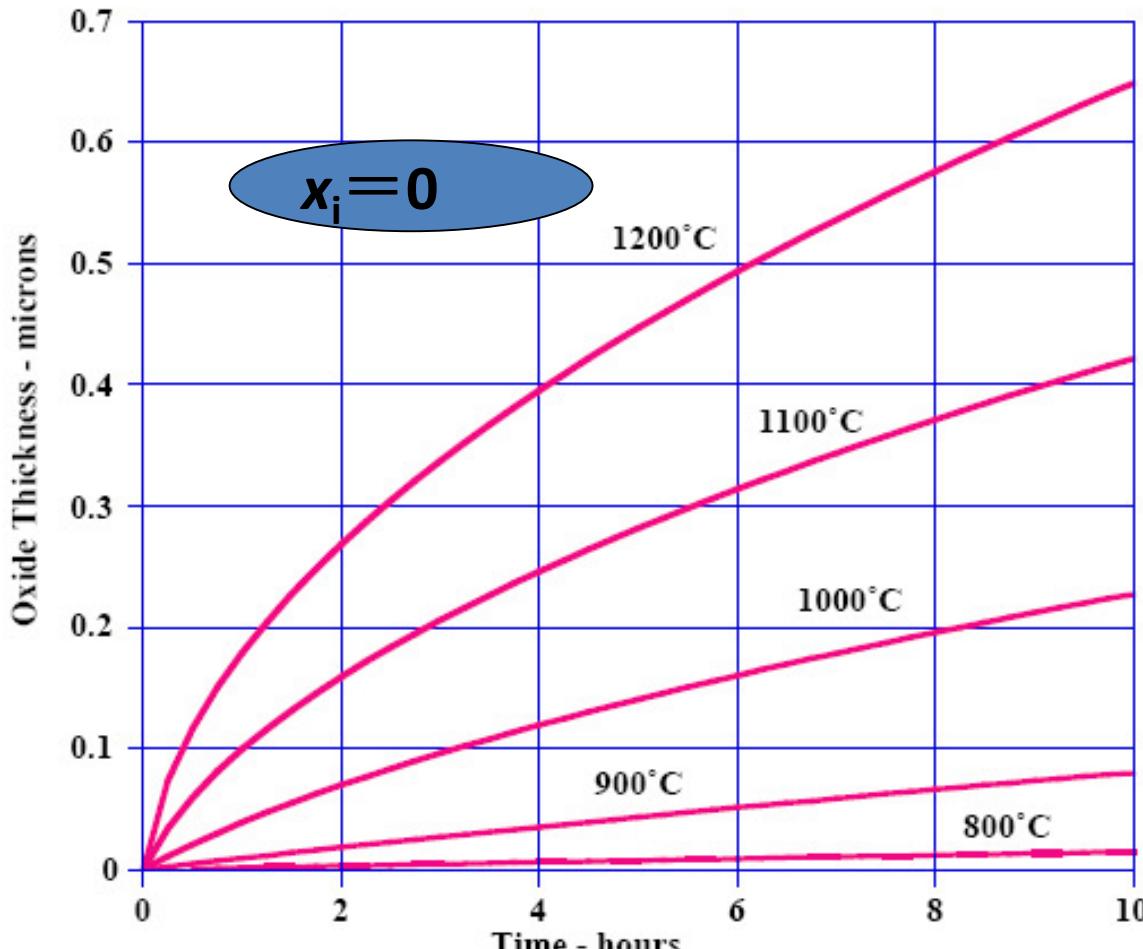


Figure 6-19

(100) silicon. The initial fast oxidation for the first ~20nm is not included (i.e. $x_i, \tau=0$).

Calculated wet (H_2O) oxidation rate from D-G model

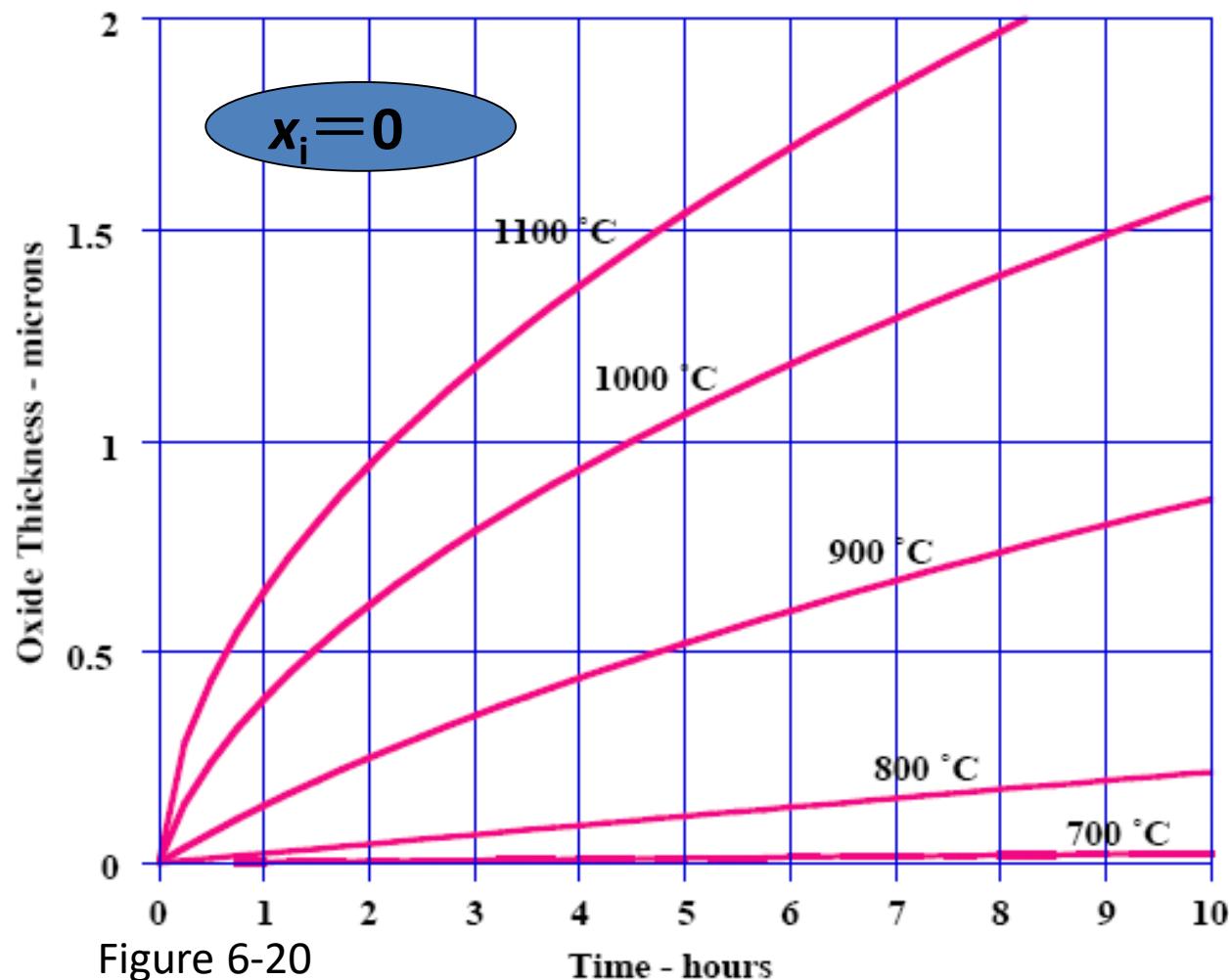


Figure 6-20

Dry vs. wet oxidation

Wet oxidation: a mixture of O₂ and H₂O (O₂ bubble through hot water) used as oxidant.

Advantage: higher growth rate than dry oxidation.

Reason for higher rate: much larger solubility in oxide (Henry's constant H) for H₂O compared with O₂. C*(bulk concentration, =HP_G) for H₂O: 3x10¹⁹/cm³; and for O₂: 5x10¹⁶/cm³.

Disadvantage: oxides grown wet are less dense, with a more open structure, because out-diffusion of H₂ creates 'voids' along its path.

Thus wet oxidation is typically used when a thick oxide is required that will not be subjected to any significant electrical stress that may lead to electrical breakdown.

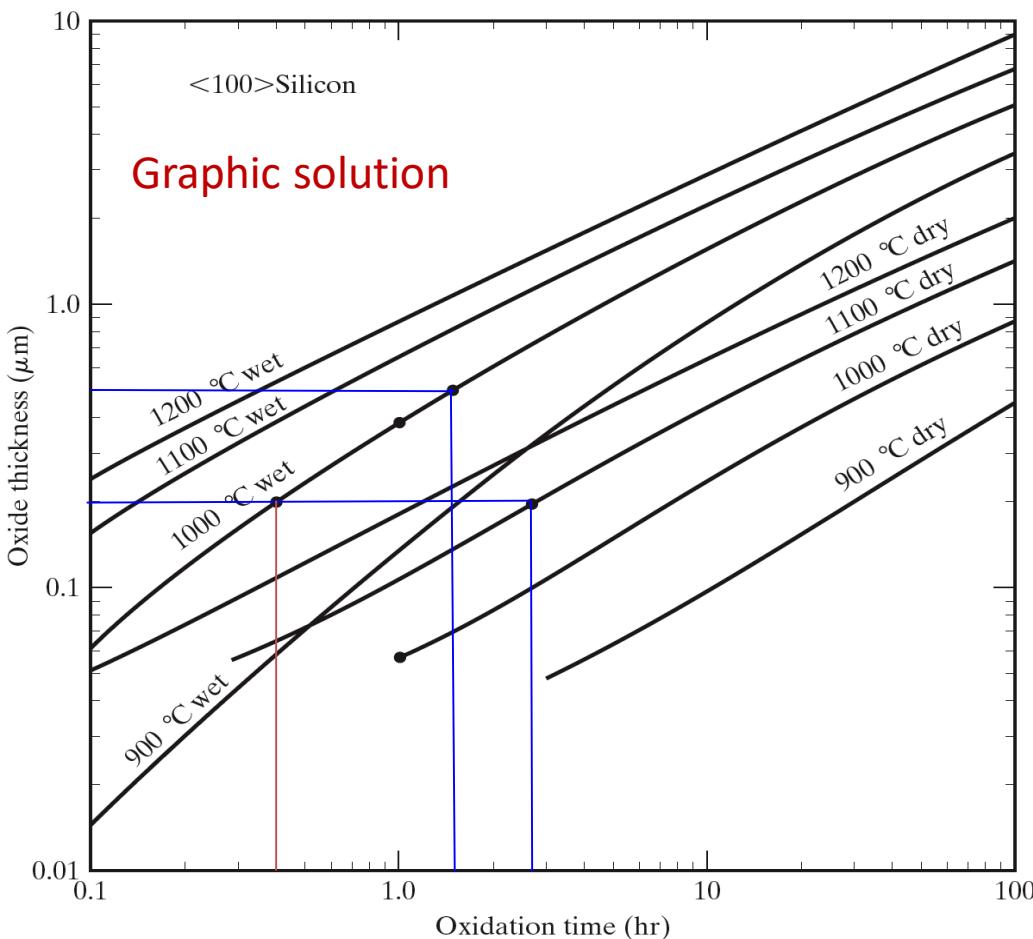
Dry oxidation: slow, higher quality than wet oxidation, used for gate oxide.

Note that dry oxidation 'appears' to always have some initial oxide present (i.e. x_i≠0).

Thermal oxidation example

A <100> silicon wafer has a 2000-Å oxide on its surface

- (a) How long did it take to grow this oxide at 1100° C in dry oxygen?
- (b) The wafer is put back in the furnace in wet oxygen at 1000° C. How long will it take to grow an additional 3000 Å of oxide?



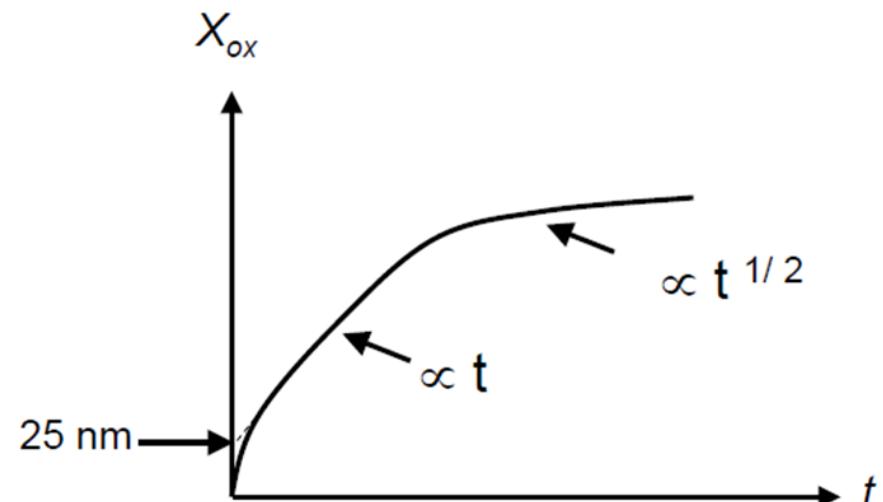
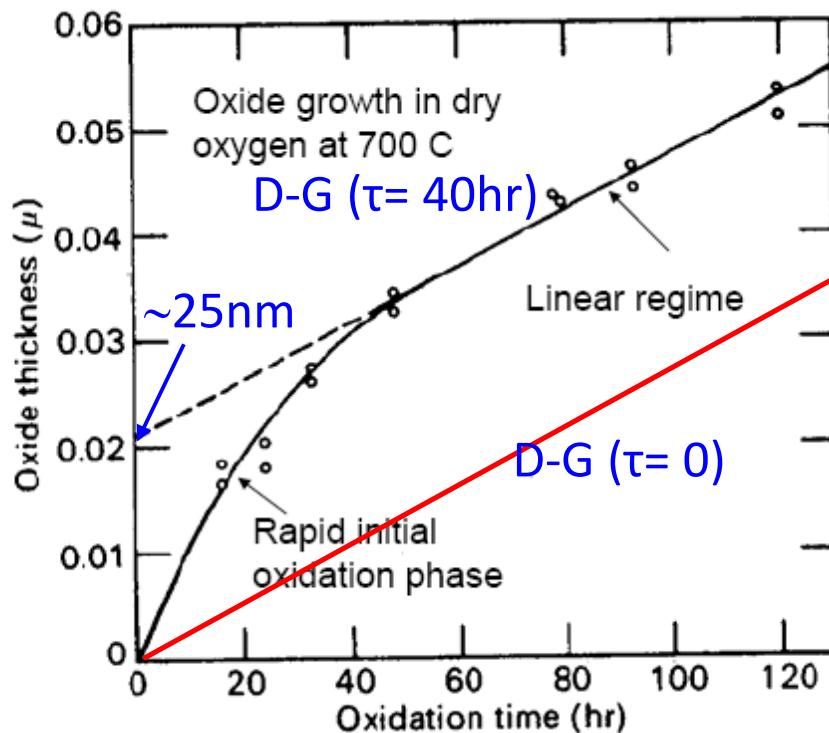
- a) According to the figure, it would take 2.8hr to grow 0.2 μm oxide in dry oxygen at 1100° C.
- b) The total oxide thickness at the end of the oxidation would be 0.5 μm which would require 1.5hr to grow if there was no oxide on the surface to begin with. However, the wafer “thinks” it has already been in the furnace 0.4hr (to grow the first 200nm oxide). Thus the additional time needed to grow the 0.3 mm oxide is 1.5-0.4 = 1.1 hr.

Chapter 6 Thermal oxidation and the Si/SiO₂ interface

1. SiO₂ properties and applications.
2. Manufacturing methods and equipment.
3. Measurement methods (mechanical, optical, electrical).
4. Deal-Grove model (linear parabolic model).
5. Thin oxide growth, dependence on gas pressure and crystal orientation.
6. Cl-containing gas, 2D growth, substrate doping effect .
7. Interface charges, dopant redistribution, rapid thermal oxidation.

Thin oxide growth

- The Deal-Grove model provides excellent agreement with experimental data except for thin (<20nm) SiO₂ grown in dry O₂.
- When using the D-G equations for thick oxides grown in O₂ on bare Si (X_i=0), one needs to “assume” X_i=25nm.
- Or equivalently, one can use τ ($\neq 0$) to correct the Deal-Grove model to compensate for the excess growth that occurs in the initial growth regime.

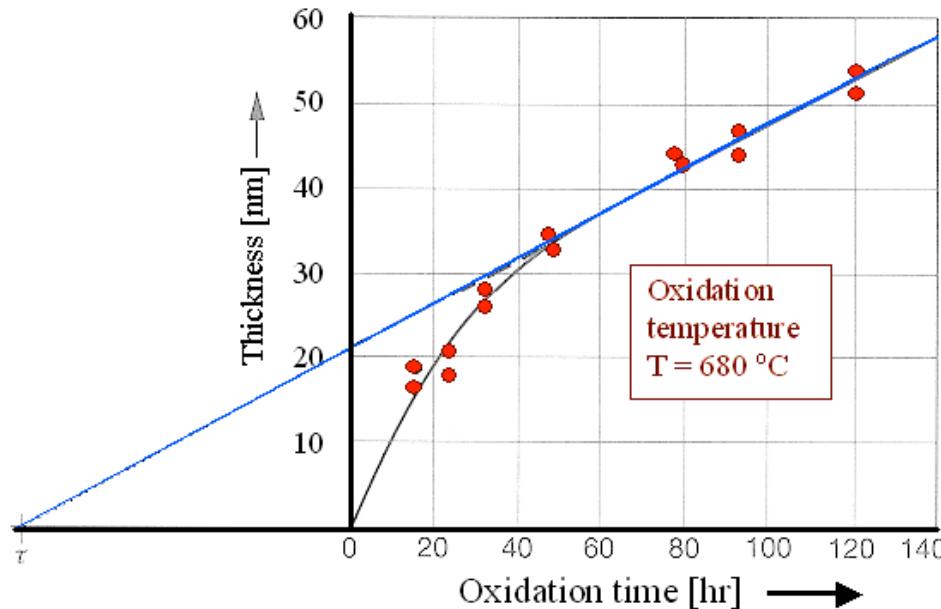


Mechanism for initial oxidation regime

Models to explain the excess growth of thin oxidation

- An electric field exists across oxide that enhances diffusion (if diffusing species is O_2^- , *but it is not.*) during early states of oxidation.
- Thin micro-channels in oxide aid in the movement of O_2 to the Si surface.
- Mismatch in thermal expansion coefficients of oxide and Si causes stress in oxide and this stress may enhance the diffusivity of the oxidizing species.

But NONE of these mechanism gained wide acceptance. The exact mechanism is still unknown.



Model of Massoud et. al.

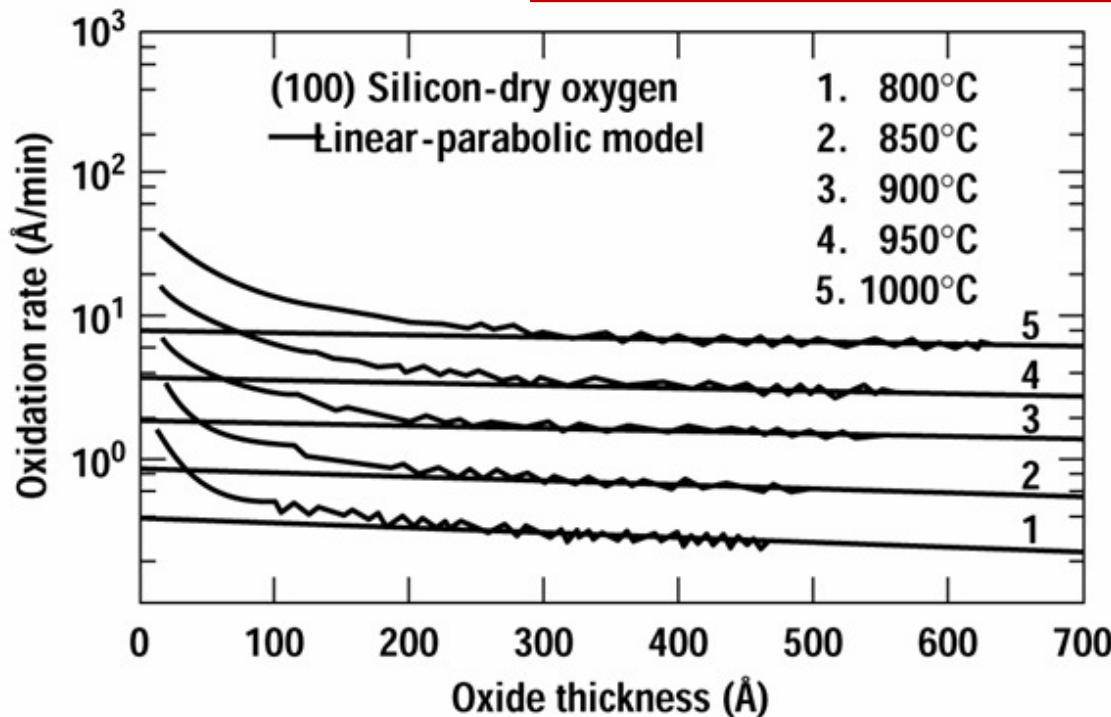
For D-G model:

$$\frac{x_o^2 - x_i^2}{B} + \frac{x_o - x_i}{B/A} = t, \text{ therefore}$$

$$\frac{2x_o}{B} \frac{dx_o}{dt} + \frac{A}{B} \frac{dx_o}{dt} = 1, \text{ thus } \frac{dx_o}{dt} = \frac{B}{2x_o + A}$$

Massoud model:

$$\frac{dx_O}{dt} = \frac{B}{2x_O + A} + C \exp\left(-\frac{x_O}{L}\right)$$



$$C = C^0 \exp(-E_A/kT)$$

$$C^0 \approx 3.6 \times 10^8 \mu\text{m/hr}$$

$$E_A \approx 2.35 \text{ eV}$$

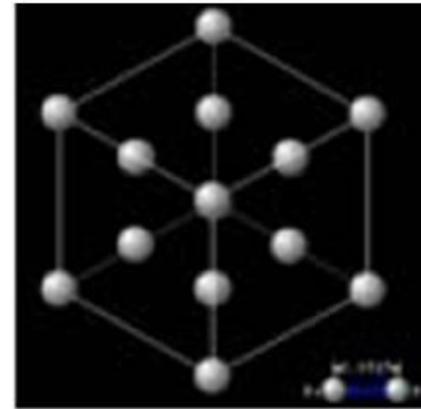
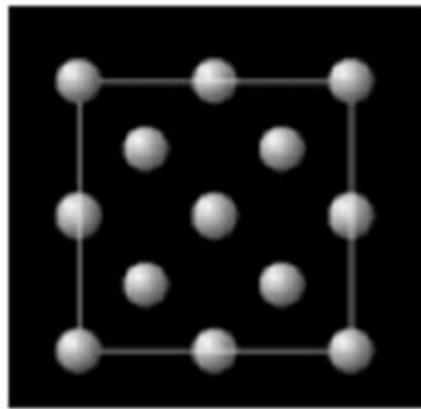
$$L \approx 7 \text{ nm}$$

For $x_o \gg 7 \text{ nm}$, Massoud model approaches DG model.

Dependence on crystal orientation

Oxidation rate depends on the availability of reaction sites on the silicon substrate.

Oxidation on the $<111>$ crystal plane occurs at a higher rate because there are a higher number of surface atoms/chemical bonds than the $<100>$ plane.



Orientation	Area of unit cell (cm ²)	Si atoms in area	Si bonds in area	Bonds available	Available bonds, N (10 ¹⁴ cm ⁻²)	N relative to $<110>$
$<110>$	$\sqrt{2}a^2$	4	8	4	9.59	1.000
$<111>$	$1/2\sqrt{3}a^2$	2	4	3	11.76	1.227
$<100>$	a^2	2	4	2	6.77	0.707

Most IC made of $<100>$ silicon.

Dependence on crystal orientation

$$k_s = k_{s0} \exp(-E_a / kT)$$

k_{s0} is a constant, roughly proportional to the number of available Si bonds for reaction per unit area.

- Interface reaction rate constant K_s (cm/sec) depends on crystal orientation.
- So the liner grown rate B/A depends on crystal orientation. $(B/A)_{111} = 1.68 (B/A)_{100}$ ($1.68 \neq 1.227/0.707 = 1.735$, see previous slide)
- The parabolic rate constant B is NOT dependent on crystal orientation.
- So this effect decreases for high temperature and/or long time oxidation when oxide become thick.

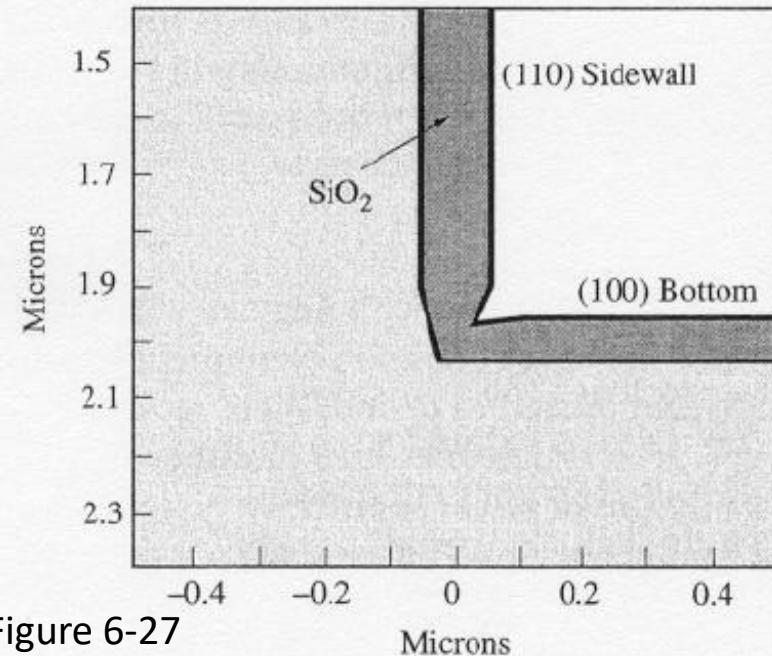
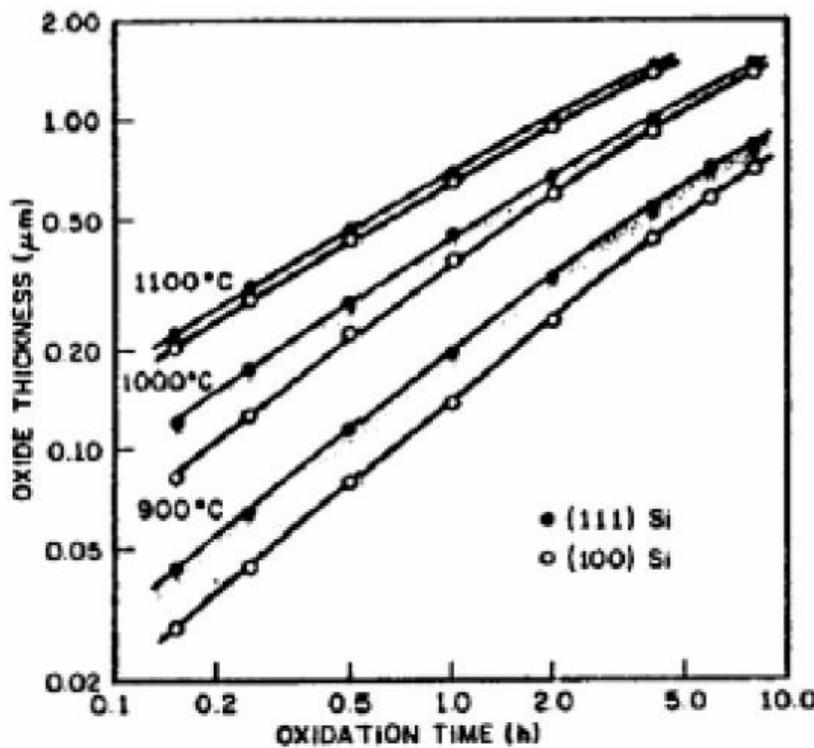


Figure 6-27

Simulation, (100) Si, in H_2O at 900°C for 30 min

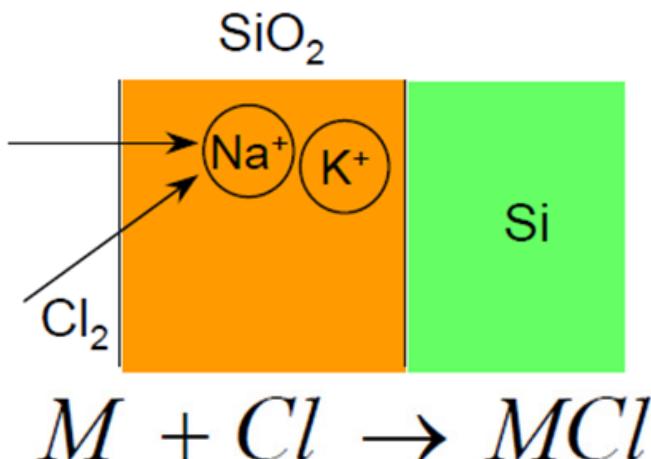
Chapter 6 Thermal oxidation and the Si/SiO₂ interface

1. SiO₂ properties and applications.
2. Thermal oxidation basics.
3. Manufacturing methods and equipment.
4. Measurement methods (mechanical, optical, electrical).
5. Deal-Grove model (linear parabolic model).
6. Thin oxide growth, dependence on gas pressure and crystal orientation.
7. Cl-containing gas, 2D growth, substrate doping effect .
8. Interface charges, dopant redistribution, rapid thermal oxidation.

Oxidation using Cl-containing gas mixture

Addition of Cl (1-3%) bearing species to oxidation ambient leads to:

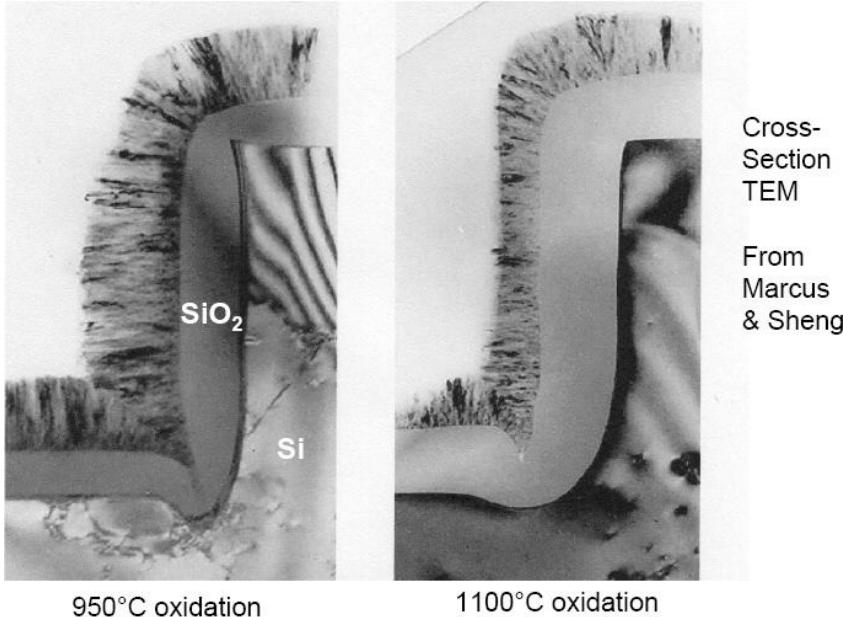
- Faster oxide growth. Both B/A and B increases, because bond energy of Si-O is 4.25eV, Si-Cl is 0.5eV, so Cl_2 reacts with Si first to form Si-Cl, which then reacts with O_2 to form SiO_2 . Here Cl_2 is a kind of catalyst.
- Cleaner oxide, less metallic contamination, since Cl is a metal getter. Most heavy metal atoms react with Cl to form *volatile* metal chloride. Metallic contaminants originate from heating elements and insulation around the fused silica flow tube in which the oxidation is done.
- Enhanced dielectric strength, reduced oxide density, improved Si/ SiO_2 interface with lower interface state density, thus improved device performance.



Chlorine species:

- Anhydrous chloride (Cl_2)
- Anhydrous hydrogen chloride (HCl)
- Trichloroethylene – TCE (C_2HCl_3)
- Trichloroethane – TCA ($\text{C}_2\text{H}_3\text{Cl}_3$)

Two dimensional oxidation



Experiment shows that:

- Oxidation is slower for convex or concave corners.
- Concave corner is even slower than convex corner.
- The smaller the curvature radius is, the slower.
- More serious for low temperature oxidation, no effect for high temperature 1200°C (when oxide can “flow”).

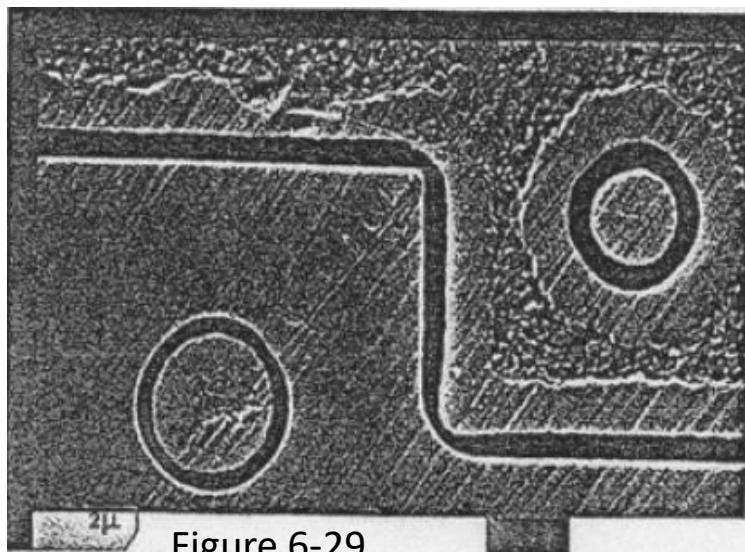
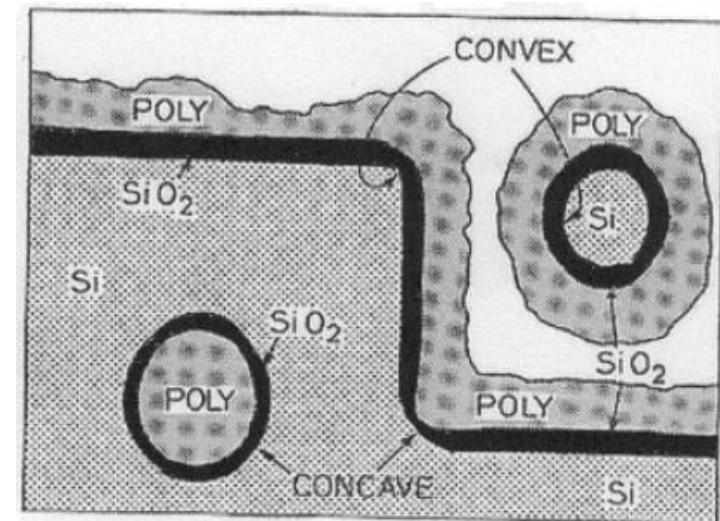


Figure 6-29

Experiment

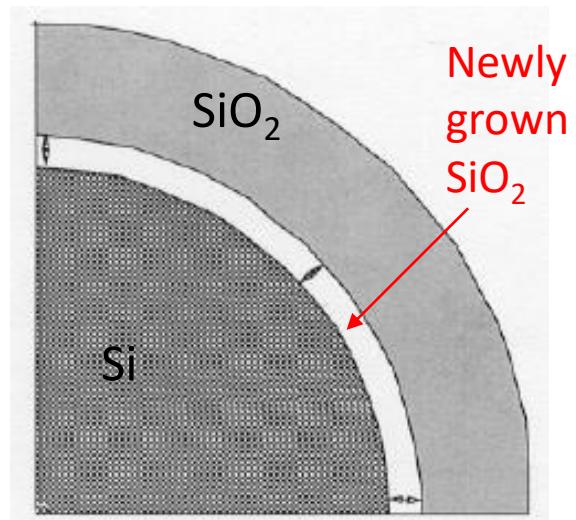
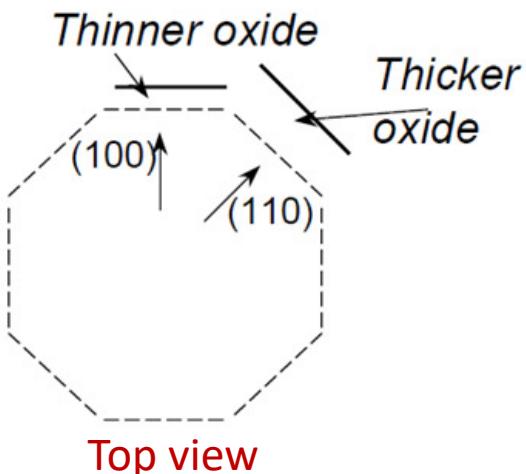
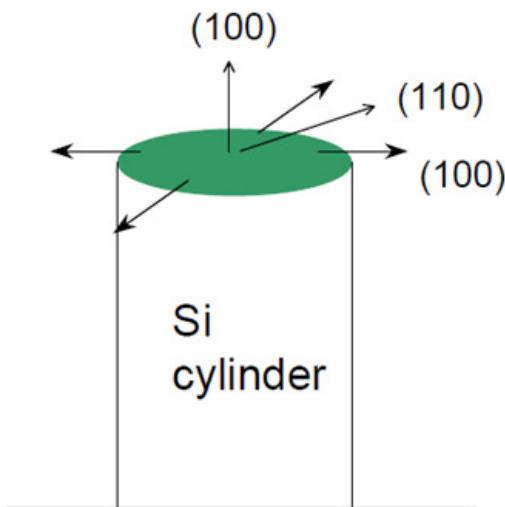


Drawing to show the structure

Two dimensional oxidation mechanism

Several physical mechanisms are important in explaining these results.

- Crystal orientation dependant rate.
- 2D oxidant diffusion (different from 1D).
- Stress due to volume expansion (we know when Si oxidizes, it expands). As the oxide grows, the “newly” formed oxide pushes out the “old” oxide which rearranges itself through viscous flow. Stress occurs typically on curved surfaces.



Dopant dependence

Common Si dopants all tend to enhance oxidation rate of Si when present in the substrate in high concentrations.

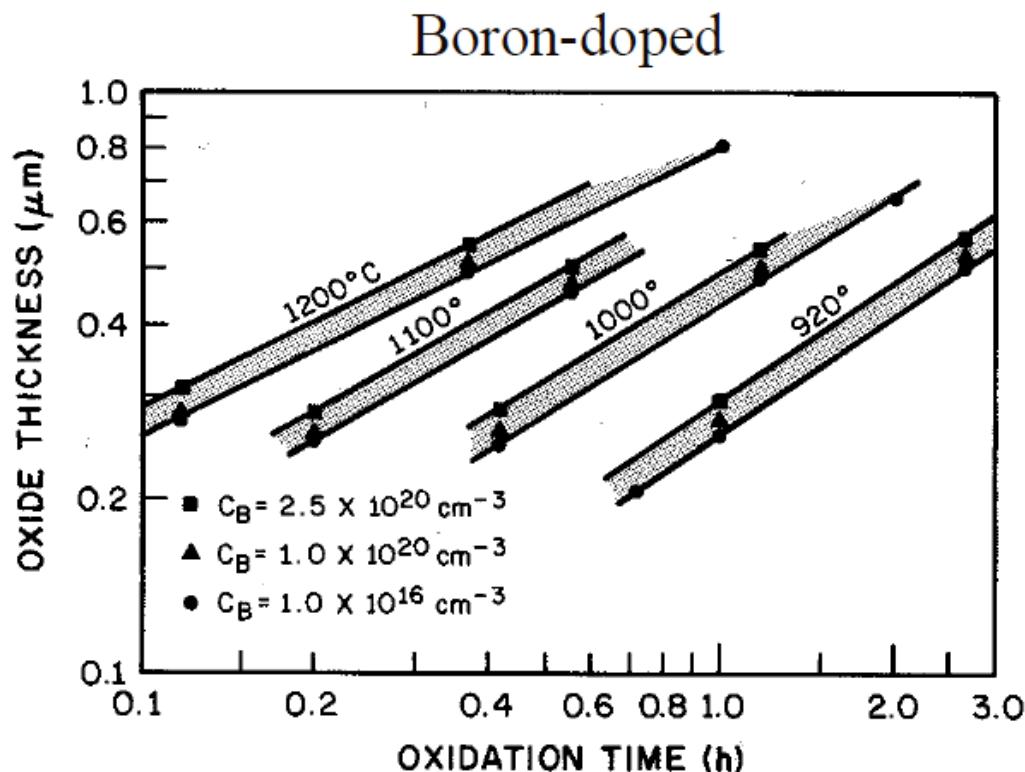
The effect is particularly important at lower temperatures and for thinner oxides, and is more important for N⁺ doping than P⁺ doping.

The oxidation rate depends on:

The dopant concentration in SiO₂ for diffusion controlled oxidation (B dominates).

The dopant concentration at Si surface for reaction controlled oxidation (B/A dominates).

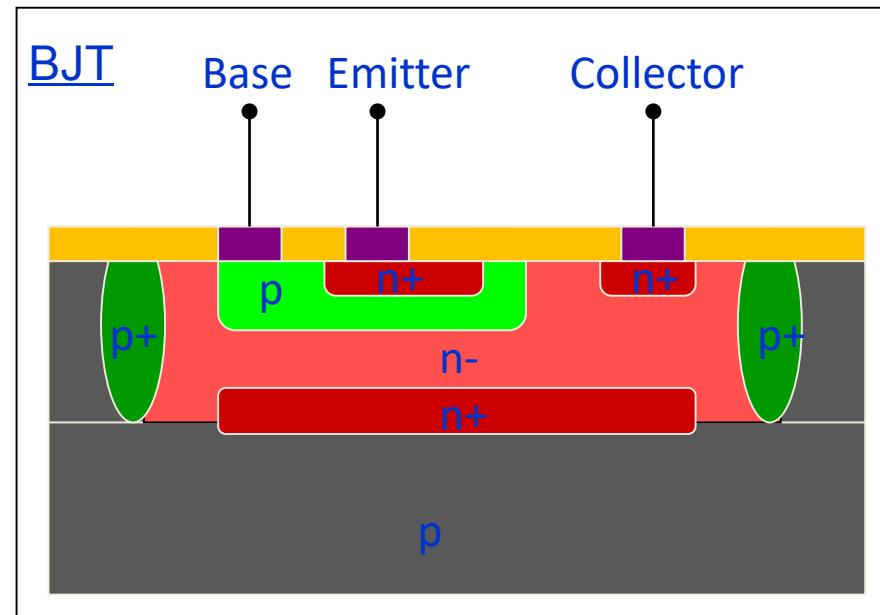
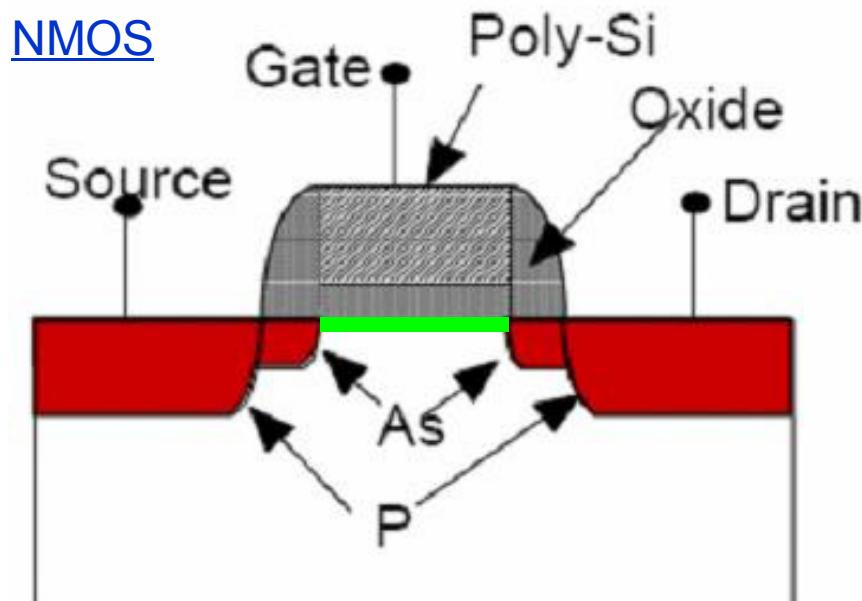
B segregates into oxide,
weakens SiO₂ bond
structure, increases oxidant
diffusivity D.



Chapter 7 Dopant Diffusion

1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

Doping in MOS and bipolar junction transistors



Doping is realized by:

- Diffusion from a gas, liquid or solid source, on or above surface.
- Ion implantation. (choice for today's IC).

In this chapter, diffusion means two very different concepts: one is to dope the substrate from source on or above surface – the purpose is doping; one is diffusion inside the substrate – the purpose is to re-distribute the dopant.

Doping profile for a p-n junction

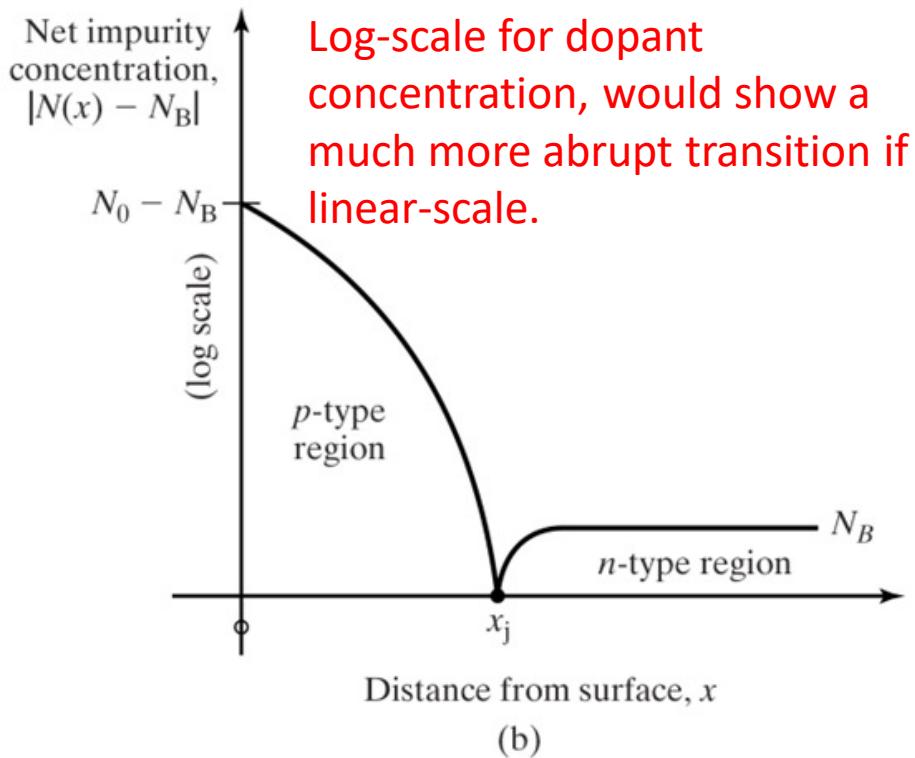
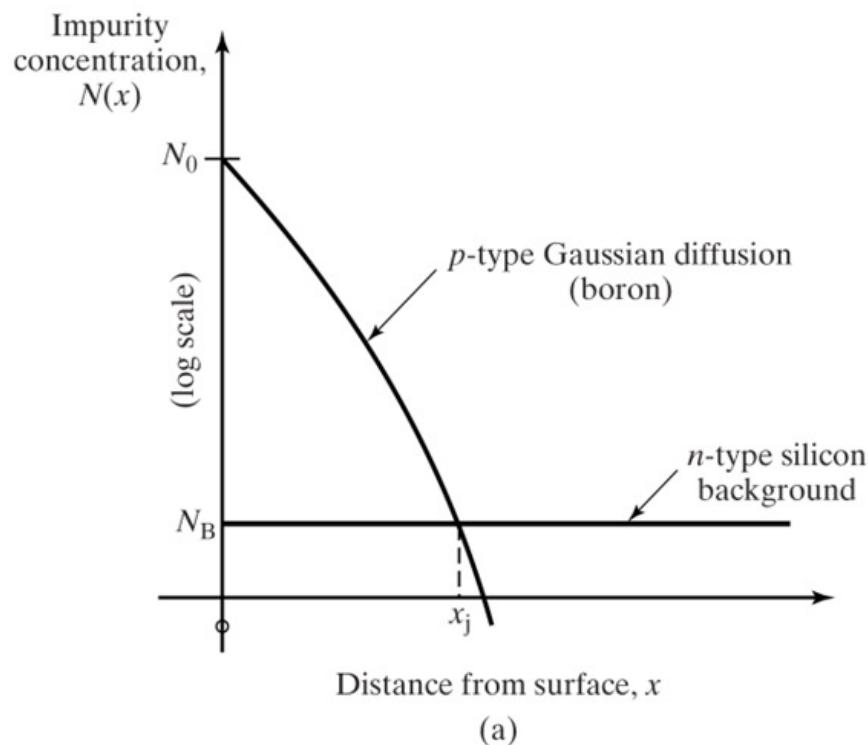
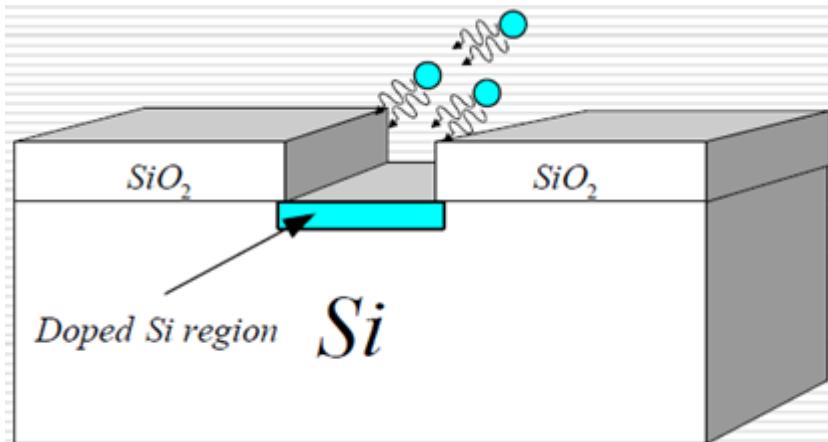


FIGURE 4.7

Formation of a *pn* junction by diffusion. (a) An example of a *p*-type Gaussian diffusion into a uniformly doped *n*-type wafer; (b) net impurity concentration in the wafer. The metallurgical junction occurs at the point $x = x_j$ where the net concentration is zero. The material is converted to *p*-type to the left of x_j and remains *n*-type to the right of x_j .

Log-scale for dopant concentration, would show a much more abrupt transition if linear-scale.

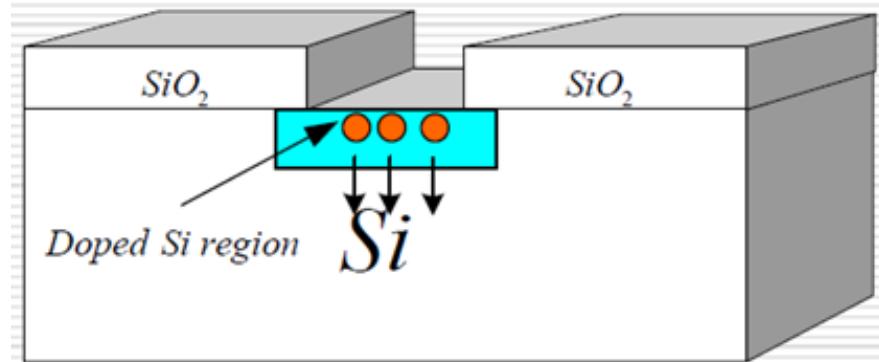
Diffusion from gas, liquid or solid source



Pre-deposition (dose control)



- Silicon dioxide is used as a mask.
- The mixture of dopant species, oxygen and inert gas like nitrogen, is passed over the wafers at order of $1000^{\circ}C$ ($900^{\circ}C$ to $1100^{\circ}C$) in the diffusion furnace.
- The dopant concentration in the gas stream is sufficient to reach the solid solubility limit for the dopant species in silicon at that temperature.
- The impurities can be introduced into the carrier gas from solid (evaporate), liquid (vapor) or gas source.



Drive-in (profile control)

Comparison of ion implantation with solid/gas phase diffusion

Table 7-2 Comparison of ion implantation versus solid- or gas-phase doping methods

Advantages

Ion Implantation and Annealing

Room temperature mask

Precise dose control

$10^{11} - 10^{16}$ atoms cm^{-2} doses

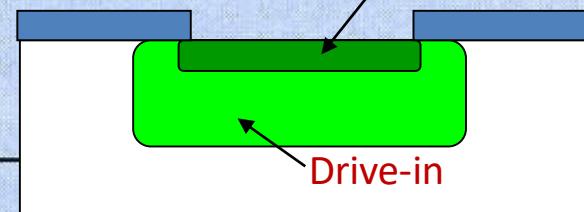
Accurate depth control

Solid-/Gas-Phase Diffusion

No damage created by doping

Batch fabrication

Pre-deposition



Disadvantages

Ion Implantation and Annealing

Implant damage enhances diffusion

Dislocations caused by damage may cause junction leakage

Implant channeling may affect profile

Solid-/Gas-Phase Diffusion

Usually limited to solid solubility

Low surface concentration hard to achieve without a long drive-in

Low dose predeps very difficult

Chapter 7 Dopant Diffusion

1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

Dopant solid solubility

Solid solubility: at equilibrium, the maximum concentration for an impurity before precipitation to form a separate phase.

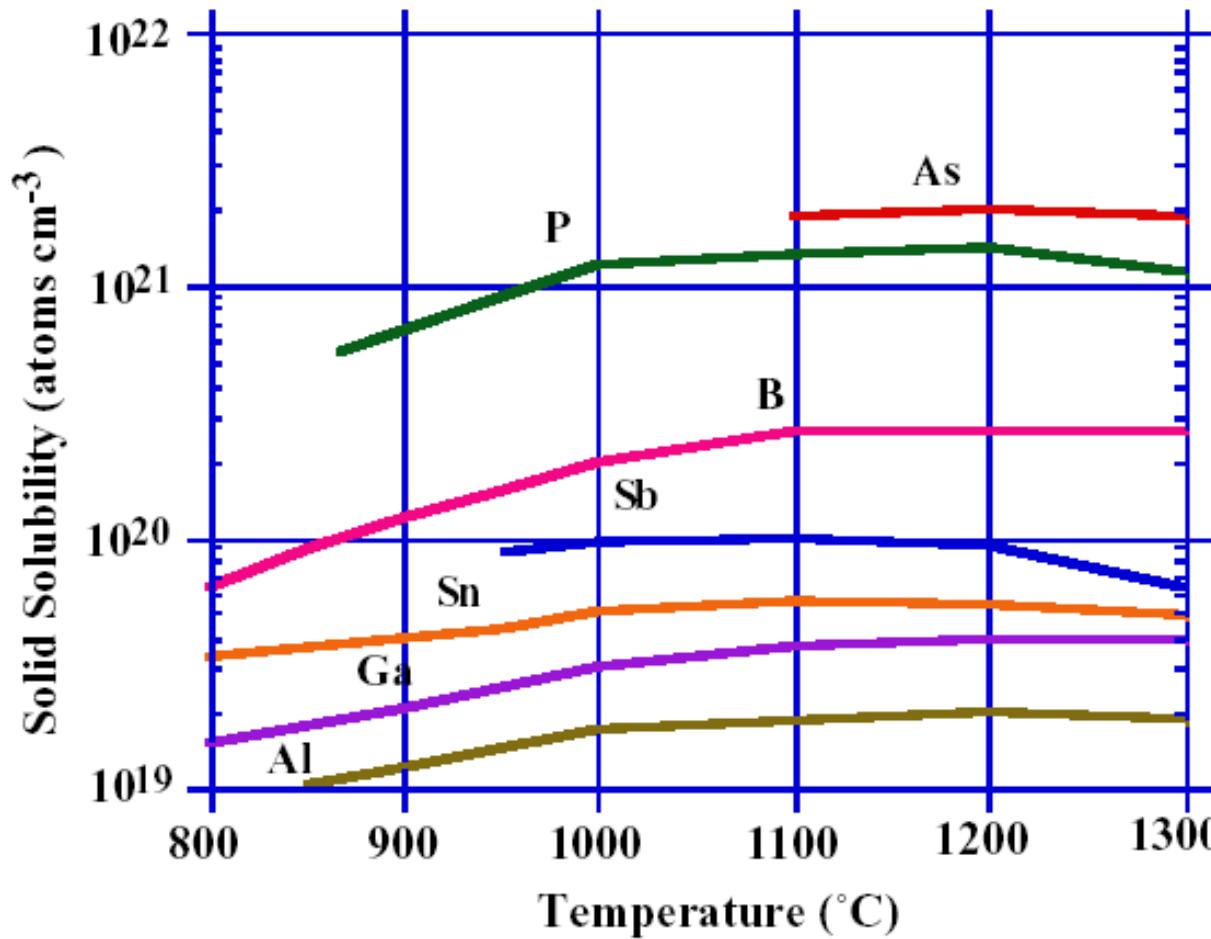
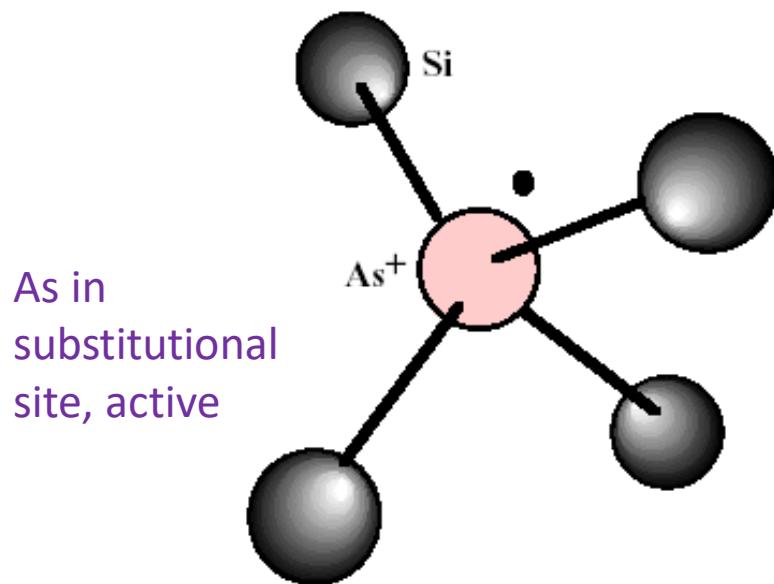


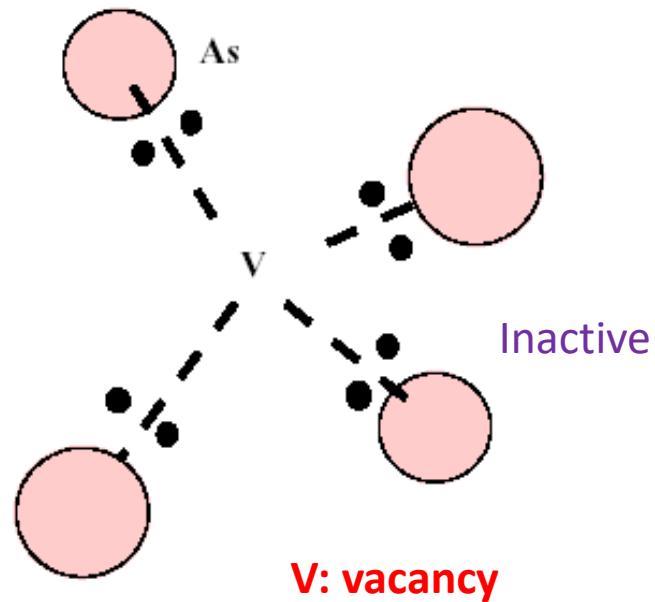
Figure 7-4

Solubility vs. electrically active dopant concentration



As in
substitutional
site, active

Figure 7-5



V: vacancy

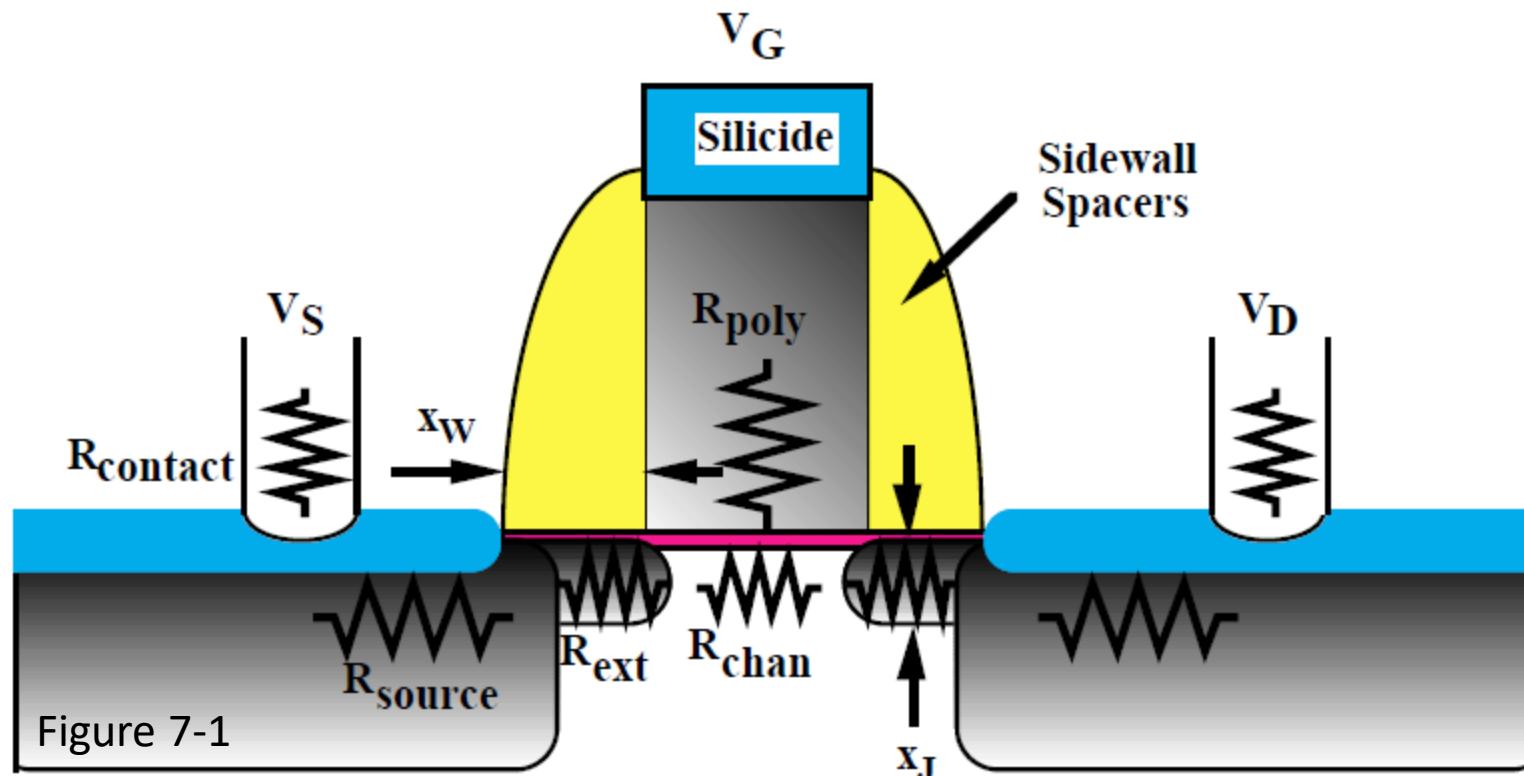
- As_4V is one possible electrically inactive form.

Not all impurities are electrically active.

As has solid solubility of $2 \times 10^{21} \text{ cm}^{-3}$.

But its maximum electrically active dopant concentration is only $2 \times 10^{20} \text{ cm}^{-3}$.

Resistance in a MOS



For thin doping layers, it is convenient to find the resistance from **sheet resistance**.

Sheet resistance R_s

$$R_s = \frac{\rho}{x_j}$$

ρ : (bulk) resistivity

x_j : junction depth, or film thickness...

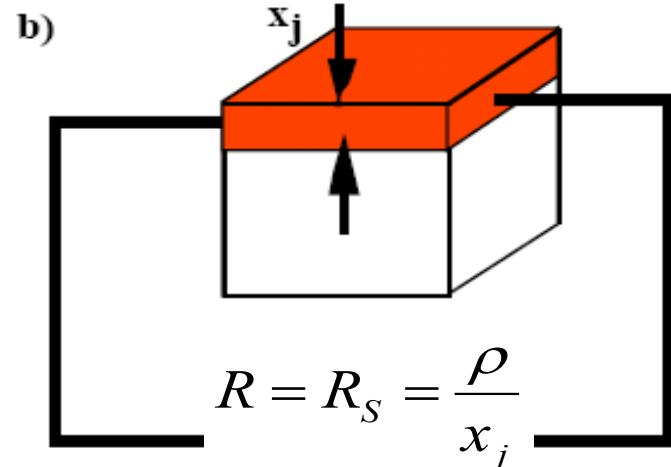
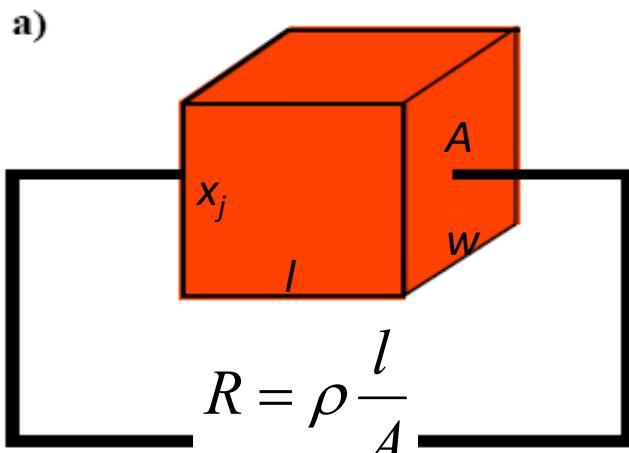


Figure 7-2

$$R = \rho \frac{l}{A} = \rho \frac{l}{wx_j} = \frac{\rho}{x_j} \frac{l}{w} = R_s \frac{l}{w}$$

$R=R_s$ when $l=w$ (square)

Important formulas

Ohm's law: $\vec{J} = \sigma \vec{E}$ $\vec{E} = \rho \vec{J}$

Mobility μ : $\vec{v} = \mu \vec{E}$

By definition: $\vec{J} = q(p\vec{v}_h - n\vec{v}_n)$

Therefore: $\sigma = q \left(p \frac{\vec{v}_h}{\vec{E}} - n \frac{\vec{v}_n}{\vec{E}} \right) = q \left(p \frac{v_{hx}}{E_x} + n \frac{v_{nx}}{E_x} \right)$

Finally: $\sigma = q(n\mu_n + p\mu_p)$

Where: $\mu_n = \frac{v_{nx}}{E_x}$ $\mu_p = \frac{v_{hx}}{E_x}$

σ : conductivity; ρ : resistivity; J : current density; E : electrical field
 v : velocity; q : charge; n, p : carrier concentration.

Sheet resistance

$$R_S = \frac{\rho}{x_j} = \frac{1}{\sigma x_j} = \frac{1}{q\mu N x_j} = \frac{1}{q\mu Q}$$

N is carrier density, Q is total carrier per unit area, x_j is junction depth

For non-uniform doping:

$$R_S = \frac{\bar{\rho}}{x_j} = \frac{1}{\bar{\sigma} x_j} = \frac{1}{q \int_0^{x_j} [n(x) - N_B] \mu[n(x)] dx}$$

This relation is calculated to generate the so-called Irvin's curves.

Chapter 7 Dopant Diffusion

1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

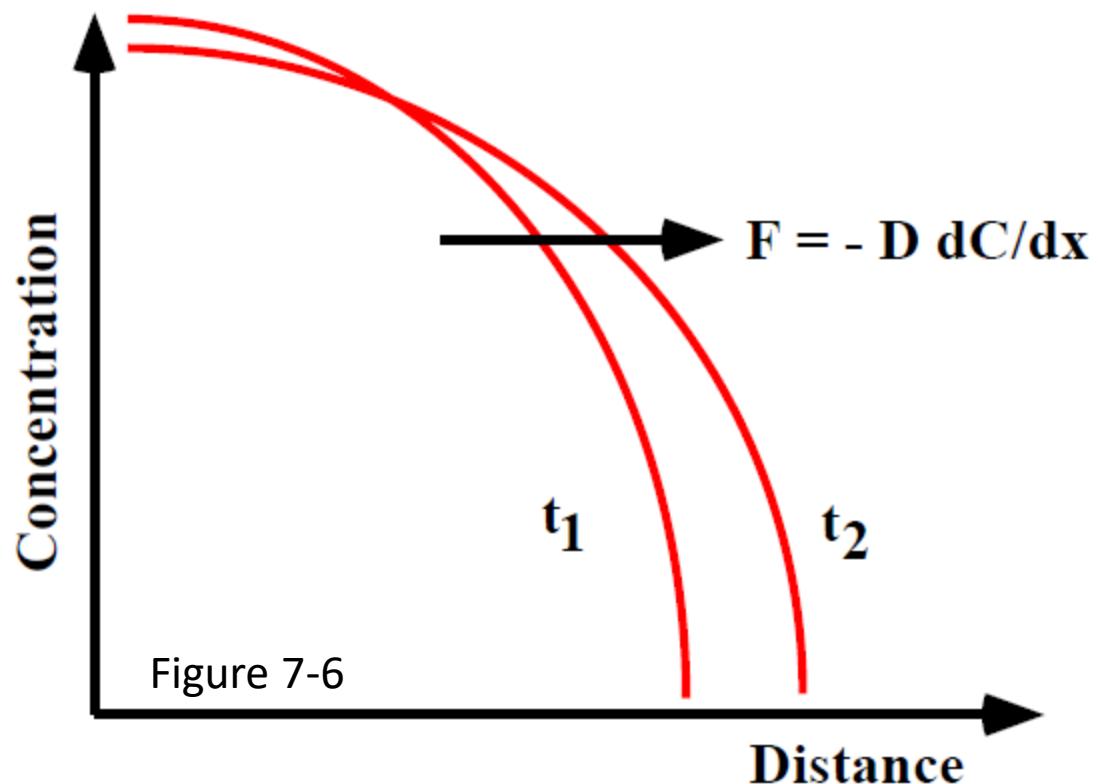
Diffusion from a macroscopic viewpoint

Fick's first law of diffusion

F is net flux.

$$F(x, t) = -D \frac{\partial C(x, t)}{\partial x}$$

This is similar to other laws where cause is proportional to effect (Fourier's law of heat flow, Ohm's law for current flow).



C is impurity concentration (number/cm³), D is diffusivity (cm²/sec).

D is related to atomic hops over an energy barrier (formation and migration of mobile species) and is exponentially (see next slide) activated.

Negative sign indicates that the flow is down the concentration gradient.

Intrinsic diffusivity D_i

Intrinsic: impurity concentration $N_A, N_D < n_i$ (intrinsic carrier density).

Note that n_i is quite high at typical diffusion temperatures, so "intrinsic" actually applies under many conditions. E.g. at 1000°C, $n_i = 7.14 \times 10^{18}/\text{cm}^3$.

$$D_i = D^0 \exp\left(-\frac{E_a}{kT}\right)$$

E_a : activation energy

	$D^0(\text{cm}^2/\text{s})$	$E_a(\text{eV})$
B	1.0	3.46
In	1.2	3.50
P	4.70	3.68
As	9.17	3.99
Sb	4.58	3.88

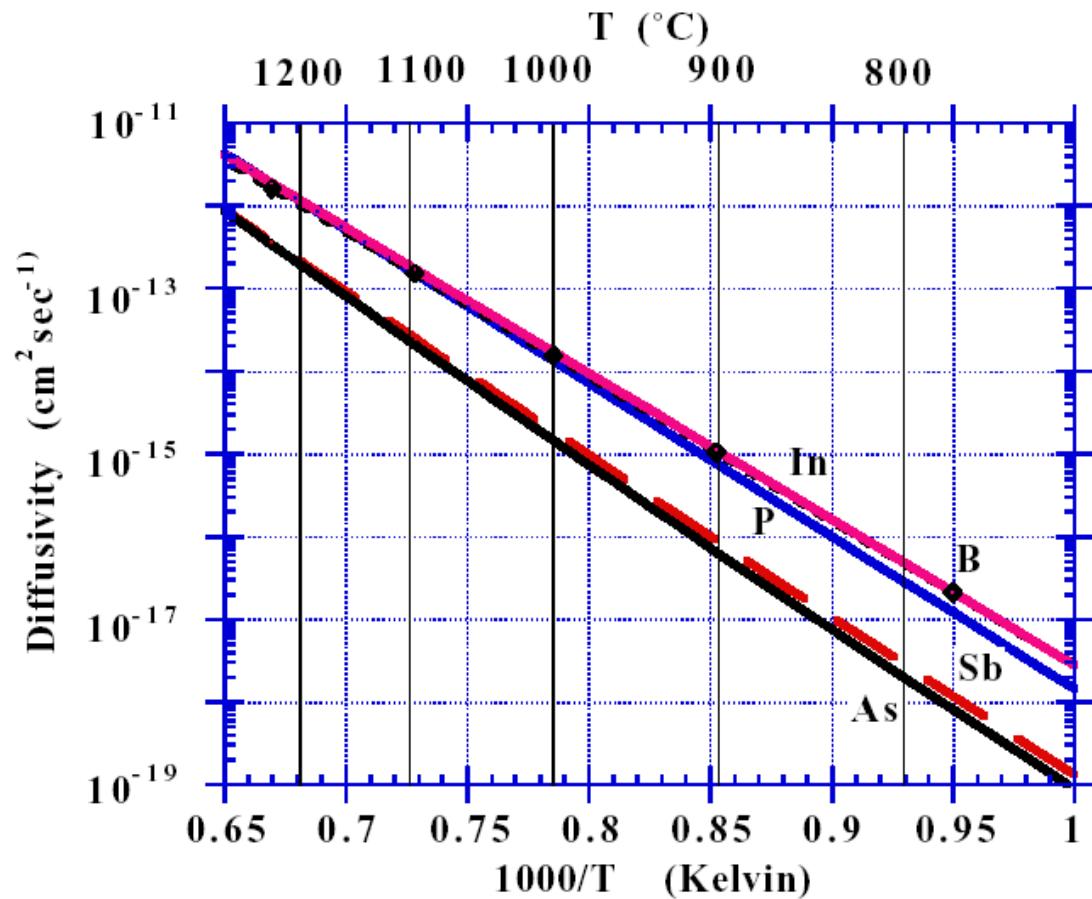


Figure 7-15, page 387

Fick's second law

The change in concentration in a volume element is determined by the change in fluxes in and out of the volume.

Within time Δt , impurity number change by:

$$[C(x, t + \Delta t) - C(x, t)] \cdot A \cdot \Delta x$$

During the same period, impurity diffuses in and out of the volume by:

$$[F(x, t) - F(x + \Delta x, t)] \cdot A \cdot \Delta t = -[F(x + \Delta x, t) - F(x, t)] \cdot A \cdot \Delta t$$

Therefore:

$$[C(x, t + \Delta t) - C(x, t)] \cdot A \cdot \Delta x = -[F(x + \Delta x, t) - F(x, t)] \cdot A \cdot \Delta t$$

Or,

$$\frac{\partial C(x, t)}{\partial t} = -\frac{\partial F(x, t)}{\partial x}$$

Since:

$$F(x, t) = -D \frac{\partial C(x, t)}{\partial x}$$

We have:

$$\frac{\partial C(x, t)}{\partial t} = -\frac{\partial F(x, t)}{\partial x} = \frac{\partial}{\partial x} \left[D \frac{\partial C(x, t)}{\partial x} \right]$$

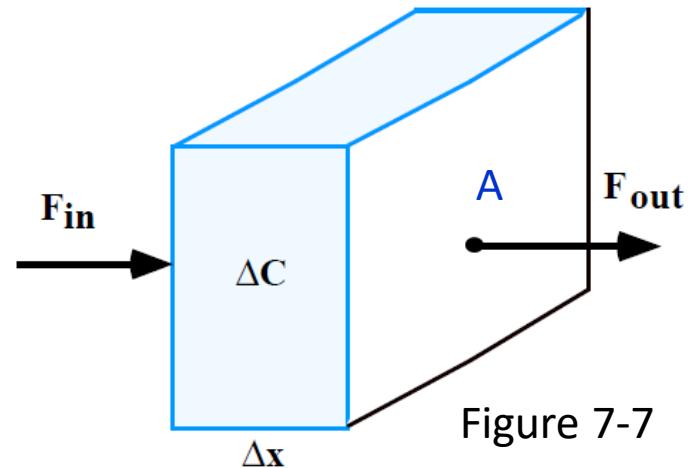


Figure 7-7

If D is constant:

$$\frac{\partial C(x, t)}{\partial t} = D \frac{\partial^2 C(x, t)}{\partial x^2}$$

Solution to diffusion equation

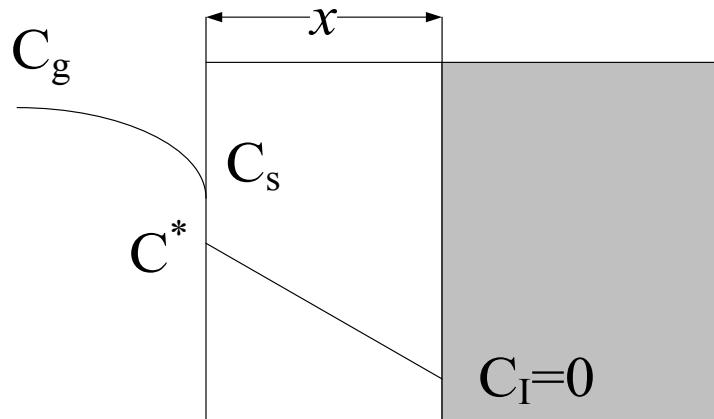
$$\frac{\partial C(x,t)}{\partial t} = D \frac{\partial^2 C(x,t)}{\partial x^2}$$

At equilibrium state, C doesn't change with time.

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} = 0$$

$$C = a + bx$$

Diffusion of oxidant (O_2 or H_2O) through SiO_2 during thermal oxidation.



SiO_2 Si

Gaussian solution in an infinite medium

$$C \rightarrow 0 \quad \text{as} \quad t \rightarrow 0 \quad \text{for} \quad x > 0$$

$$C \rightarrow \infty \quad \text{as} \quad t \rightarrow 0 \quad \text{for} \quad x = 0$$

$$\int C(x,t)dx = Q \quad (\text{limited source})$$

Figure 7-8 A delta function of dopant containing a dose Q is introduced into an infinite medium and subsequently diffused.

$$C(x,t) = \frac{Q}{2\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) = C(0,t) \exp\left(-\frac{x^2}{4Dt}\right)$$

This corresponds to, e.g. implant a very narrow peak of dopant at a particular depth, which approximates a delta function.

Important consequences:

- Dose Q remains constant
- Peak concentration (at $x=0$) decreases as $1/\sqrt{t}$
- Diffusion distance from origin increases as $2\sqrt{Dt}$

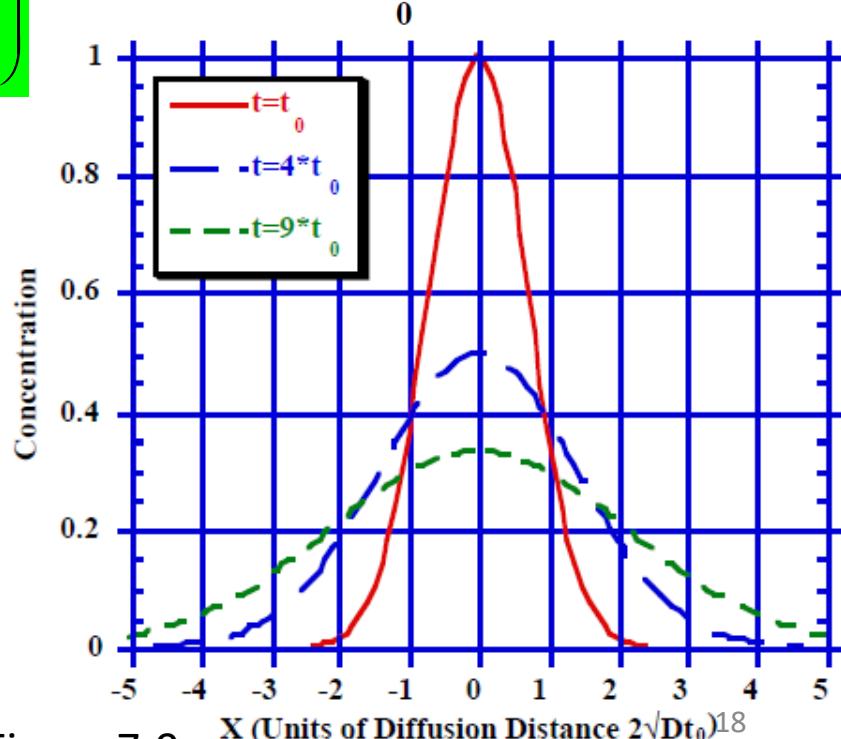
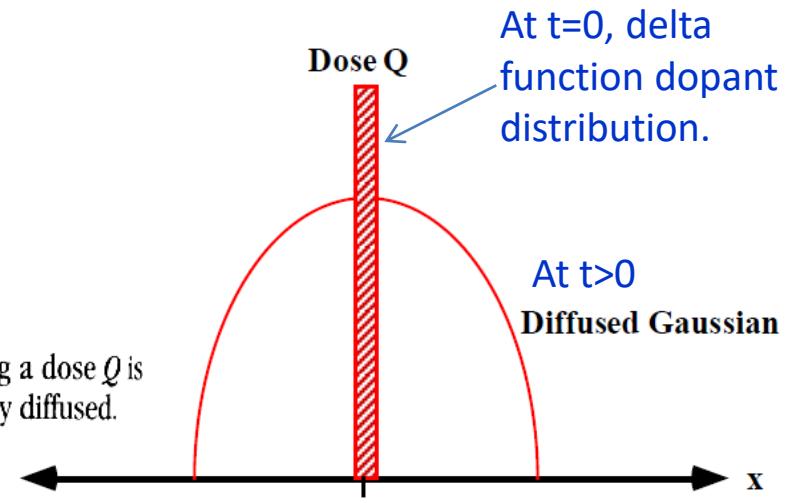
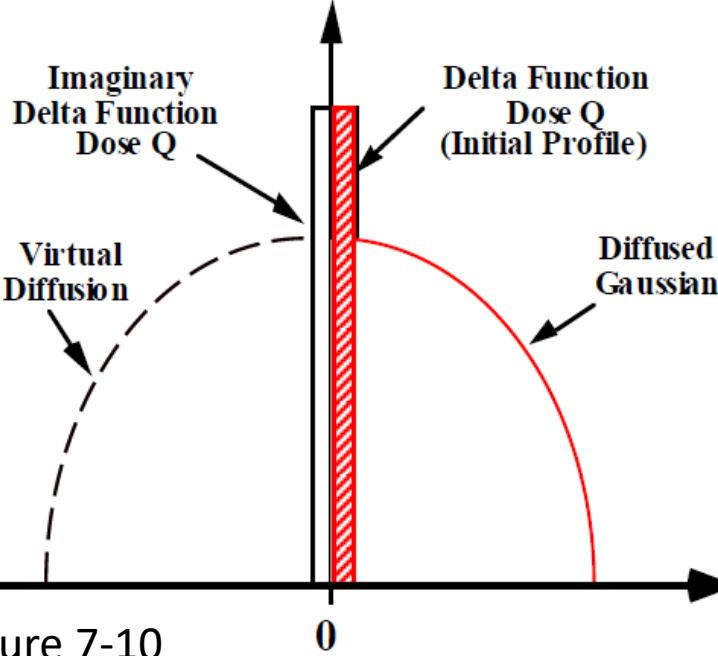


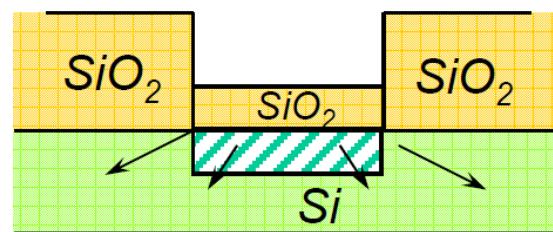
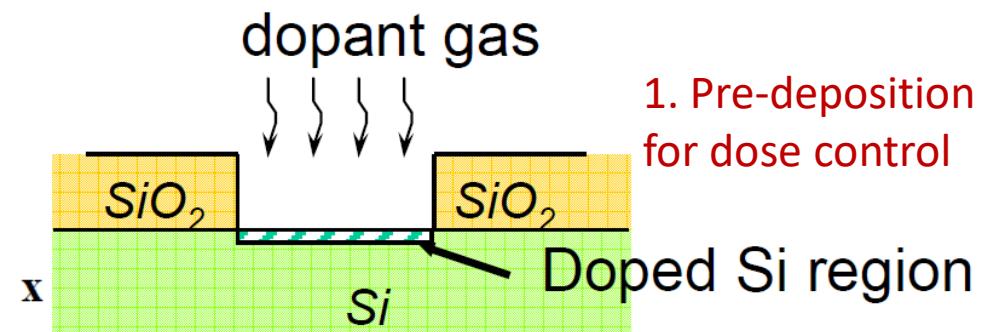
Figure 7-9

Gaussian solution near a surface



$$C(x,t) = \frac{Q}{\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right)$$

No "2" here



A surface Gaussian diffusion can be treated as a Gaussian diffusion with dose $2Q$ in an infinite bulk medium.

The top is capped with SiO_2 layer, thus no out-diffusion!

Gaussian solution near a surface

Surface concentration decreases with time

$$C_s = C(0, t) = \frac{Q_T}{\sqrt{\pi D t}}$$

Note: this C_s has nothing to do with solid solubility! Here C_s decreases with time, whereas solid solubility is a constant.

Junction depth

$$\begin{aligned}x_j &= 2 \sqrt{D t \ln \left(\frac{C_s}{C_B} \right)} \\&= 2 \sqrt{D t \ln \left(\frac{Q_T}{C_B \sqrt{\pi D t}} \right)}\end{aligned}$$

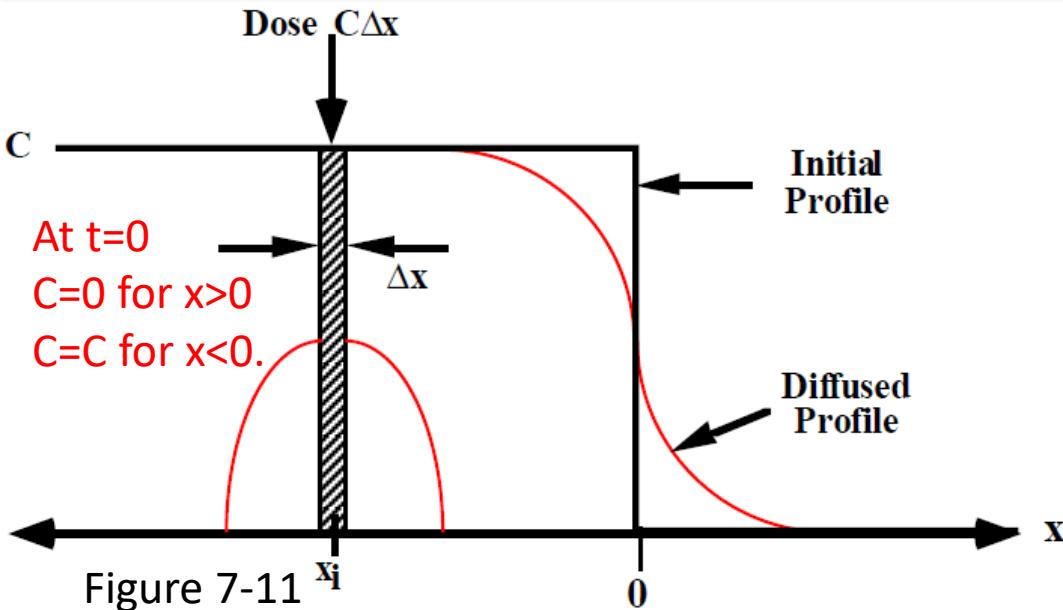
Concentration gradient

$$\frac{\partial C(x, t)}{\partial x} = -\frac{x}{2 D t} C(x, t)$$

At p-n junction

$$\left. \frac{\partial C(x, t)}{\partial x} \right|_{x_j} = -\frac{2 C_B}{x_j} \ln \left(\frac{C_s}{C_B} \right)$$

Error function solution in an infinite medium



$$C(x, t) = \frac{C}{2\sqrt{\pi Dt}} \sum_{i=1}^n \Delta x_i \exp\left[-\frac{(x - x_i)^2}{4Dt}\right]$$

$$C(x, t) = \frac{C}{2\sqrt{\pi Dt}} \int_{-\infty}^0 \exp\left[-\frac{(x - \alpha)^2}{4Dt}\right] d\alpha$$

$$\frac{(x - \alpha)}{2\sqrt{Dt}} = \eta$$

$$C(x, t) = \frac{C}{\sqrt{\pi}} \int_{x/2\sqrt{Dt}}^{+\infty} \exp[-\eta^2] d\eta = \frac{C}{\sqrt{\pi}} \left[\int_0^{+\infty} - \int_0^{x/2\sqrt{Dt}} \right]$$

This corresponds to, e.g. putting a thick heavily doped epitaxial layer on a lightly doped wafer.

An infinite source of material in the half-plane can be considered to be made up of a sum of Gaussians. The diffused solution is also given by a sum of Gaussians, known as the error-function solution.

$$\text{erf}(z) \equiv \frac{2}{\sqrt{\pi}} \int_0^z \exp(-\eta^2) d\eta$$

$$C(x, t) = \frac{C}{2} \left[1 - \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \right]$$

$$\text{erfc}(x) \equiv 1 - \text{erf}(x)$$

$$C(x, t) = \frac{C}{2} \left[\text{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) \right]$$

erfc: complementary error function

Error function solution in an infinite medium

Evolution of erfc diffused profile

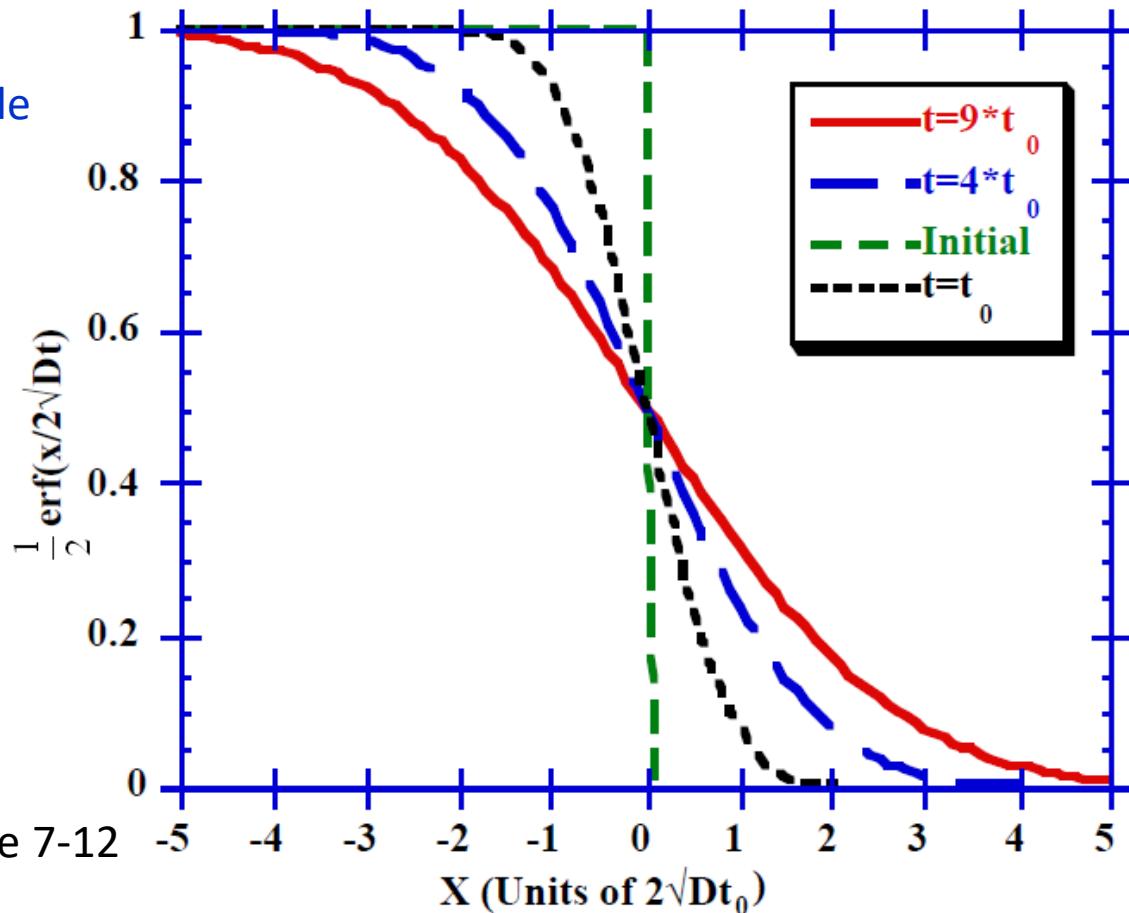


Figure 7-12

Important consequences of error function solution:

- Symmetry about mid-point allows solution for constant surface concentration to be derived.
- Error function solution is made up of a sum of Gaussian delta function solutions.
- Dose beyond $x=0$ continues to increase with annealing time.

Error function solution in an infinite medium

Properties of Error Function $\text{erf}(z)$ and Complementary Error Function $\text{erfc}(z)$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2) du$$

$$\text{erf}(0) = 0$$

$$\text{erf}(\infty) = 1$$

$$\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-u^2) du$$

$$\text{erfc}(0) = 1$$

$$\text{erfc}(\infty) = 0$$

$$\frac{d[\text{erf}(x)]}{dx} = \frac{2}{\sqrt{\pi}} \exp(-x^2)$$

$$\int_0^\infty \text{erfc}(x) dx = \frac{1}{\sqrt{\pi}}$$

$$\frac{d^2[\text{erf}(x)]}{dx^2} = -\frac{4}{\sqrt{\pi}} x \exp(-x^2)$$

For $x \ll 1$ $\text{erf}(x) \approx \frac{2}{\sqrt{\pi}} x$

For $x \gg 1$ $\text{erfc}(x) = \frac{1}{\sqrt{\pi}} \frac{\exp(-x^2)}{x}$

Error function solution near a surface

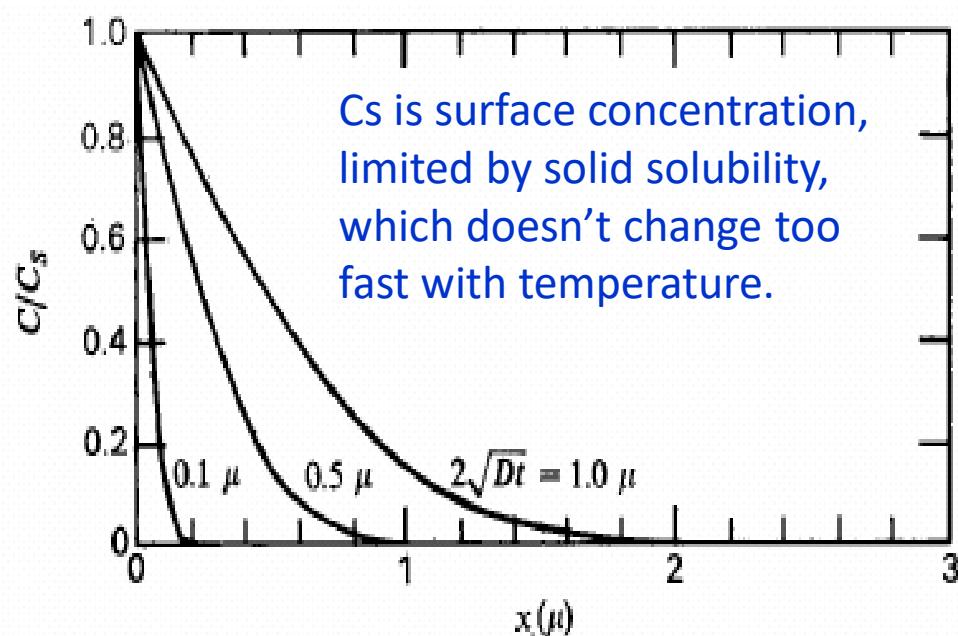
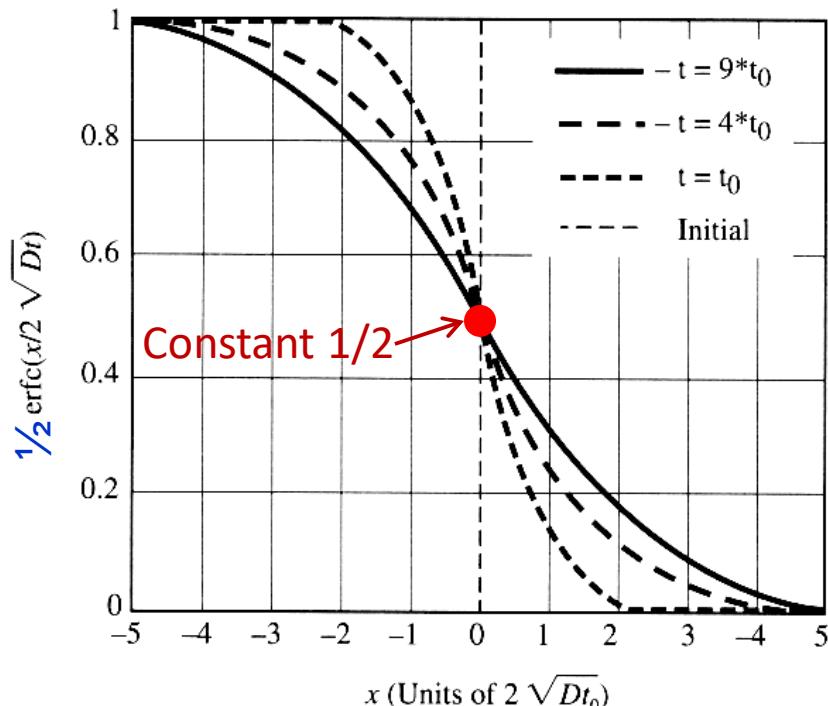
Constant surface concentration at all times, corresponding to, e.g., the situation of diffusion from a gas ambient, where dopants “saturate” at the surface ($C_s \equiv$ solid solubility). This is for doping, total dose Q increases with time; rather than drive-in (to spread out the dopant) for which the total dose Q is constant. Drive-in usually follows this doping process.

Boundary condition: $C(x,0)=0, x \neq 0; C(0,t)=C_s; C(\infty,t)=0$

$$C(x,t) = C_s \operatorname{erfc} \left(\frac{x}{2\sqrt{Dt}} \right) = \frac{2C_s}{\sqrt{\pi}} \int_{x/2\sqrt{Dt}}^{\infty} e^{-u^2} du$$

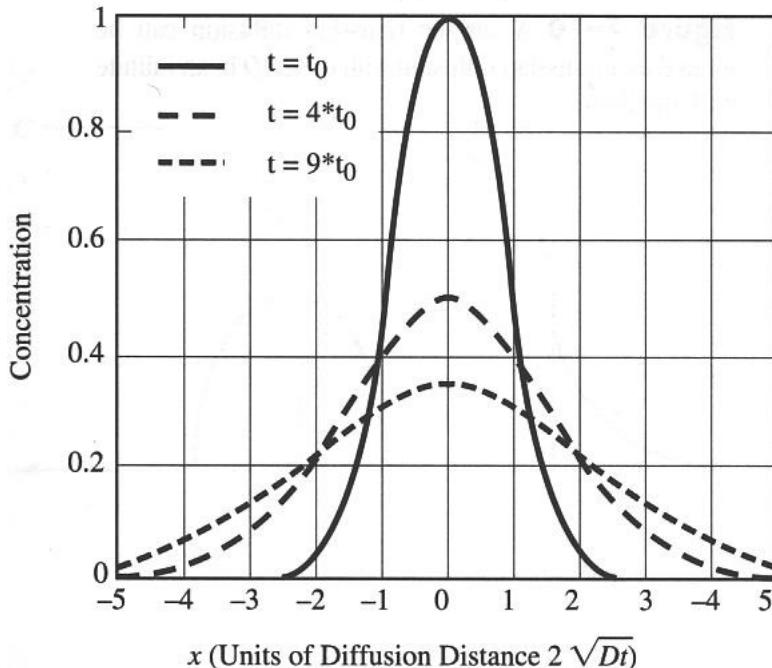
Pre-deposition dose

$$Q = \int_0^{\infty} C_s \operatorname{erfc} \left(\frac{x}{2\sqrt{Dt}} \right) dx = \frac{2C_s}{\sqrt{\pi}} \sqrt{Dt}$$



Successive diffusions

- Successive diffusions using different times and temperatures
- Final result depends upon the total Dt product



When D is the same (same temperature)

When diffused at different temperatures

$$(Dt)_{tot} = \sum D_i t_i$$

For example, the profile is a Gaussian i function at time $t=t_0$, then after further diffusion for another $3t_0$, the final profile is still a Gaussian with $t=4t_0=t_0+3t_0$.

(The Gaussian solution holds only if the Dt used to introduce the dopant is small compared with the final Dt for the drive-in i.e. if an initial δ/delta function approximation is reasonable)

$$(Dt)_{eff} = D(t_1 + t_2 + \dots + t_n)$$

$$(Dt)_{eff} = D_1 t_1 + D_2 t_2 + \dots = D_1 t_1 + D_1 t_2 \left(\frac{D_2}{D_1} \right) + \dots$$

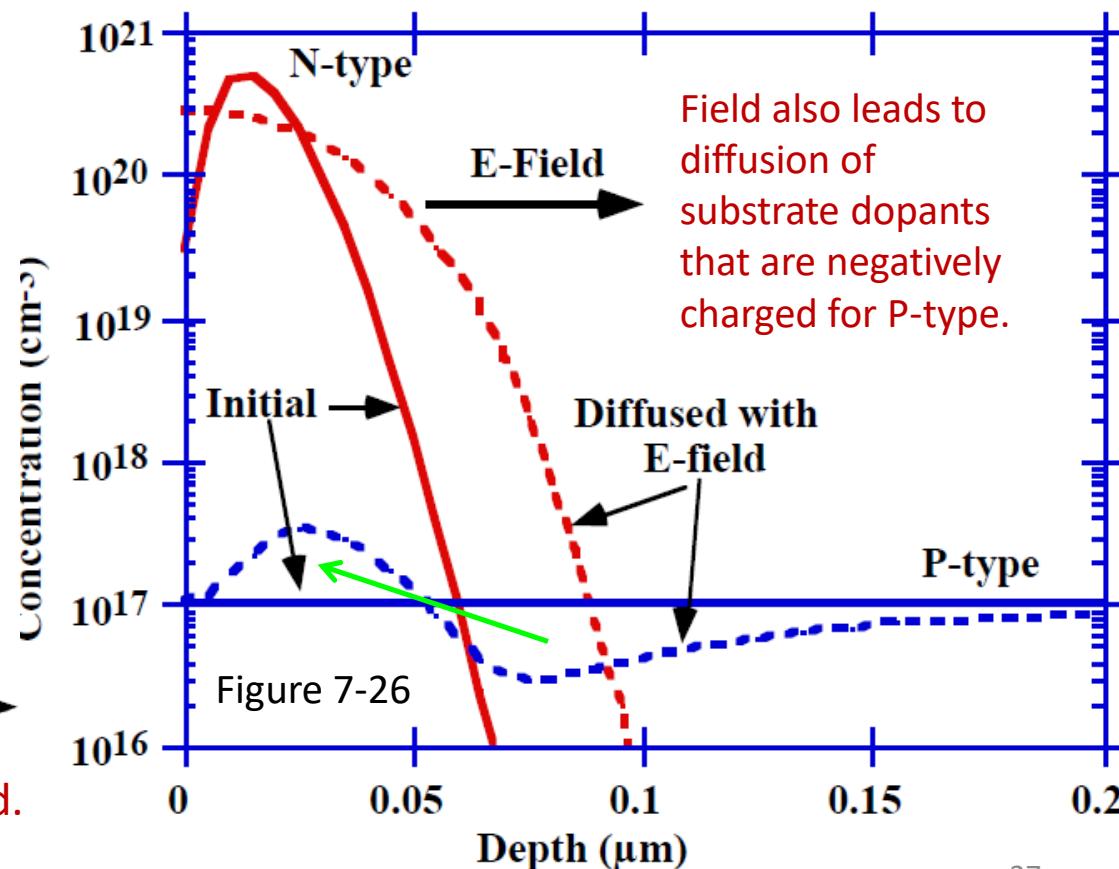
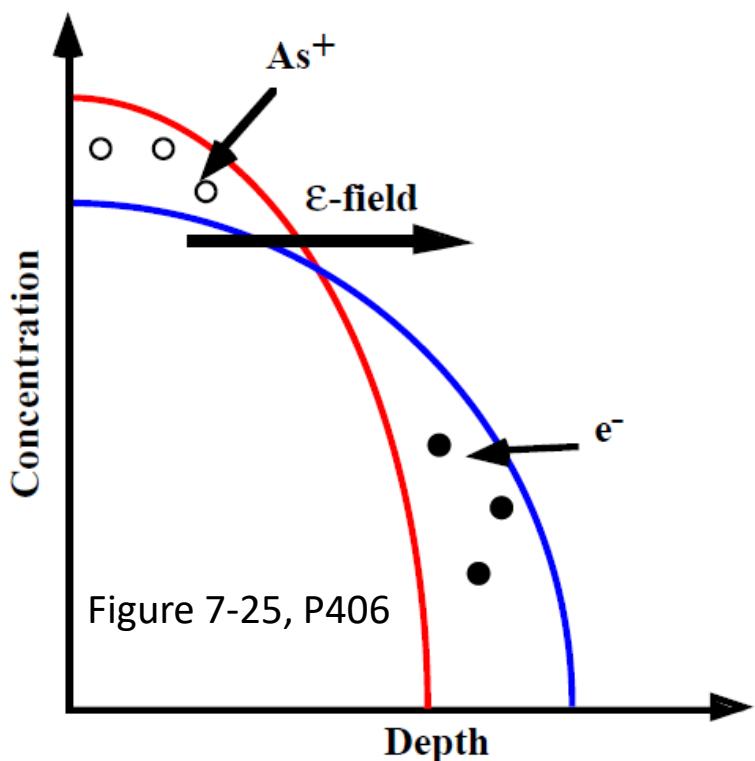
As D increases exponentially with temperature, total diffusion (thermal budget) is mainly determined by the higher temperature processes.

Chapter 7 Dopant Diffusion

1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

Electric field effect: non-intrinsic/extrinsic diffusion

- Fick's Laws: only valid for diffusion under special conditions.
- When the doping is higher than n_i , ϵ -field effects become important .
- The field comes from higher mobility of electrons and holes compared to dopant atoms.
- Carries move faster than dopants, until an equilibrium is attained due to the electric field that slows down carrier and speeds up dopant.



Concentration dependent diffusion

At high doping concentrations, the Fick's equation must be solved numerically since $D \neq \text{constant}$.

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D_A^{\text{eff}} \frac{\partial C}{\partial x} \right)$$

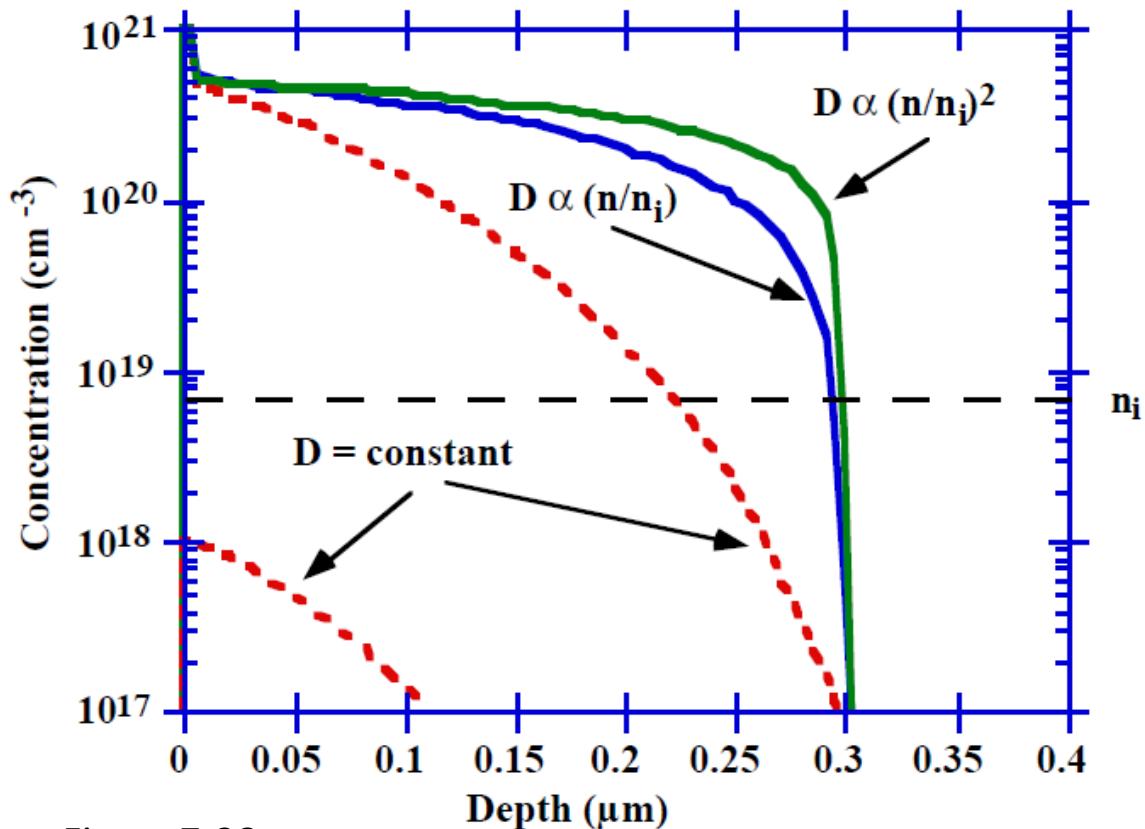


Figure 7-28

The dash line show the erfc profiles.

The solid lines are numerical simulations, which agree with experimental results.

Those box-like (steep) doping profile is often desirable.

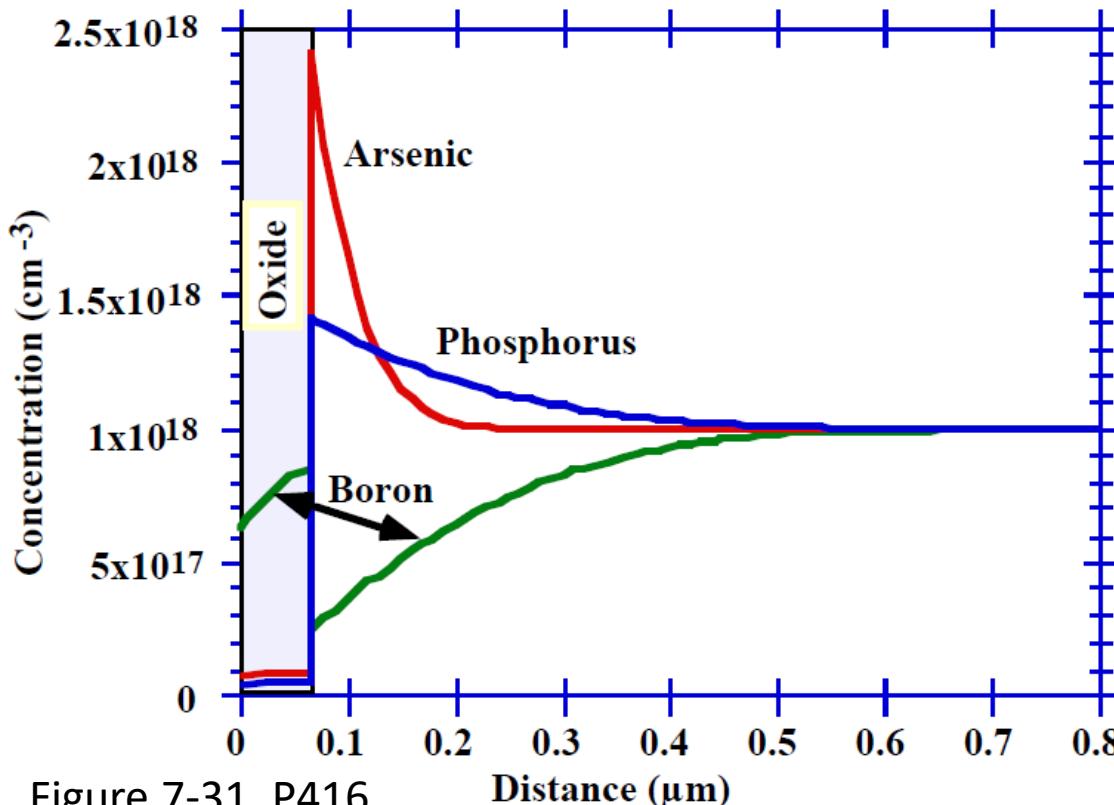
Chapter 7 Dopant Diffusion

1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

Dopant segregation

Dopants segregate coefficient k_0 .

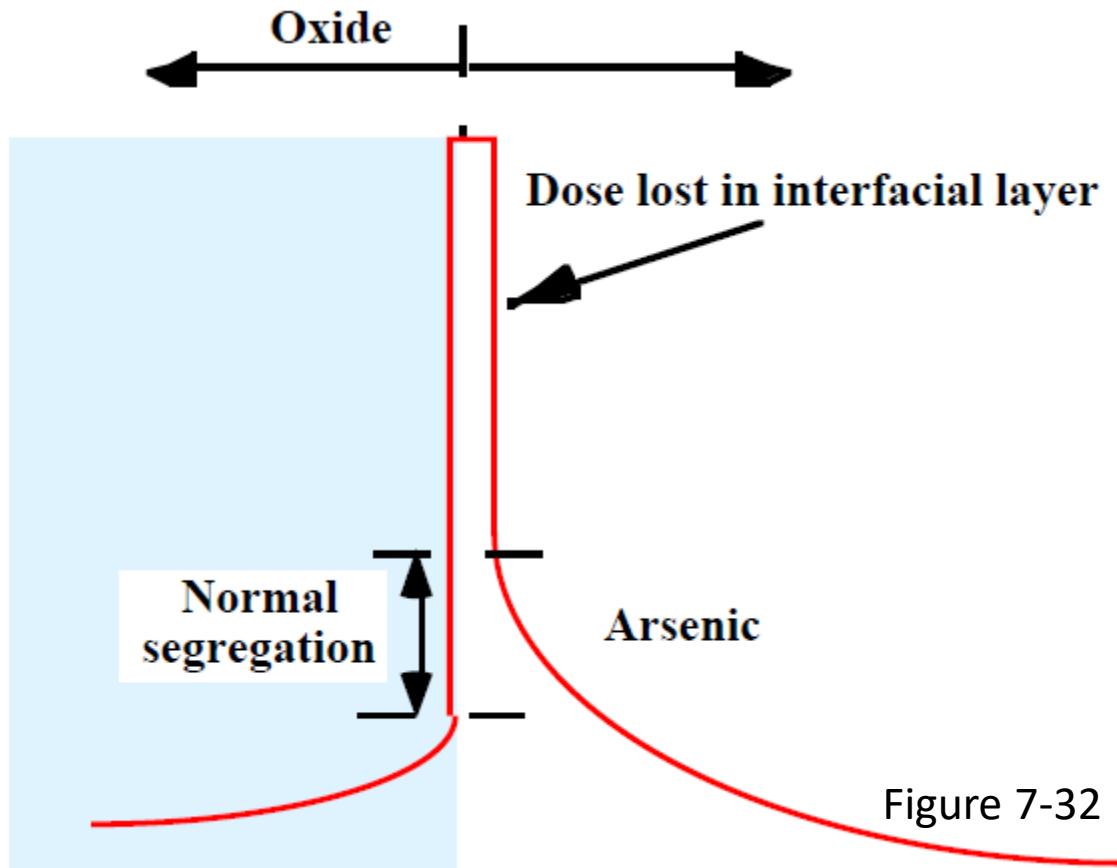
$$k_0 = \frac{C_{Si}}{C_{SiO_2}} \approx \begin{cases} 0.3 \text{ for Boron, B} \\ 10 \text{ for Arsenic, As} \\ 10 \text{ for Antimony, Sb} \\ 10 \text{ for Phosphorus, P} \end{cases}$$



Oxidation of a uniformly doped boron substrate depletes the boron into the growing SiO_2 . Whereas N-type dopants tend to pile-up near the interface.

Figure 7-31, P416

Interfacial dopant pile-up during oxidation



Dopants may also segregate to an interface layer, perhaps only a monolayer thick.

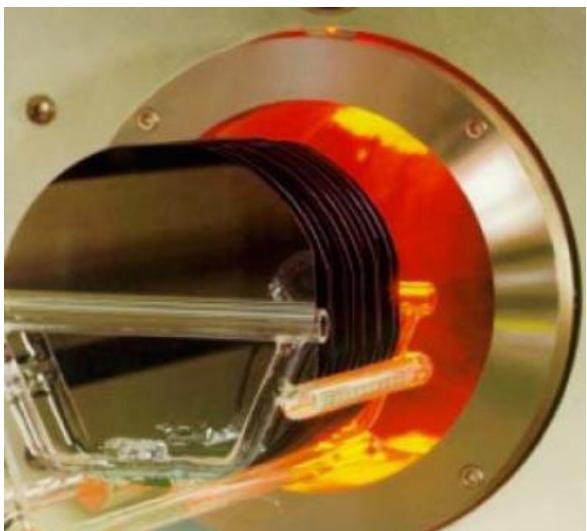
They are not active (do not contribute electrons/holes).

This may consume up to 50% of the dose in a shallow layer.

Chapter 7 Dopant Diffusion

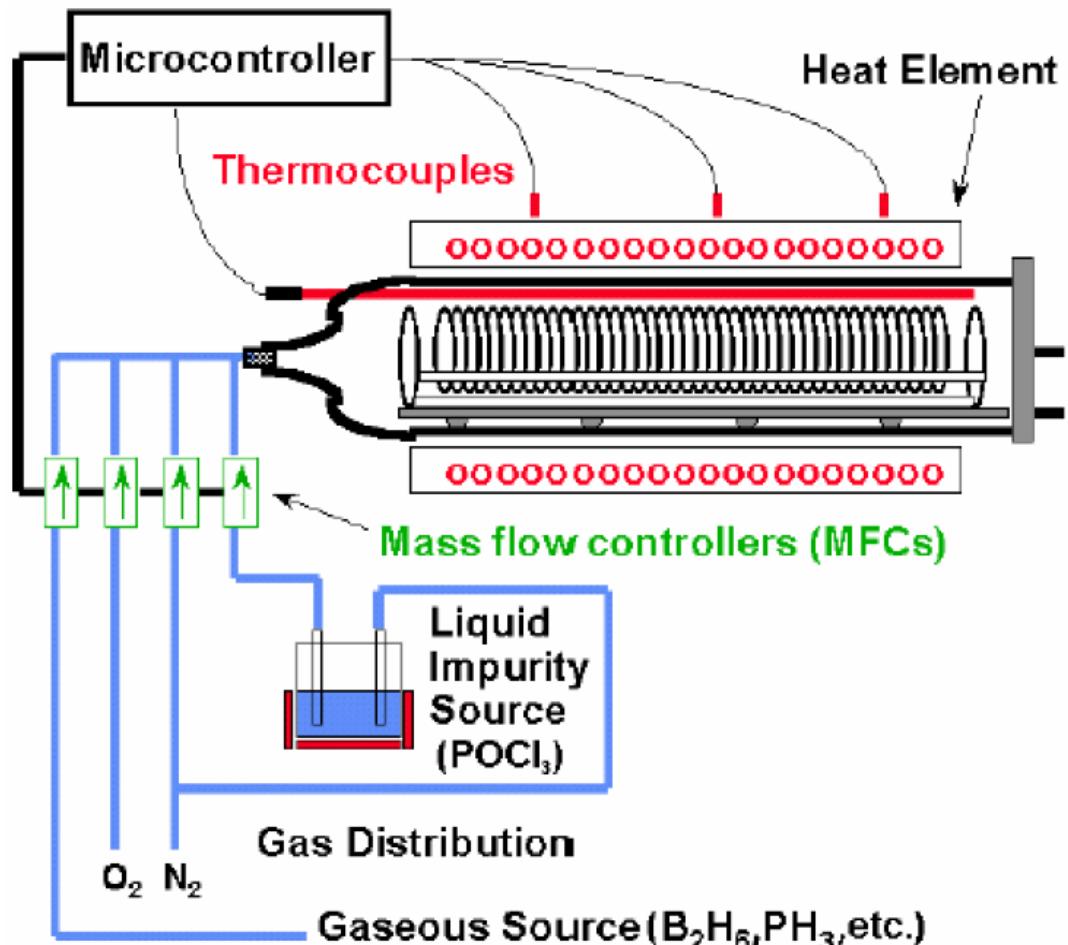
1. Introduction and application.
2. Dopant solid solubility and sheet resistance.
3. Microscopic view point: diffusion equations.
4. Physical basis for diffusion.
5. Non-ideal and extrinsic diffusion.
6. Dopant segregation and effect of oxidation.
7. Manufacturing and measurement methods.

Diffusion furnace: horizontal



Temperature control in 3 to 5 zones by spike thermocouples.

Diffusion tube and boat made from SiO_2 , SiC ...



Horizontal furnaces were the diffusion work horses up to 200 mm wafers

Diffusion furnace: vertical

Vertical furnaces are used for 200mm and above wafers.



Vertical furnace for 300mm wafers.



Close view into three tubes

Diffusion methods

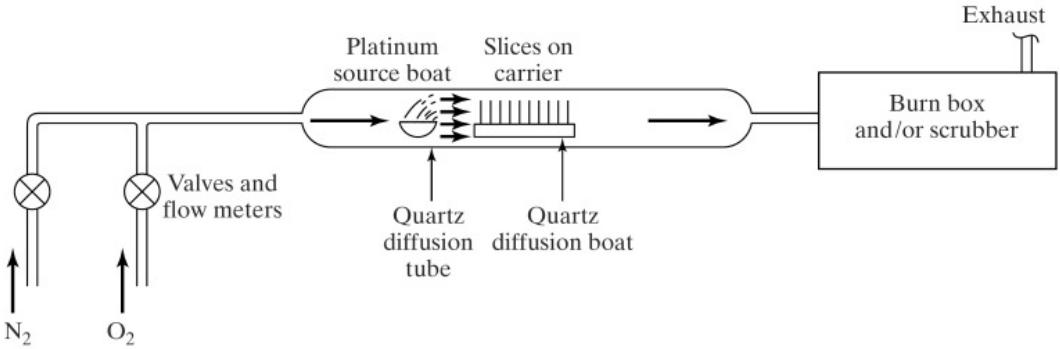
Gas source

Liquid source

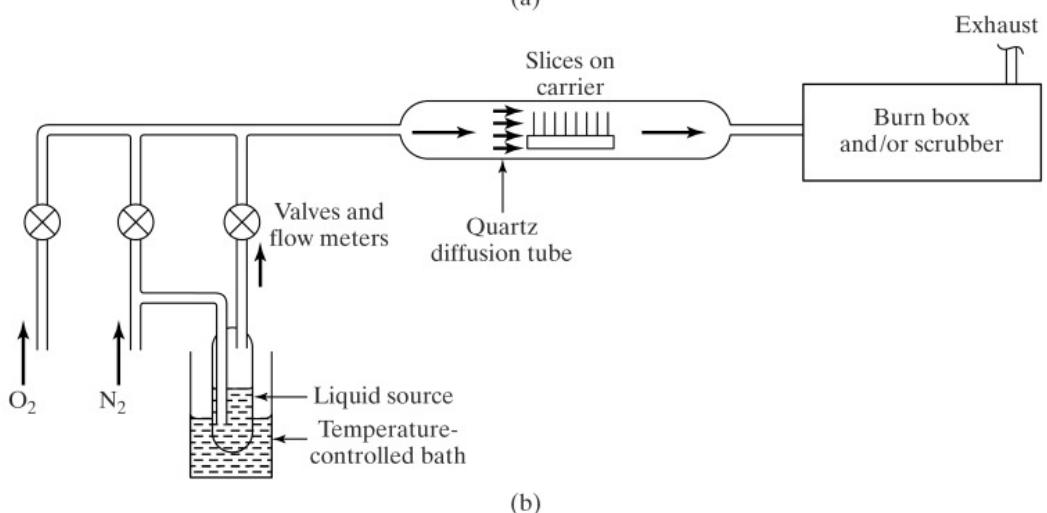
Solid source

Spin-on glass

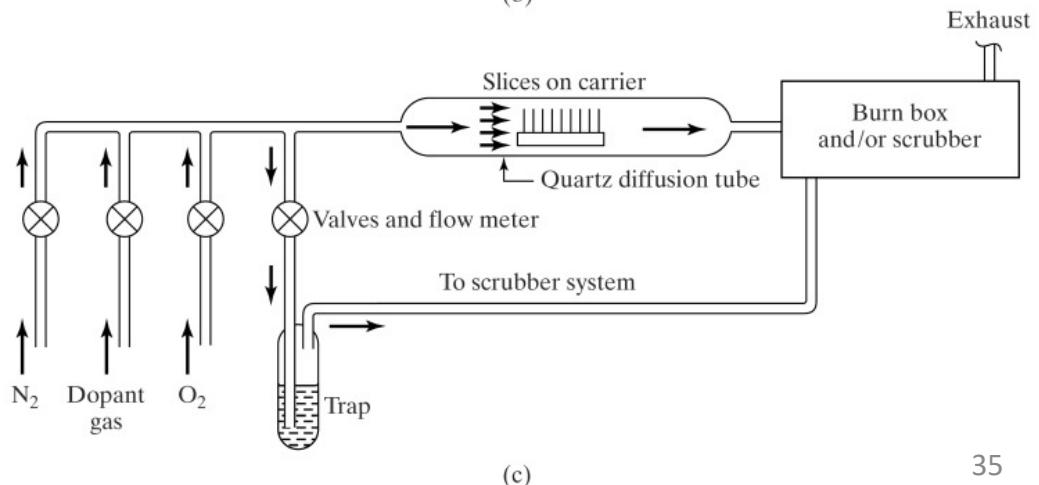
Note: for drive-in, a capping layer ($\text{SiO}_2\ldots$) is often used to prevent out-diffusion into air.



(a)



(b)



(c)

Open furnace tube systems

- (a) Solid source in platinum source boat
- (b) Liquid Source - carrier gas passing through bubbler
- (c) Gaseous impurity source

Gas source doping, very toxic gases

TABLE 4.4 Threshold Limit Recommendations for Common Gaseous Sources [21] *

Source	8-h exposure level (ppm)	Life-threatening exposure	Comments
Diborane (B_2H_6)	0.10	160 ppm for 15 min	Colorless, sickly sweet, extremely toxic, flammable.
Phosphine (PH_3)	0.30	400 ppm for 30 min	Colorless, decaying fish odor, extremely toxic, flammable. A few minutes' exposure to 2000 ppm can be lethal.
Arsine (AsH_3)	0.05	6–15 ppm for 30 min	Colorless, garlic odor, extremely toxic. A few minutes' exposure to 500 ppm can be lethal.
Silane (SiH_4)	0.50	Unknown	Repulsive odor, burns in air, explosive, poorly understood.
Dichlorosilane (SiH_2Cl_2)	5.00	...	Colorless, flammable, toxic. Irritating odor provides adequate warning for voluntary withdrawal from contaminated areas.

Silane and dichlorosilane used for
polysilicon deposition

*Data from the 1979 American Conference of Governmental Hygienists (ACGIH).

Spin on glass doping

- Spin-coat an oxide on the wafer (room temperature).
- Bake at 200°C for 15min to remove solvent.
- Effectively, the film can be considered as a mixture of SiO_2 and dopant oxide.
- Then diffuse into Si at high temperature.
- Good for many types of dopant, with a wide range of dose.
- This is finite source, whereas gas/liquid/solid are infinite source.

Sources:

As: arsenosilica

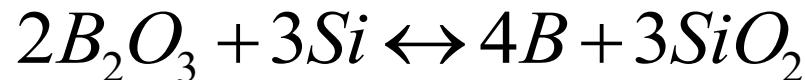
Sb: antimonysilica

B: borosilica

P: phosphorosilica

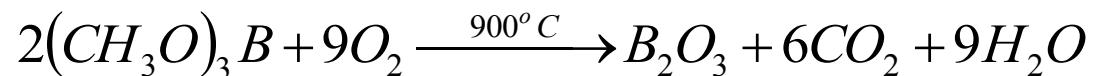
Diffusion system: Boron

Surface reaction:

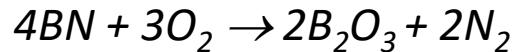


Solid sources: boron nitride (BN) and trimethylborate (TMB).

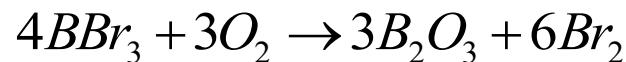
TMB has high vapor pressure at room temperature,
so placed outside of furnace.



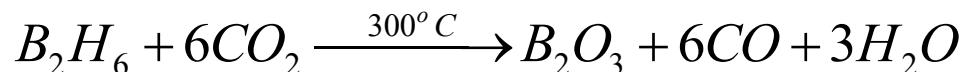
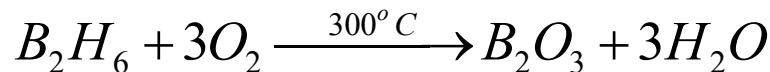
One can also use BN wafers, pass O₂, 900°C.



Liquid sources: boron tribromide BBr₃.



Gaseous sources: diborane B₂H₆.



Diffusion system: Phosphorus

Surface reaction:



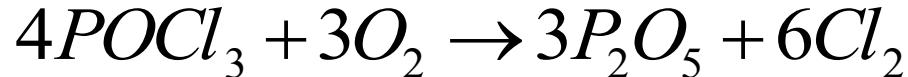
Solid sources: (can be made into wafers like BN, but not popular)

Phosphorus pentoxide

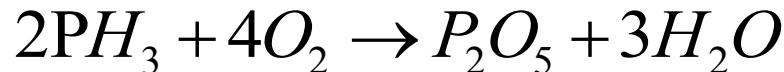
Ammonium monophosphate $NH_4H_2PO_4$

Ammonium diphosphate $(NH_4)_2HPO_4$

Liquid source: phosphorus oxychloride $POCl_3$

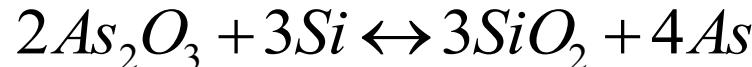


Gaseous source: phosphine PH_3 .



Diffusion system: Arsenic & Antimony

Arsenic surface reaction:

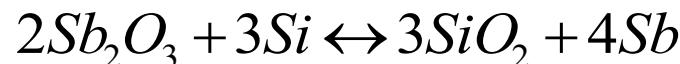


Solid sources: possible, low surface concentrations.

Gaseous source: arsine AsH_3 .

Ion implantation is normally used for deposition.

Antimony surface reaction:



Liquid source: antimony pentachloride Sb_3Cl_5 .

Ion implantation is normally used for deposition.

Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

Ion implantation and its history

A gas is ionized, and the ions are accelerated by a high electric field, and injected into the target wafer to hundreds of nm depth.

Typical ion implantation parameters:

Ion: P, As, Sb, B, In, O

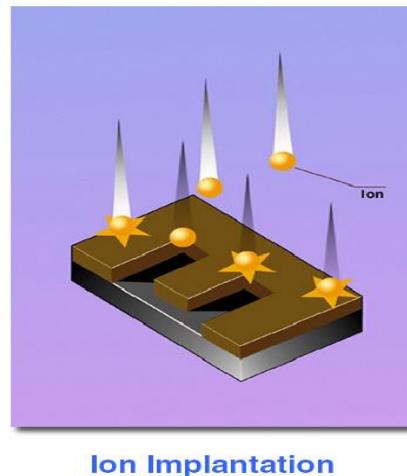
Dose: 10^{11} - 10^{18} cm⁻²

Ion energy: 1 - 400 keV

Uniformity and reproducibility: $\pm 1\%$

Temperature: room temperature

Ion flux: 10^{12} - 10^{14} ions/(cm²·s)



Dose (ϕ) = # of atoms/cm².
Concentration (C) = # of atoms/cm³.

$$\text{dose } \phi = \int_0^{\infty} C(x) dx$$

- The idea was proposed by Shockley in 1954, but used for mass production only after late 1970s.
- Before ion implantation, doping is achieved by diffusion into the bulk silicon from gaseous source above surface, or pre-deposited chemical source on wafer surface.
- Diffusion lacks the flexibility and control required by CMOS processing, and ion implantation quickly gained popularity for the introduction of dopant atoms.
- Modern ion implanters were originally developed from particle accelerator technology. The implantation is always followed by a thermal activation (600-1100°C).

Advantages of ion implantation

- Very precise dose control.

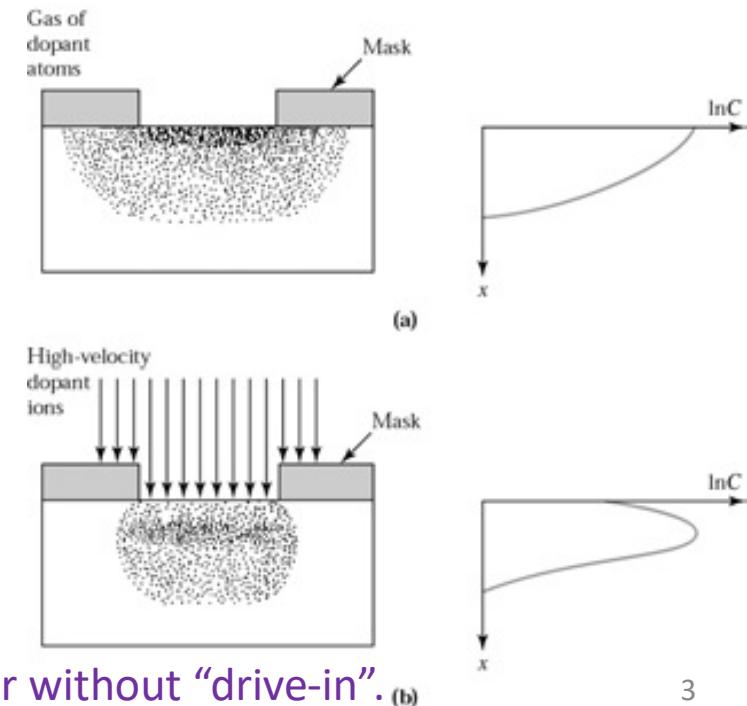
The ion implanter forms a simple electrical circuit. By monitoring the current in the circuit (or by a monitoring circuit with Faraday cups), significant accuracy in the implanted dose can be maintained. Assuming a current sensitivity of nA, and a minimum required implantation time of 10 seconds, it can be shown that doses as low as 10^{11}cm^{-2} , can be measured.

On the contrary, in chemical source predeposits, dose values less than $5\times10^{13}/\text{cm}^2$ are not achievable.

- High dose introduction is not limited to solid solubility limit values.
- Excellent doping uniformity is achieved across the wafer (< 1% variation across 12" wafer) and from wafer to wafer.
- Less dopant lateral diffusion, good for small device.
- Done in high vacuum, it is a very clean process step.
- Besides precise dose control, one can also control the profile (peak depth and spread range) better than diffusion (peak concentration always near surface).

(top) Doping by diffusion and “drive-in”.

(bottom) Doping by ion implantation with or without “drive-in”. (b)



Advantages/disadvantage of ion implantation

Advantage:

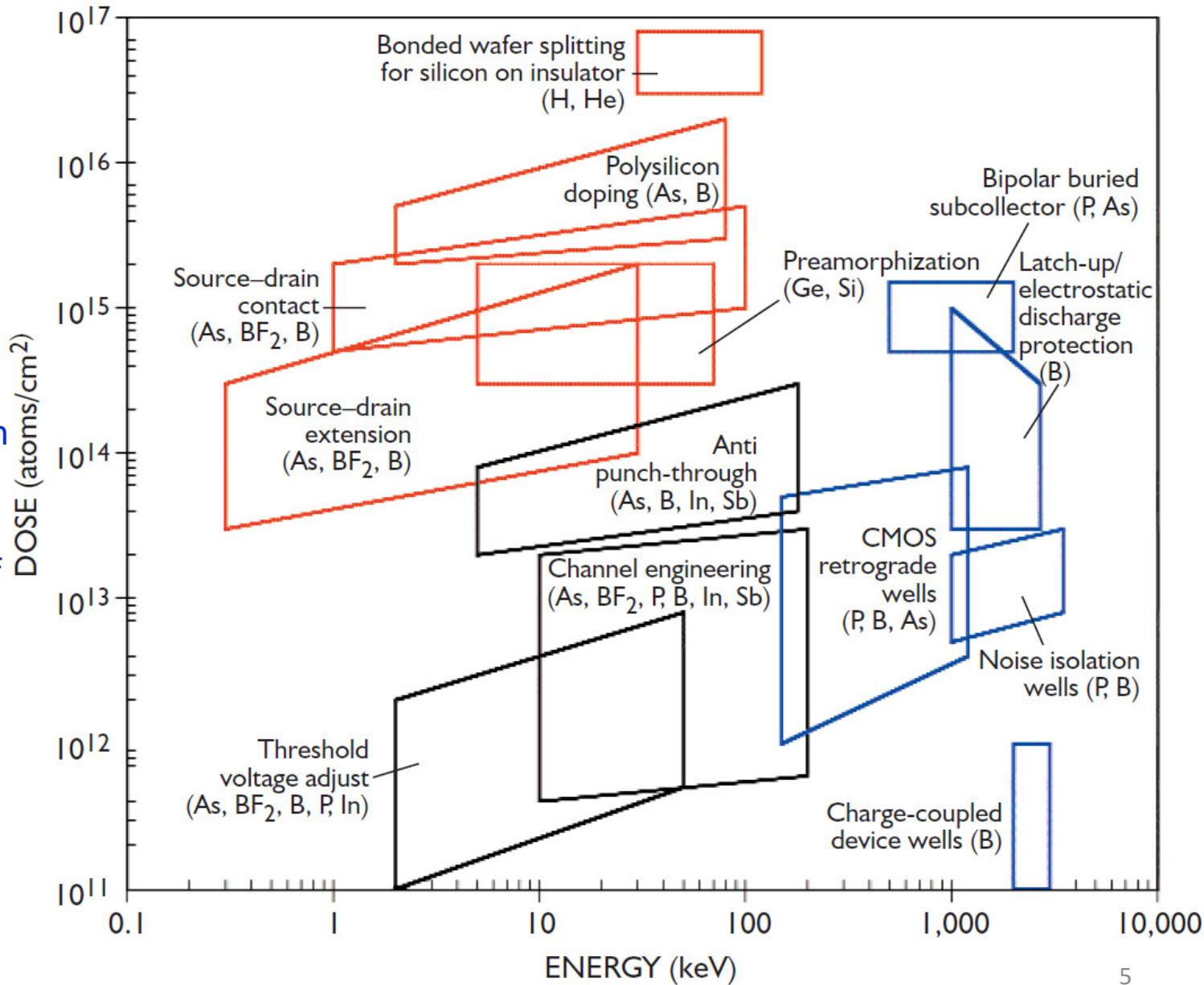
- Low-temperature process (can use photoresist as mask)
- Wide selection of masking materials, e.g. photoresist, oxide, poly-Si, metal
- Less sensitive to surface cleaning procedures.
- Very fast (6" wafer can take as little as 6 seconds for a moderate dose)
- Complex profiles can be achieved by multi-energy implants.

Disadvantage:

- Very expensive equipment (\$1M or more).
- At high dose values, throughput is less than diffusion (chemical source pre-deposition on surface).
- Ions damage the semiconductor lattice. Not all the damage can be corrected by annealing.
- Very shallow doping are difficult or impossible.
- Masking materials can be “knocked” into the wafer creating unwanted impurities.

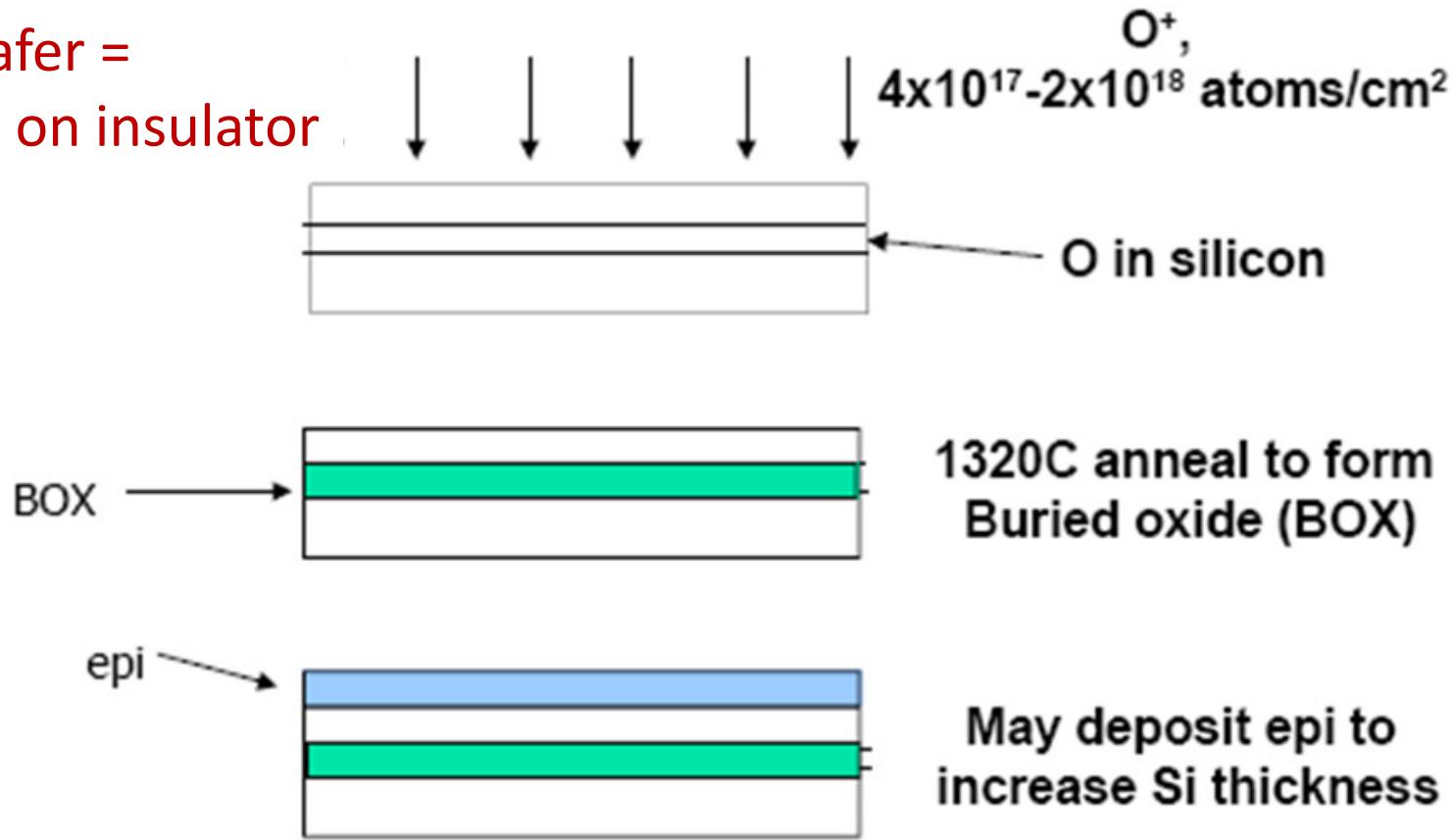
Dose-energy application space

Dose and energy requirements of major implantation applications
(species shown roughly in order of decreasing usage).



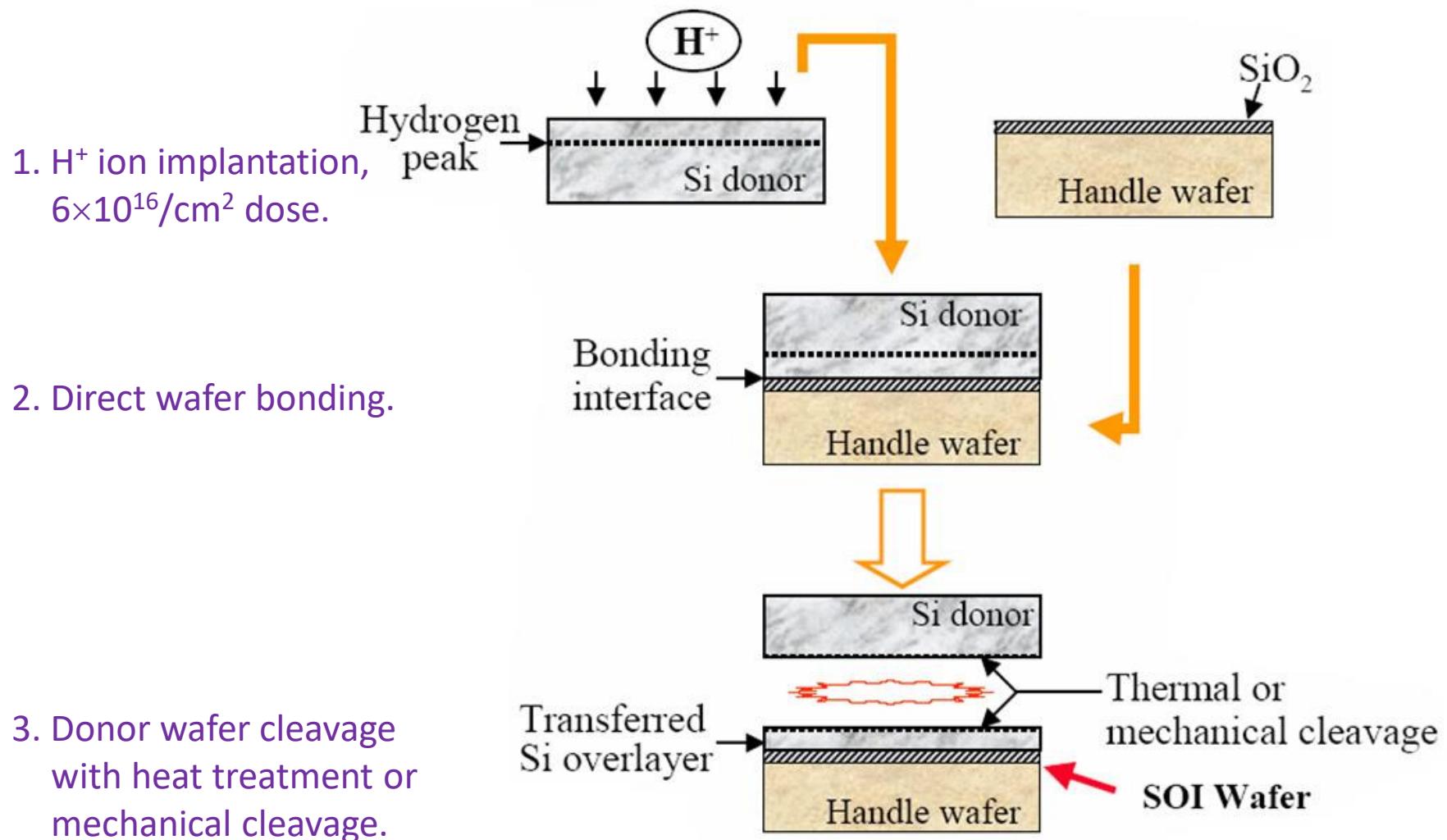
SIMOX (Separation by IMplantation of OXygen) for SOI wafer

SOI wafer =
silicon on insulator



- SOI wafers provide better performance for high speed circuits than conventional wafers.
- SOI wafer by SIMOX has better yield, but more expensive than wafer bonding and CMP process. (CMP = chemical mechanical polishing)
- 400nm buried oxide requires a high oxygen dose of $2 \times 10^{18}/\text{cm}^2$, which is a slow process.

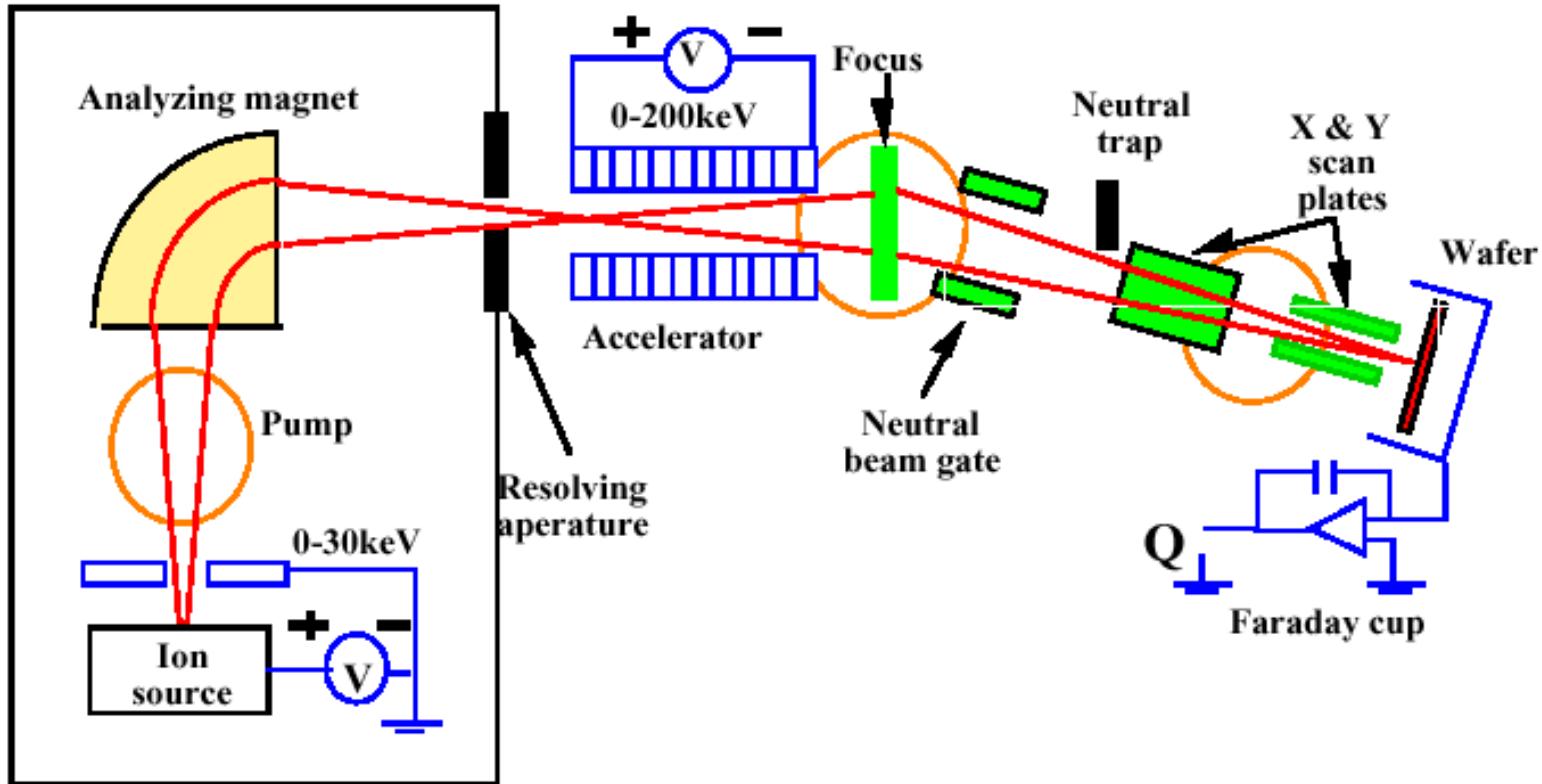
SOI wafer by smart-cut using hydrogen implantation



Chapter 8 Ion implantation

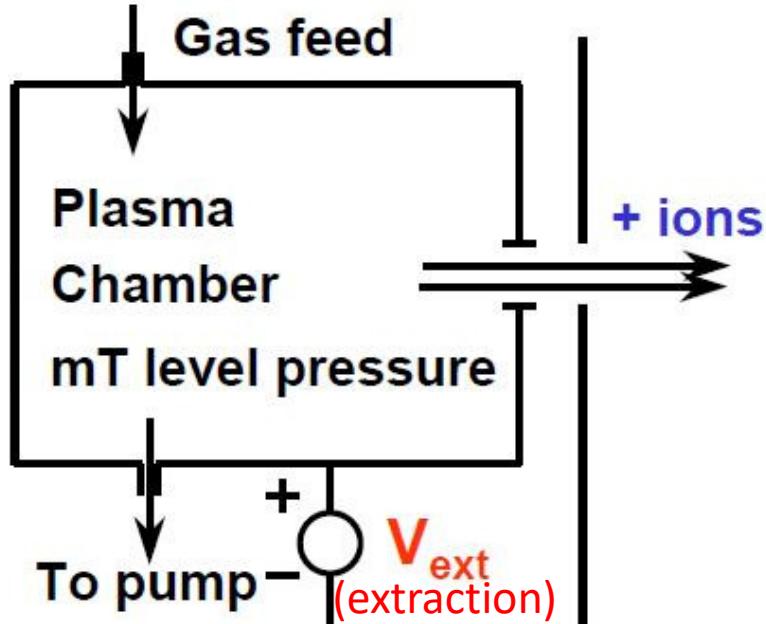
1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

Schematic of an ion implanter



- **Ion source:** operates at a high voltage(25kV) and convert the electrically neutral dopant atoms in the gas phase into plasma ions and undesired species.
- **Mass spectrometer:** a magnet bend the ion beam through right angle, and select the desired impurity ion and purge undesired species. Selected ion passes through an aperture.
- **Accelerator:** add energy to beam up to 200keV. (contained, to shield possible x-ray).
- **Scanning system:** x and y axis deflection plates are used to scan the beam across the wafer to produce uniform implantation of desired dose. The beam is bended to prevent the neutral particles from hitting the target .

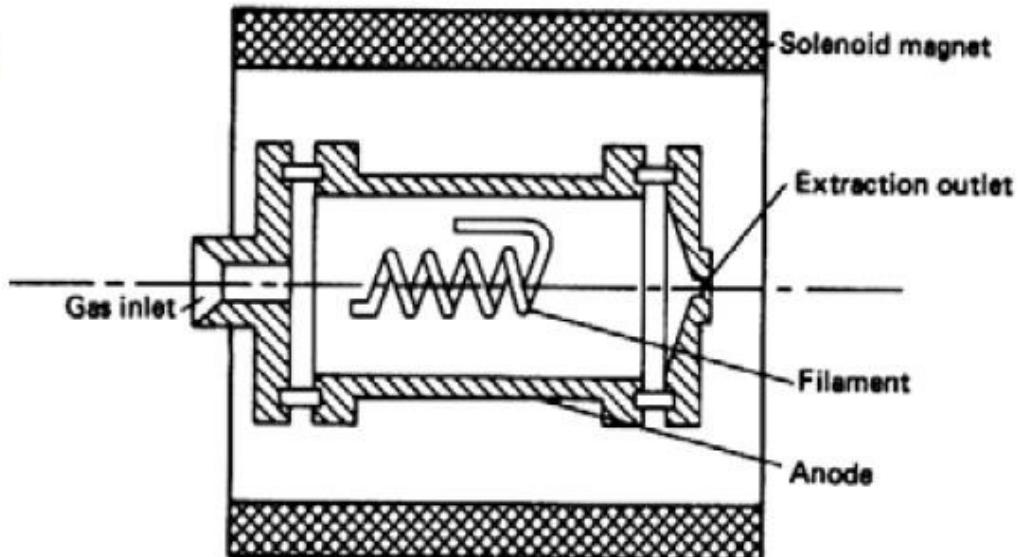
Plasma ion source and ion extraction



Variable extraction voltage
(typically $\sim 30\text{KV}$)

Plasma ion source

Nielsen-type gaseous source
(no plasma)



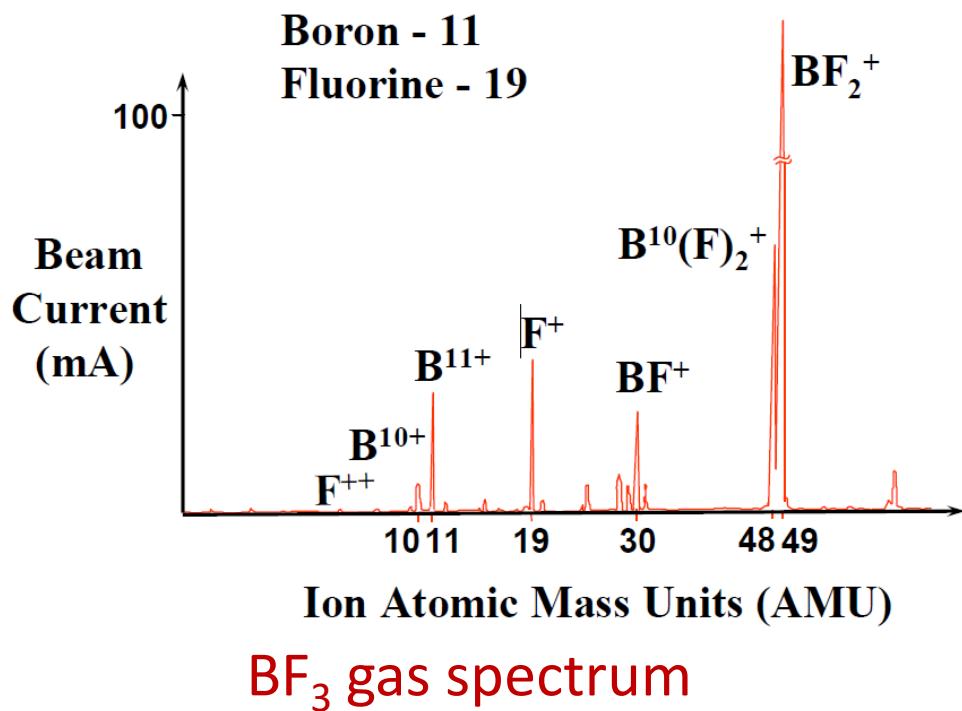
Positive ions are attracted to the exit side of the source chamber, which is biased at a large negative potential with respect to the filament.



- Filament emits electrons, which are accelerated to gain enough energy.
- The electrons collide with the molecules or atoms, and ionize them.
- The ions are extracted, rough-focused, then travel toward the magnetic analyzer.

Select implantation species

- Dopant gas containing desired species, BF_3 , BCl_3 , B_2H_6 , PH_3 , AsH_3 , AsF_5 .
- If using solid or liquid source, they need to be heated to vaporize .
- Plasma provides positive ions, $(\text{B}^{11})^+$, BF_2^+ , $(\text{P}^{31})^+$, $(\text{P}^{31})^{++}$.

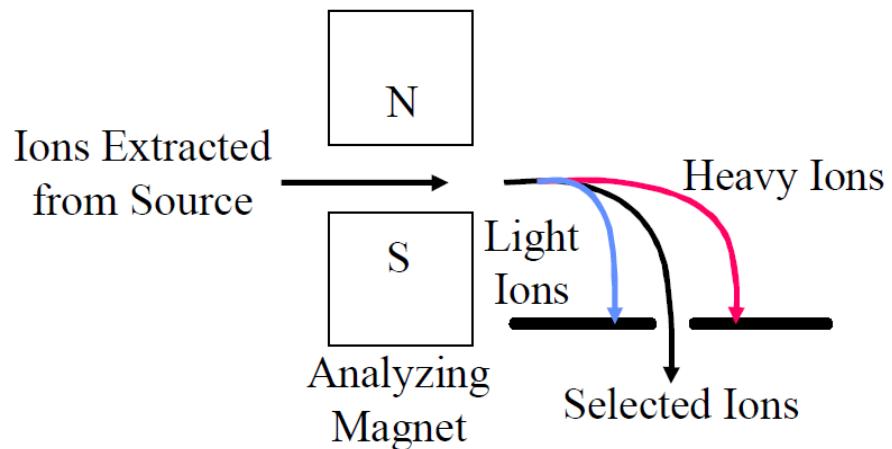


$$\frac{mv^2}{r} = qvB \quad v = \sqrt{\frac{2E}{m}} = \sqrt{\frac{2qV}{m}}$$

$$R = \frac{mv}{qB} = \frac{1}{B} \sqrt{\frac{2mV}{q}}$$

Mass to charge ratio of the selected ions:

$$m/q = R^2 B^2 / (2 V_{ext})$$



Only one mass will have exactly the correct radius of curvature to exit the source through the slit.

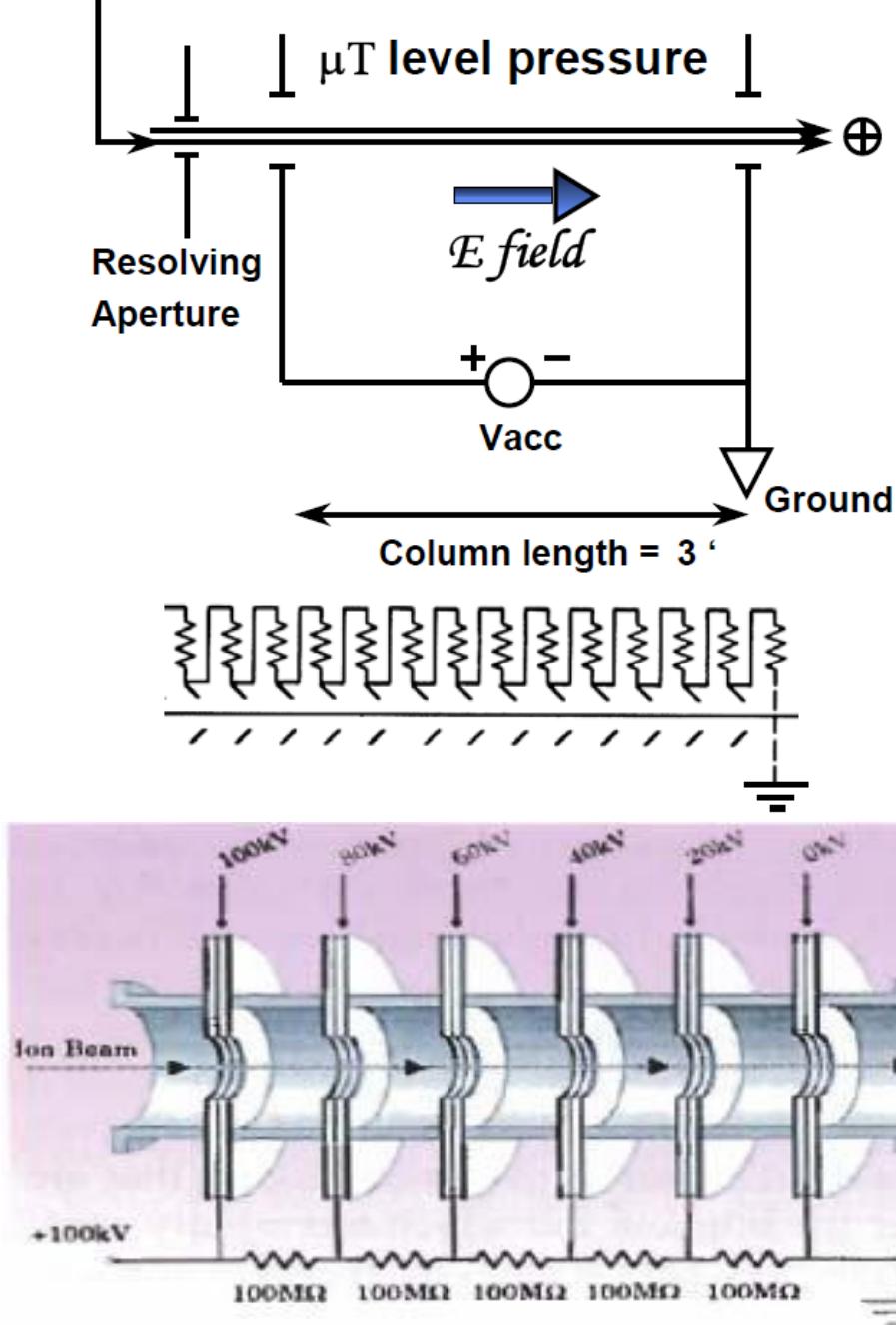


Fig. 12-23 Acceleration-column of an ion implanter.

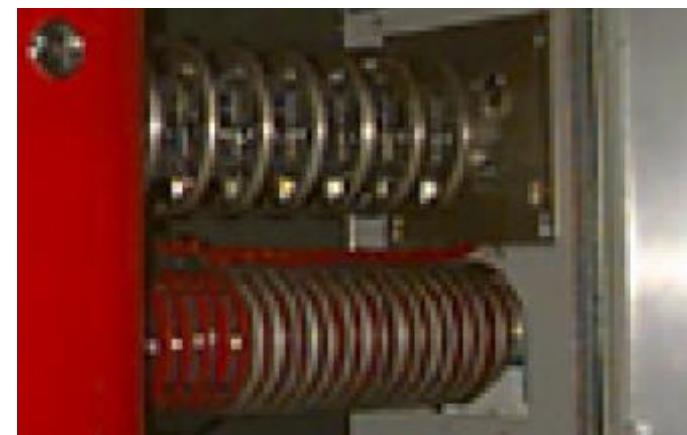
Ion acceleration

Final kinetic energy of the ion =

$$Q (V_{ext} + V_{acc})$$

Example: $V_{ext} = 30\text{KV}$, $V_{acc} = 70\text{KV}$

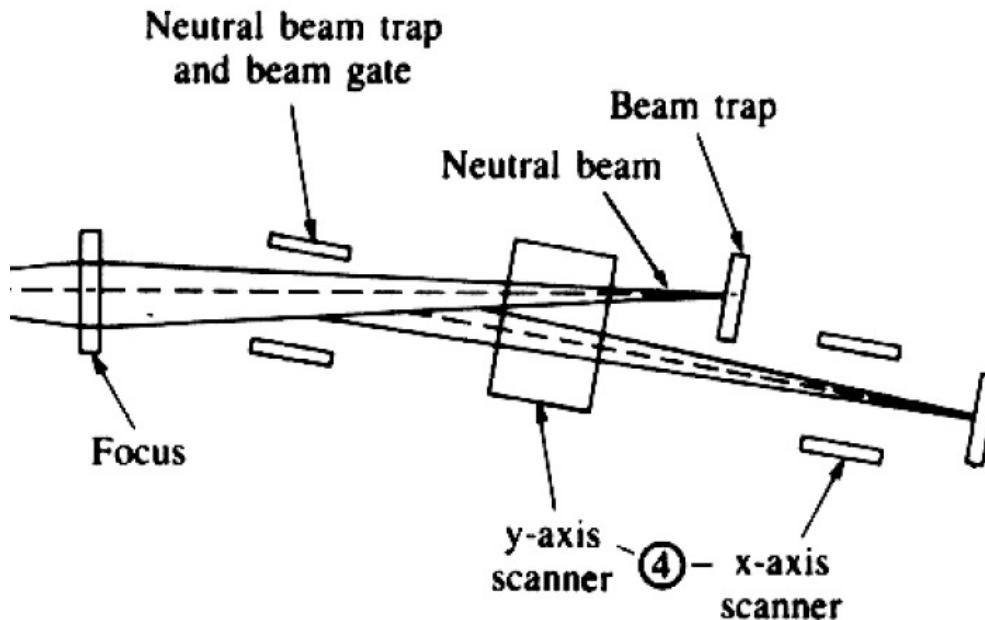
Energy of the ion = 100 keV



This shows 14 equal acceleration plates.

For example, if the desired acceleration is 70keV, each section would contribute 5k volts.

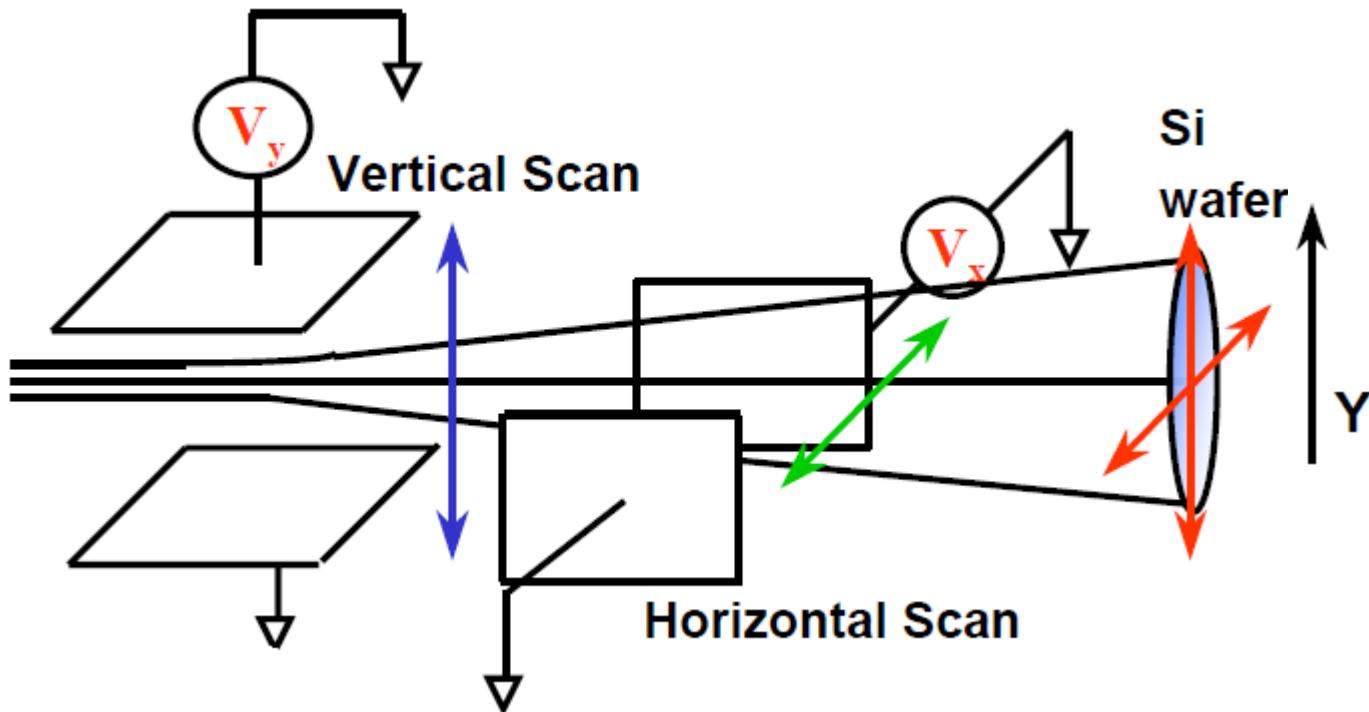
Neutral trap



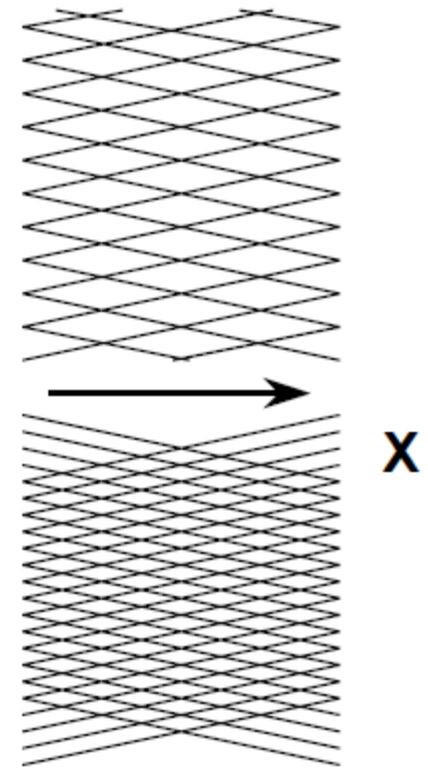
- There may be ions that collided with other ions in the beam, undergoing a charge exchange.
- Neutrals are highly undesirable since they will not be deflected in electrostatic end-station scanning mechanisms.
- To avoid this problem, most systems are equipped with a bend, i.e. parallel plates of an electrostatic deflection system.
- Neutrals are not deflected and so do not follow the beam, but instead strike a beam stop.
- The ions are sufficiently deflected by the plates to continue to travel down the tube.

Beam scanning: electrostatic scanning

Electrostatic scanning: low/medium beam current implanters, $I < 1\text{mA}$.



Scan Patterns

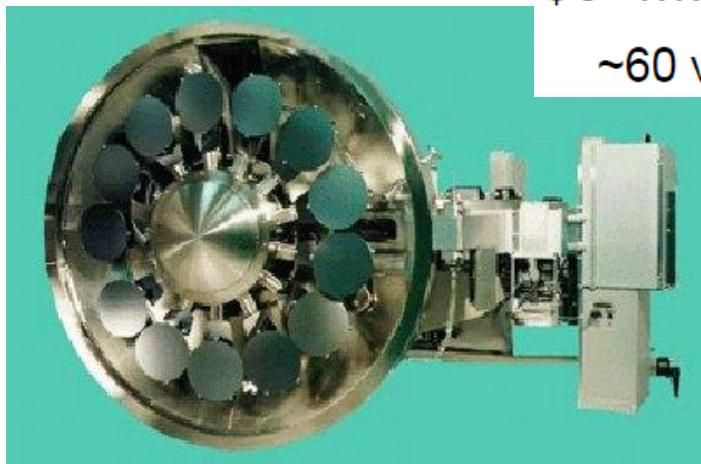


This type of implanter is suitable for low dose implants. With scan frequencies in the 100 Hz range, good implant uniformity is achieved with reasonable throughput.

The focused ion beam is scanned over the wafer in a highly controlled manner in order to achieve uniform doping.

Mechanical scanning

- 15 - 25 wafers /disc.
- Excellent throughput for high dose implants.
- High current (~20mA)
- Scan speed adjustment to insure uniform dose.

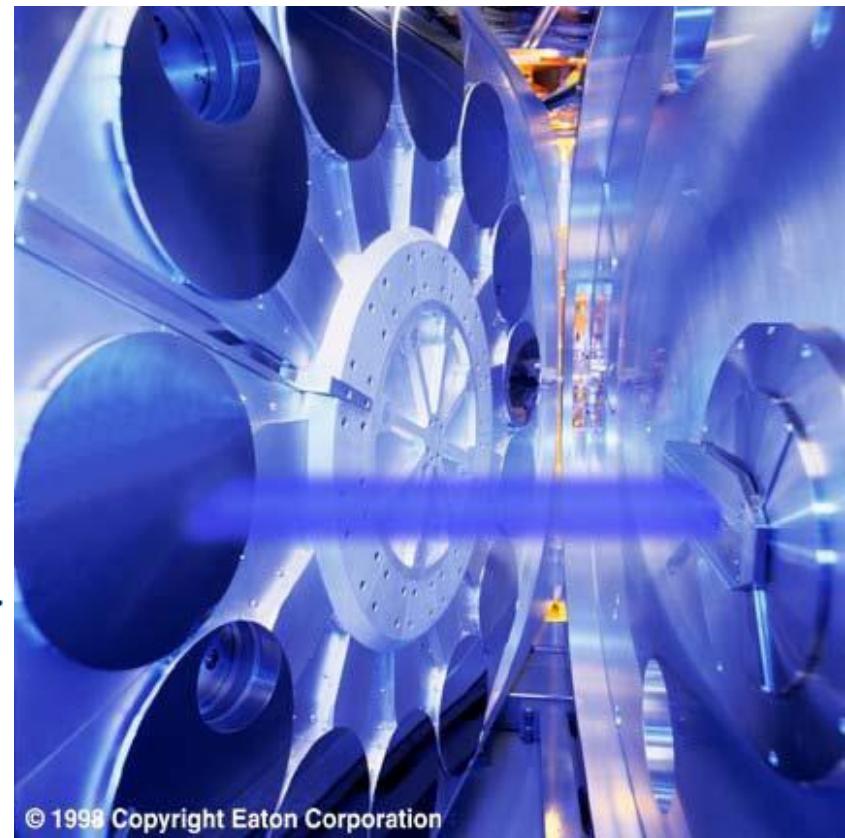


Multi-wafer, mechanical scanning end-station

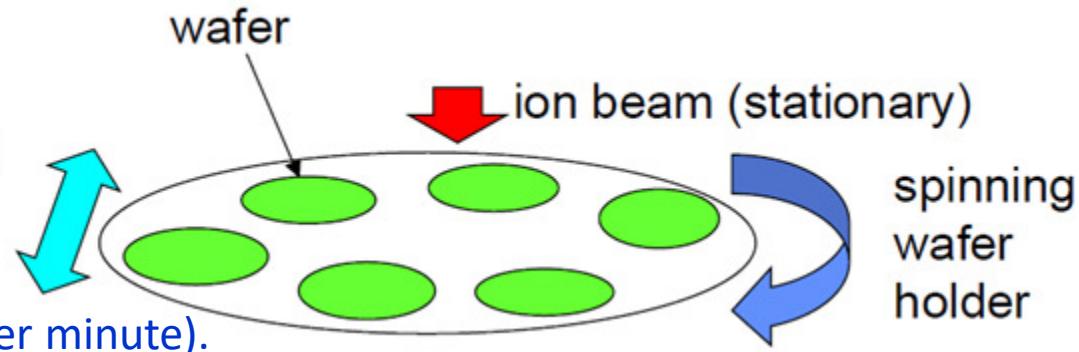
Translational wafer holder motion.

Rotating disk, 1000rpm (rotations per minute).
Stationary ion beam, beam diameter ~5cm.

\$3-4M/implanter
~60 wafers/hour

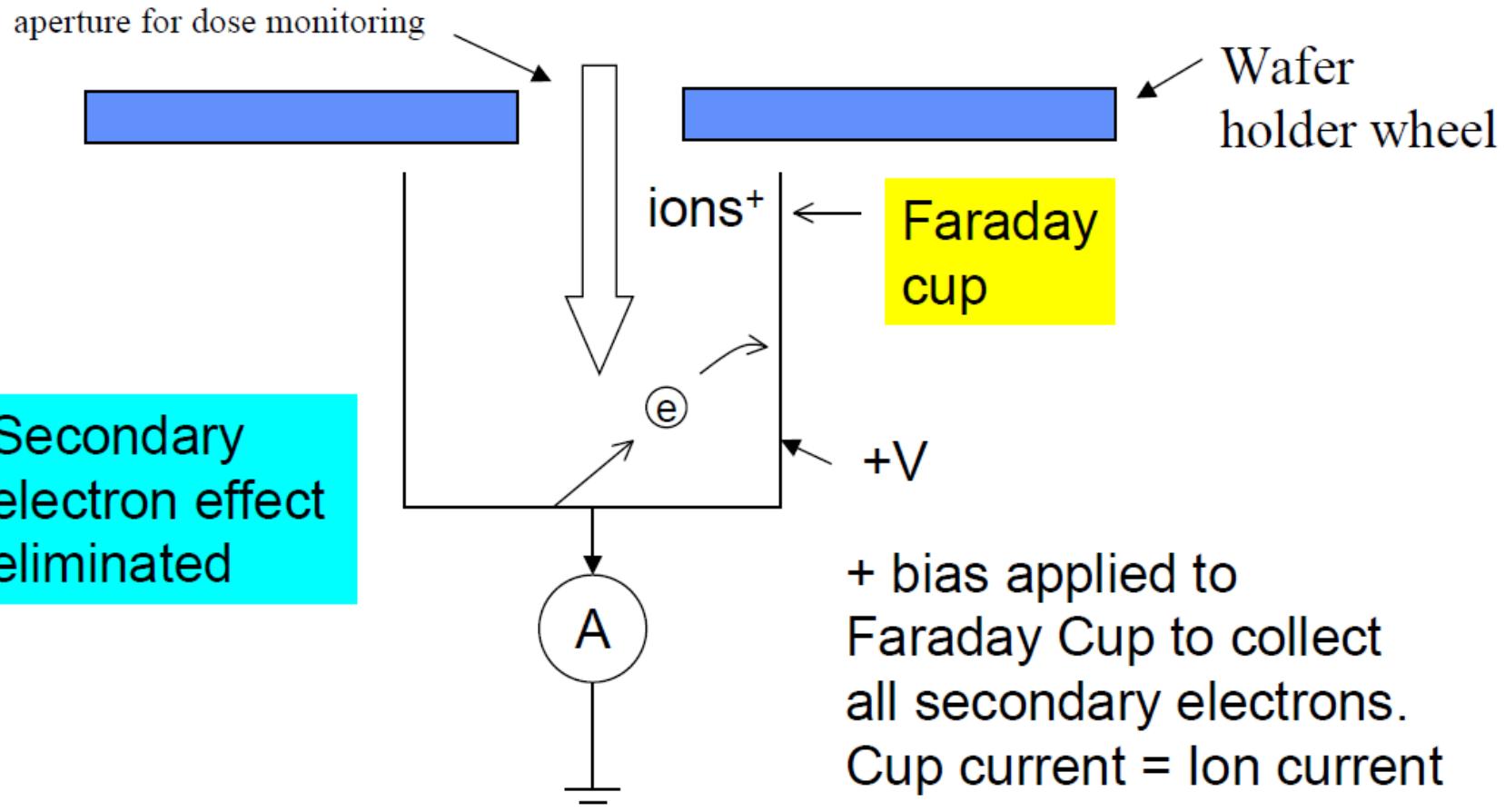


Eaton HE3 High Energy Implanter, showing the ion beam hitting the 300mm wafer end-station.



Practical implantation dosimetry (dose measurement)

The implant dose ϕ is the number of ions implanted per unit area (cm^2) of the wafer. If a beam current I is scanned for a time t , the total implanted charge $Q = (I \times t)$.



(Charge collected by integrating cup current) / (cup area) = dose

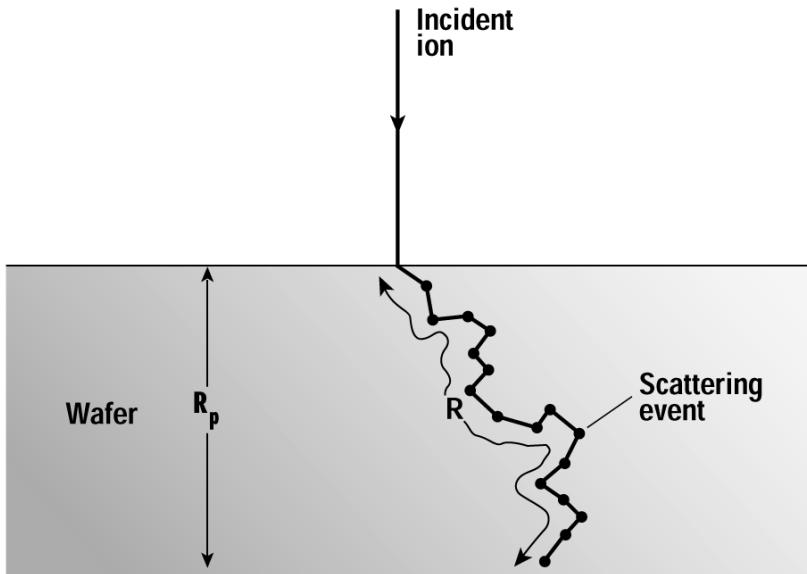
Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution profile.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

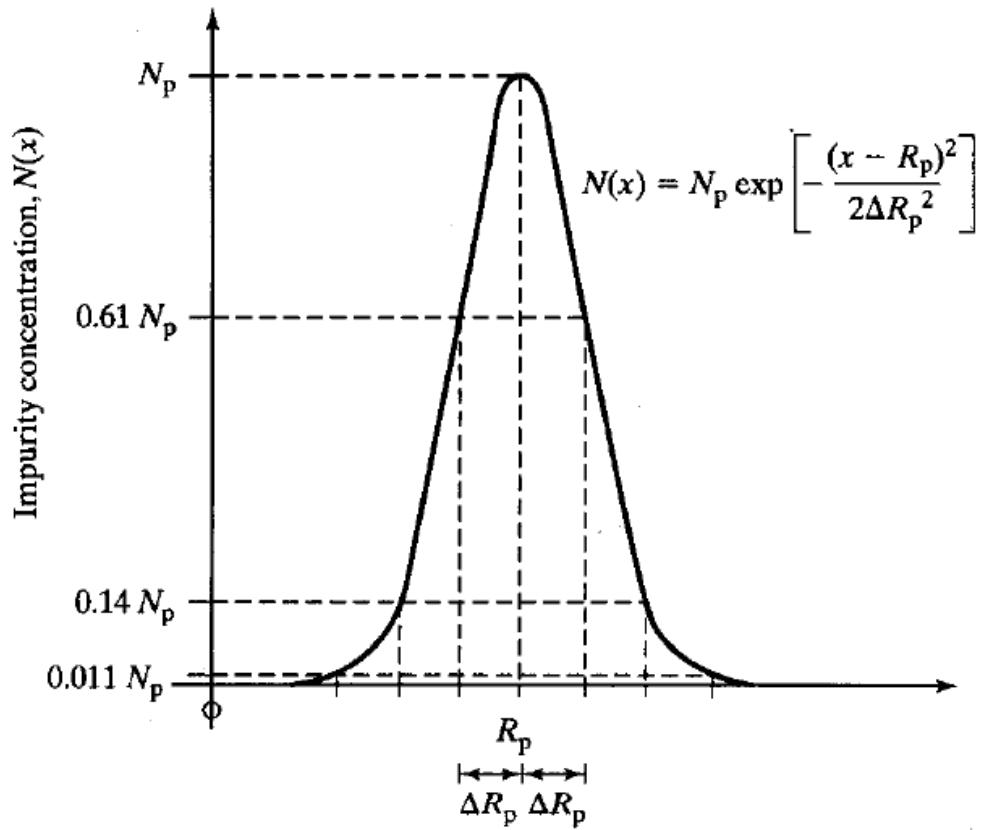
Dopant (impurity) concentration profile

- Gaussian distribution for first order approximation.
- R_p = projected range, is a function of ion energy and mass, and atomic number of impurity as well as target material.
- ΔR_p = straggle = standard deviation.
- N_p = peak concentration at $x=R_p$.
- Dose $Q=\int N(x)dx=(2\pi)^{1/2}N_p \Delta R_p$.

In textbook, C is used for concentration, to replace N used here.

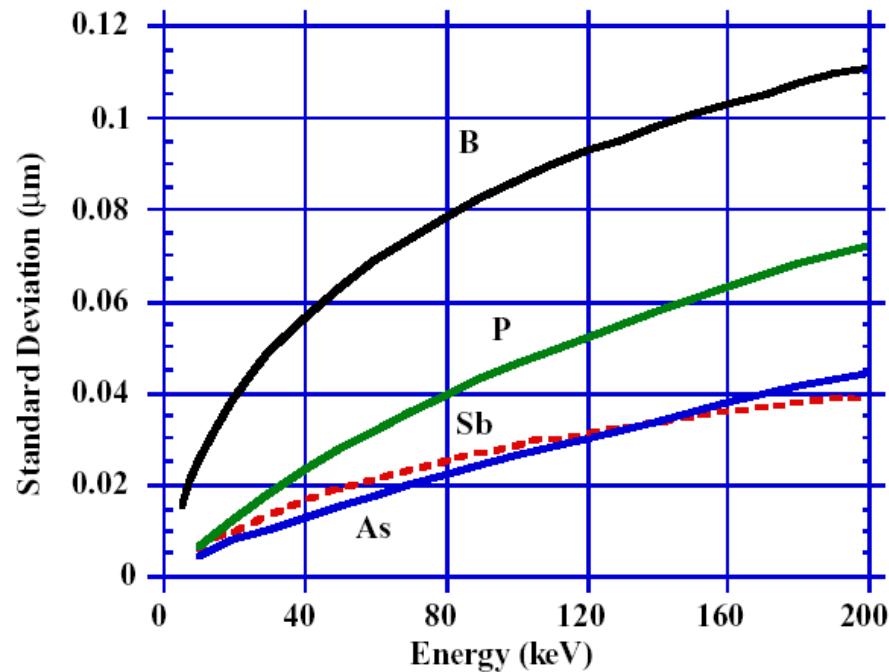
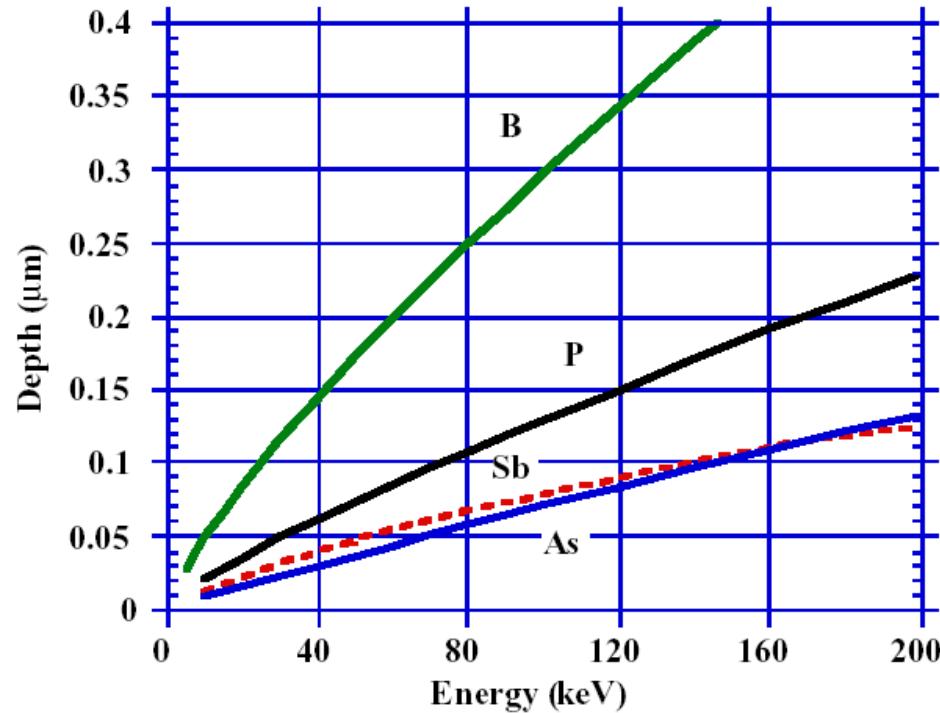


At wafer surface, $x=0$.



$$N(x) = N_p e^{-\frac{(x-R_p)^2}{2\Delta R_p^2}} = \frac{Q}{\sqrt{2\pi}\Delta R_p} e^{-\frac{(x-R_p)^2}{2\Delta R_p^2}}$$

Projection range (depth) and straggle (standard deviation)



$$\Delta R_p \cong \frac{2}{3} R_p \left[\frac{\sqrt{M_i M_t}}{M_i + M_t} \right]$$

M_i and M_t are the masses for incident ion and target atom.

Example

Example: Phosphorus with energy of 100keV is implanted into a silicon wafer

- A) What should the implanted dose be if peak concentration of $1 \times 10^{17} \text{ cm}^{-3}$ is desired.
- For 100Kev from figure: $R_p = 0.12 \mu\text{m}$ $\Delta R_p = 0.045 \mu\text{m}$

$$Q = \sqrt{2\pi} N_p \Delta R_p = \sqrt{2\pi} (1 \times 10^{17} \text{ cm}^{-3}) (0.45 \times 10^{-6} \text{ cm}) = 1.13 \times 10^{12} \text{ cm}^{-2}$$

- B) What length of time is required to implant this dose into a 200mm wafer using $2 \mu\text{A}$ beam current with single ionized phosphorus.

$$T = \frac{mqA Q}{I} = \frac{1 \times 1.6 \times 10^{-19} \text{ coul} (\pi) (10 \text{ cm})^2 (1.13 \times 10^{-12} \text{ cm}^{-2})}{2 \times 10^{-6} \text{ coul/sec}} = 24 \text{ sec}$$

Junction depth in Si

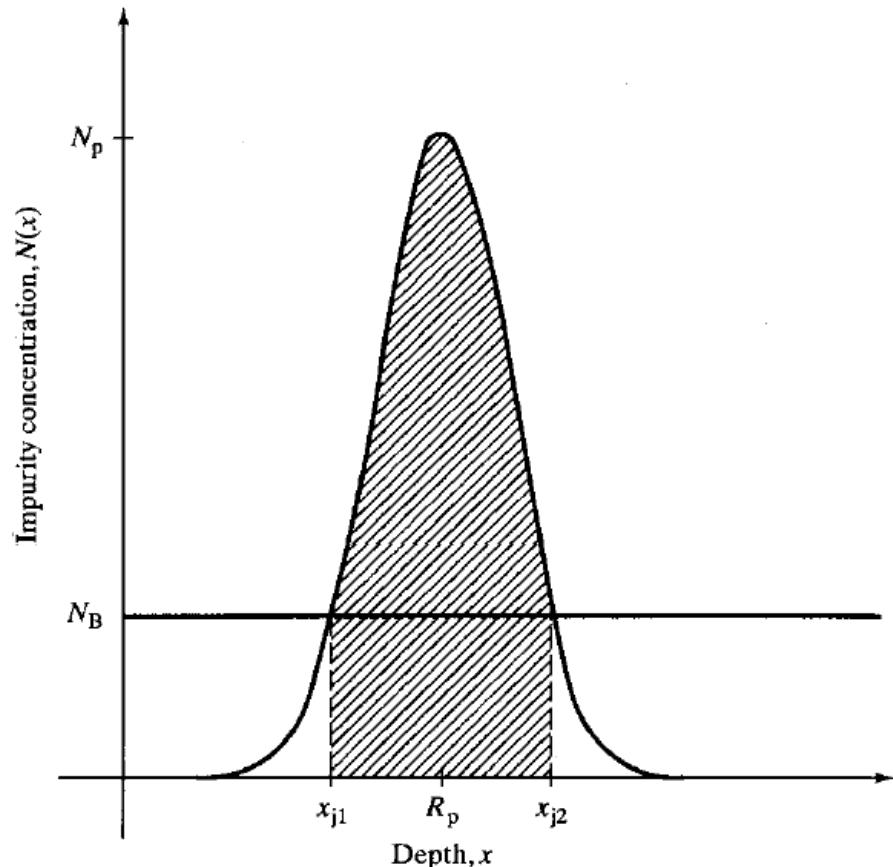
Implant into Si already doped at N_B .
 E.g. implant P into B-doped Si.
 P: n-type doping; B: p-type.

$$N(x) = N_p e^{-\frac{(x_j - R_p)^2}{2\Delta R_p^2}} = N_B$$

$$\frac{N_B}{N_p} = e^{-\frac{(x_j - R_p)^2}{2\Delta R_p^2}}$$

$$\ln \frac{N_p}{N_B} = \frac{(x_j - R_p)^2}{2\Delta R_p^2}$$

$$x_j = R_p \pm \Delta R_p \sqrt{2 \ln \frac{N_p}{N_B}}$$



Junction formation by impurity implantation. Two pn junctions are formed at x_{j1} and x_{j2} .

Example calculations

200 keV Phosphorus is implanted into a p-Si ($C_B = 10^{16}/\text{cm}^3$) with a dose of $10^{13}/\text{cm}^2$.

From graphs or tables , $R_p = 0.254 \mu\text{m}$, $\Delta R_p = 0.0775 \mu\text{m}$

(a) Find peak concentration

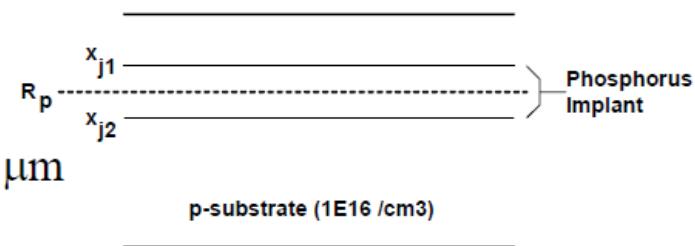
$$C_p = (0.4 \times 10^{13}) / (0.0775 \times 10^{-4}) = 5.2 \times 10^{17}/\text{cm}^3$$

(b) Find junction depths

$$(b) C_p \exp[-(x_j - 0.254)^2 / 2 \Delta R_p^2] = C_B \quad \text{with } x_j \text{ in } \mu\text{m}$$

$$\therefore (x_j - 0.254)^2 = 2 \times (0.0775)^2 \ln [5.2 \times 10^{17} / 10^{16}]$$

$$\text{or } x_j = 0.254 \pm 0.22 \mu\text{m} ; x_{j1} = 0.032 \mu\text{m} \text{ and } x_{j2} = 0.474 \mu\text{m}$$



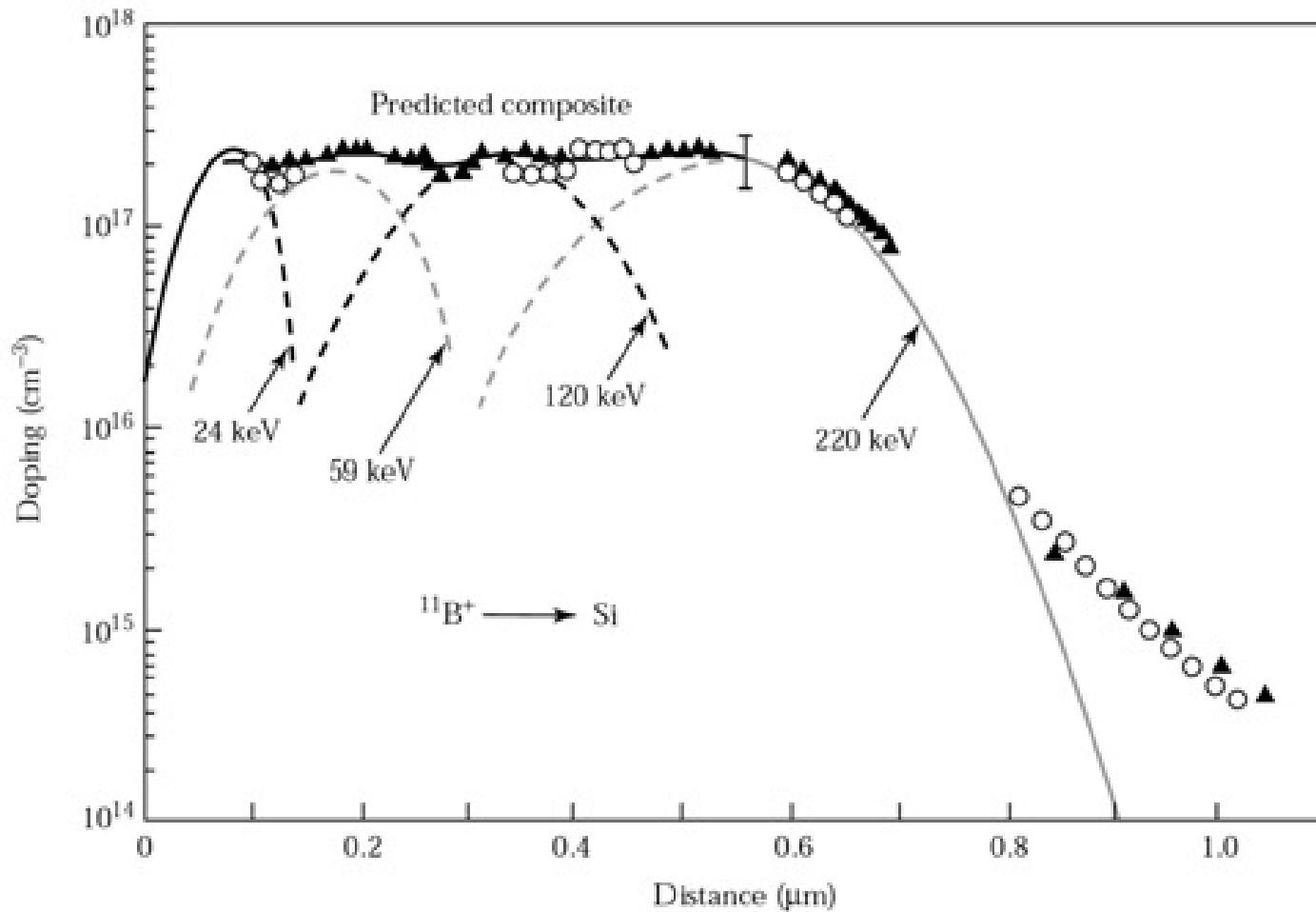
(c) Find sheet resistance

From the mobility curve for electrons (using peak conc as impurity conc), $\mu_n = 350 \text{ cm}^2/\text{V}\cdot\text{sec}$

$$R_s = \frac{1}{q\mu_n \phi} = \frac{1}{1.6 \times 10^{-19} \times 350 \times 10^{13}} \approx 1780 \Omega/\text{square.}$$

Composite doping profile using multiples implants.

In many application doping profiles other than the simple Gaussian are required.



Diffusion during subsequent anneals

During high temperature steps after implant (most commonly an activation anneal), the implanted impurities will begin to diffuse, broadening the implantation profile.

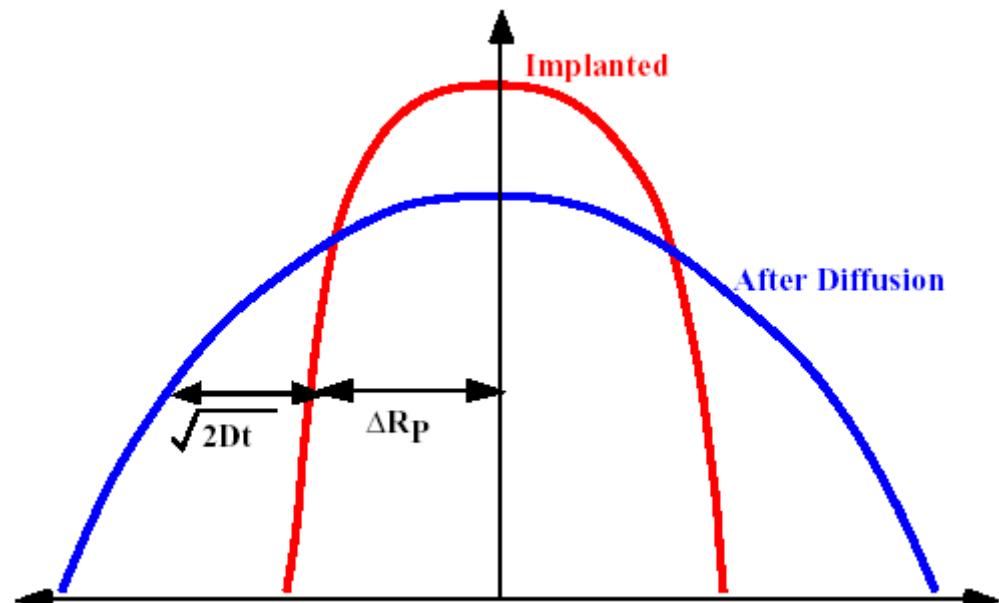
For implantations far away from the surface and for reasonable short characteristic diffusion lengths, the new profile can be approximated by:

$$C(x, t) = \frac{Q}{\sqrt{2\pi(\Delta R_p^2 + 2Dt)}} \exp\left[-\frac{(x - R_p)^2}{2(\Delta R_p^2 + 2Dt)}\right]$$

$$C(x, t) = \frac{Q}{2\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right)$$

$$C(x) = \frac{Q}{\sqrt{2\pi} \Delta R_p} \exp\left[-\left(\frac{x - R_p}{\sqrt{2\Delta R_p}}\right)^2\right]$$

$$\frac{(\Delta R_p)^2}{2} = D_0 t_0$$



$$Dt \Leftrightarrow D_0 t_0 + Dt$$

Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution profile.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

How thick mask is needed?

In order to act as an efficient mask, the thickness of the mask should be large enough that the tail of the implant profile in the silicon should not significantly alter the doping concentration (C_B).

$$C^*(x_m) = C_P^* \left[\exp - \frac{(x_m - R_p^*)^2}{2\Delta R_p^{*2}} \right] \leq C_B$$

("*" means for value in the mask layer)

Mask thickness x_m

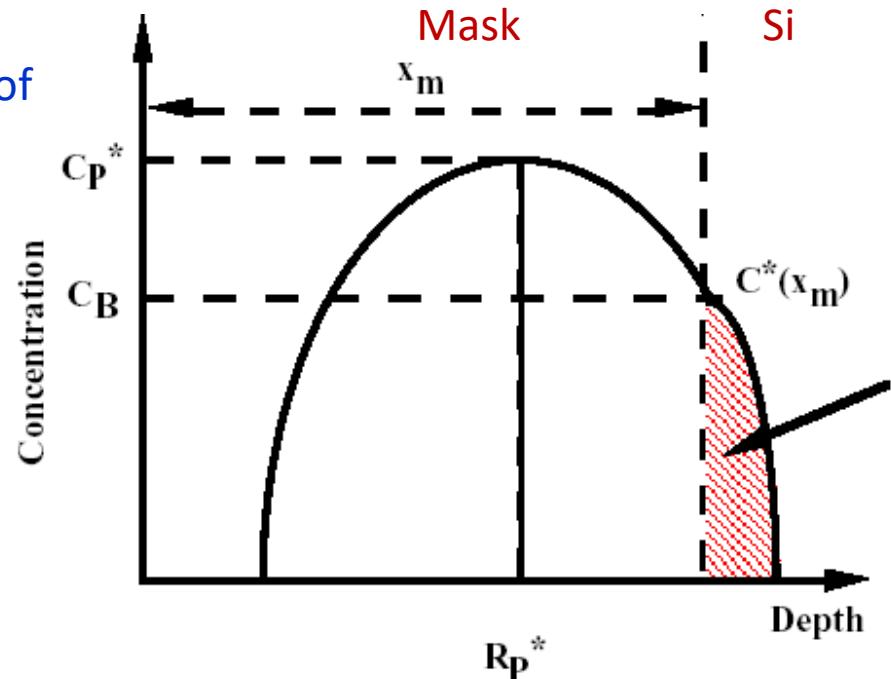
$$x_m = R_p^* + \Delta R_p^* \sqrt{2 \ln \left(\frac{C_P^*}{C_B} \right)} = R_p^* + m \Delta R_p^*$$

Dose penetrating the mask:

$$Q_p = \frac{Q}{\sqrt{2\pi}\Delta R_p^*} \int_{x_m}^{\infty} \exp \left[-\frac{1}{2} \left(\frac{x - R_p^*}{\Delta R_p^*} \right)^2 \right] dx = \frac{Q}{2} \operatorname{erfc} \left(\frac{x_m - R_p^*}{\sqrt{2}\Delta R_p^*} \right)$$

Rule of thumb: good masking thickness

$$X_m = R_p + 4.3 \Delta R_p, C(x=X_m)/C(x=R_p) \sim 10^{-4}.$$



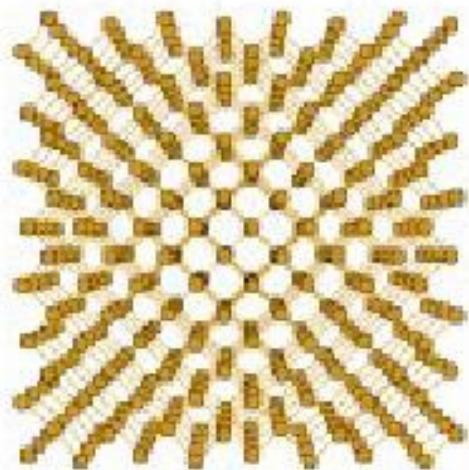
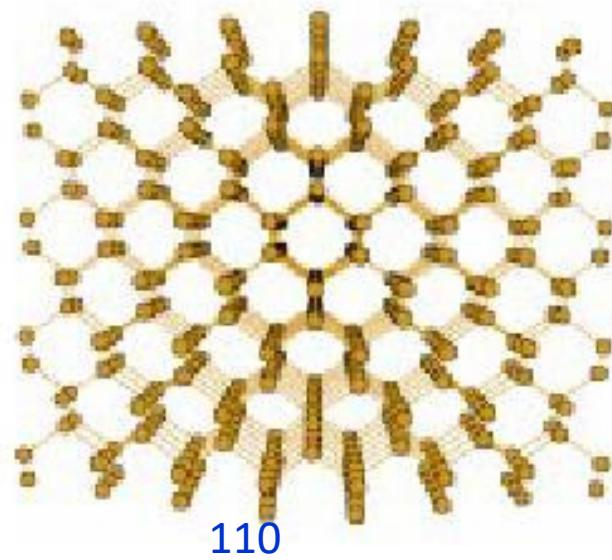
Mask material thickness

MATERIAL THICKNESS NEEDED TO MASK					
At 200 KeV	Poly	SiO2	Si3N4	Al	Resist
Boron	0.9μm	1.0μm	0.61μm	0.9μm	1.0μm
Phosphorous	0.7μm	0.6μm	0.42μm	0.55μm	0.8μm
Arsenic	0.3μm	0.3μm	0.18μm	0.28μm	0.35μm
Antimony	0.2μm	0.2μm	0.16μm	0.18μm	0.25μm
At 100 KeV	Poly	SiO2	Si3N4	Al	Resist
Boron	0.65μm	0.7μm	0.42μm	0.7μm	0.7μm
Phosphorous	0.4μm	0.36μm	0.25μm	0.3μm	0.45μm
Arsenic	0.18μm	0.16μm	0.1μm	0.16μm	0.20μm
Antimony	0.12μm	0.11μm	0.07μm	0.10μm	0.14μm
FOR 0.0001% TRANSMISSION					

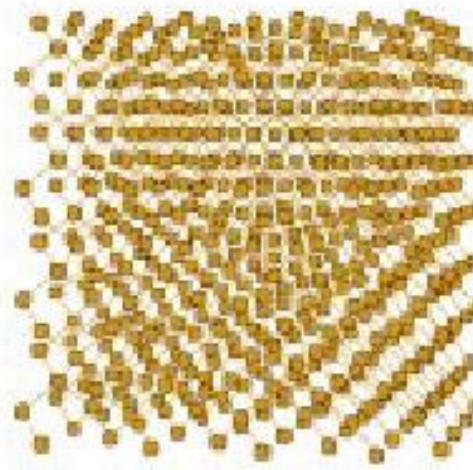
Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution profile.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

Looking at Si at different orientations

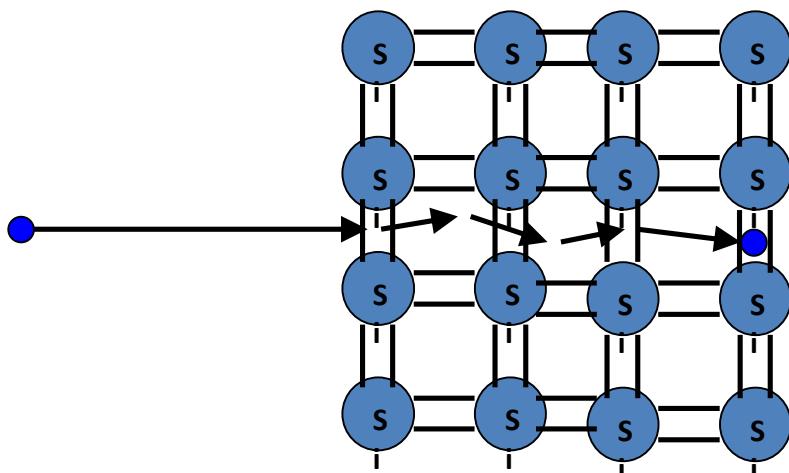


100

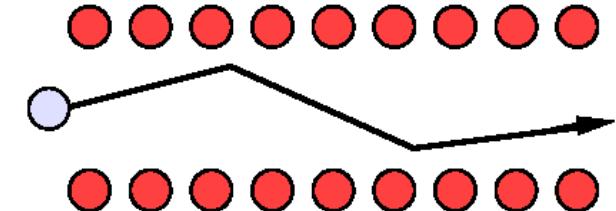
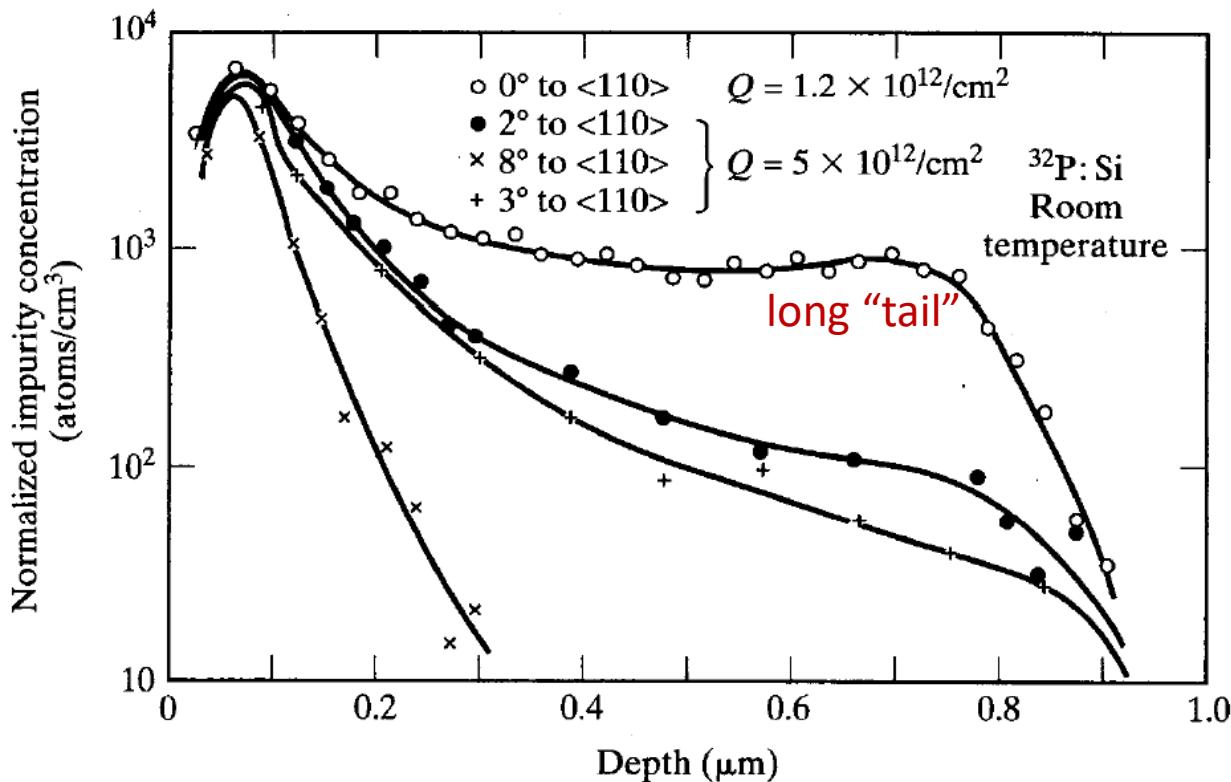


Random tilt and rotation

Channeling effect



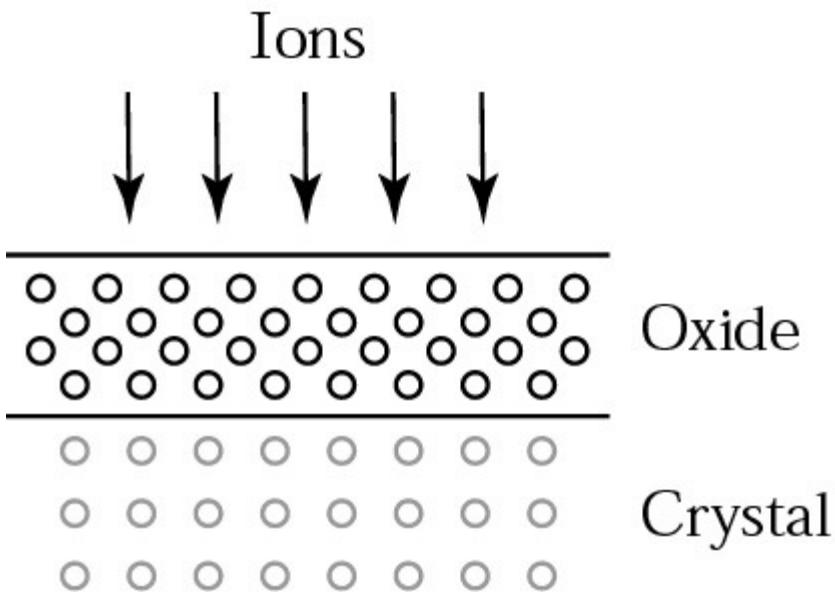
- Occur when ion velocity is parallel to a major crystal orientation.
- Some ions may travel considerable distances with little energy loss.
- Once in a channel, ion will continue in that direction, making many glancing internal collisions that are nearly elastic (their stopping is then dominated by electronic drag only), until it comes to rest or finally dechannels.



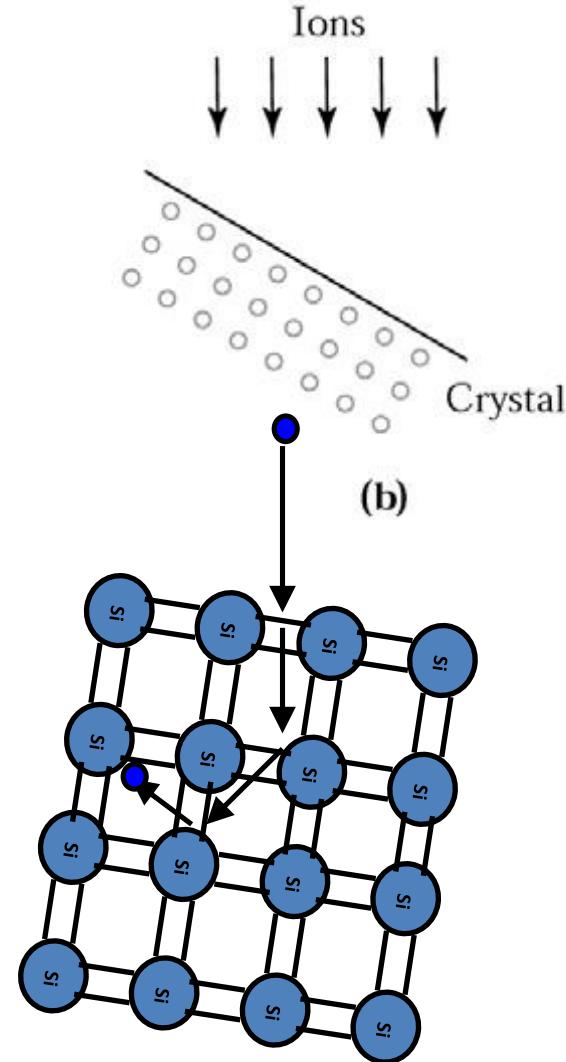
Phosphorus impurity profiles for 40keV ion implantations to silicon at various angles from the <110> axis.

Methods to reduce channeling effect

A thin screen oxide which is amorphous is often used, causing some randomization of incident beam before it enters the lattice.

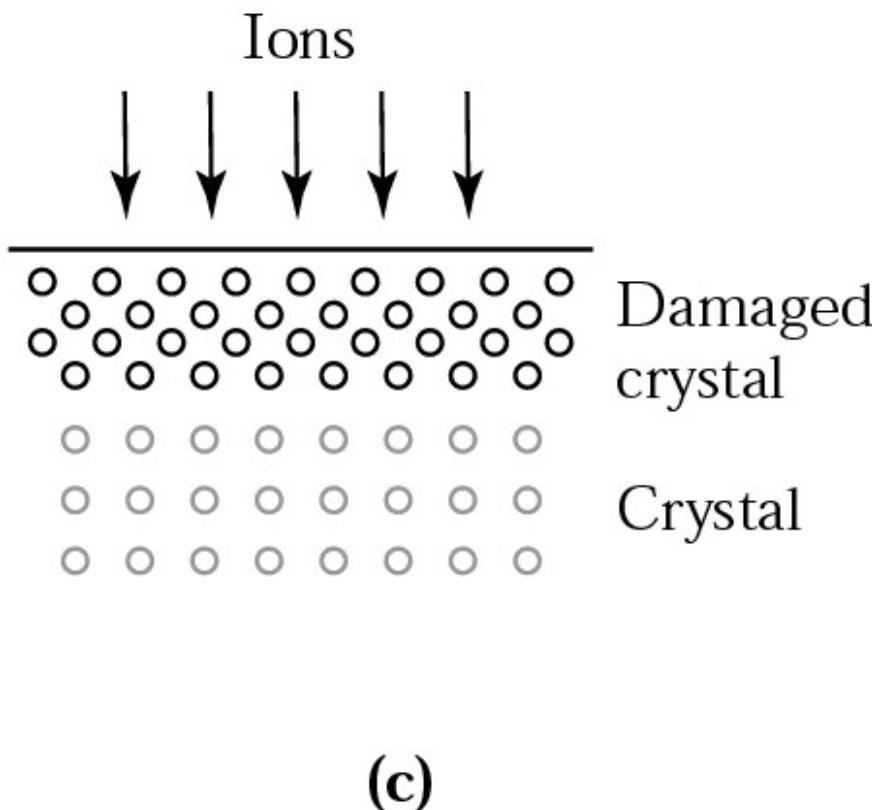


Most IC implantation is done off axis. A typical tilt angle is 7° .

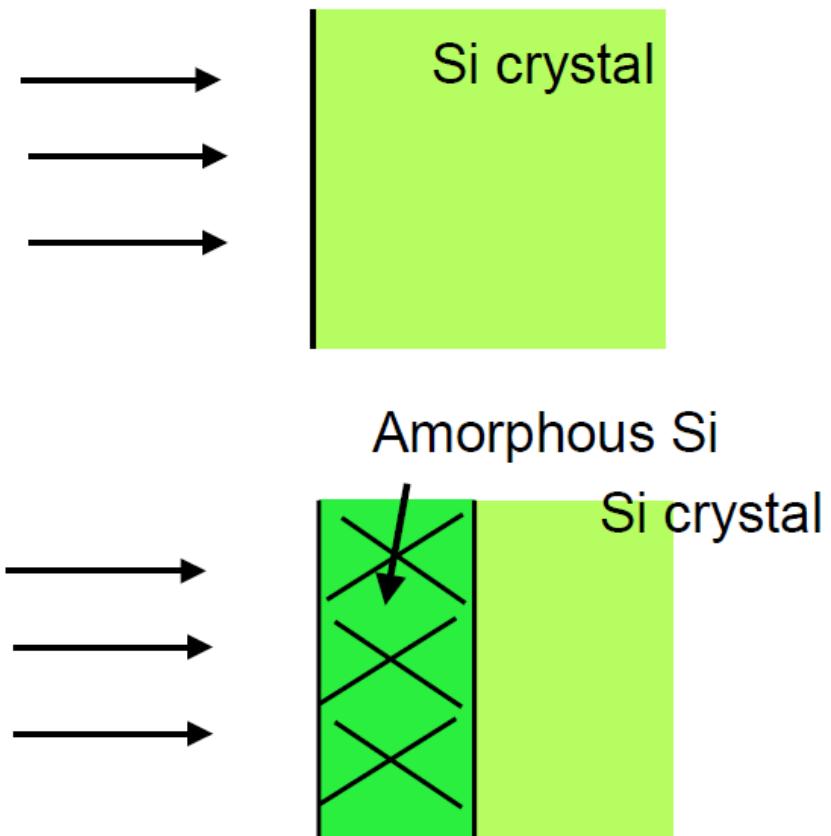


Methods to reduce channeling effect

Destroy the lattice before implantation



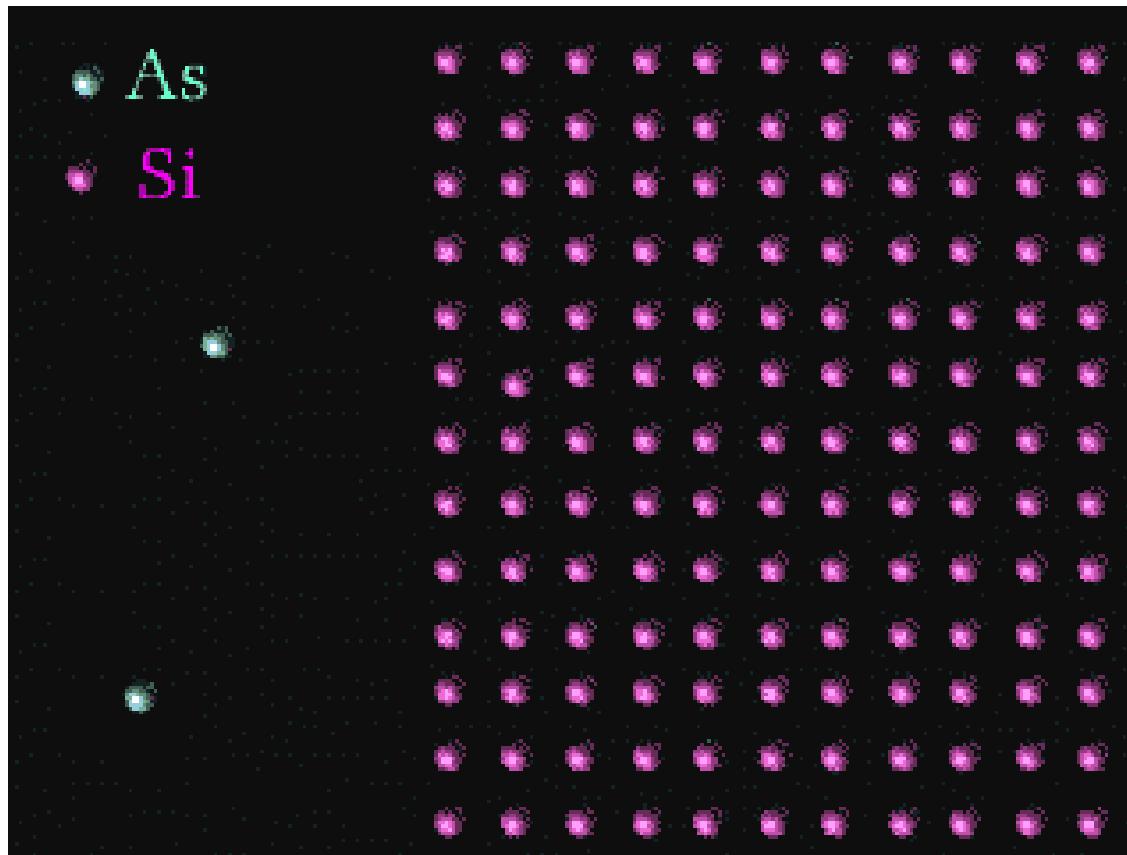
1. High dose Si^+ implantation to convert the surface layer into amorphous Si.
2. Implantation of desired dopant



Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution profile.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

Ion – substrate interaction



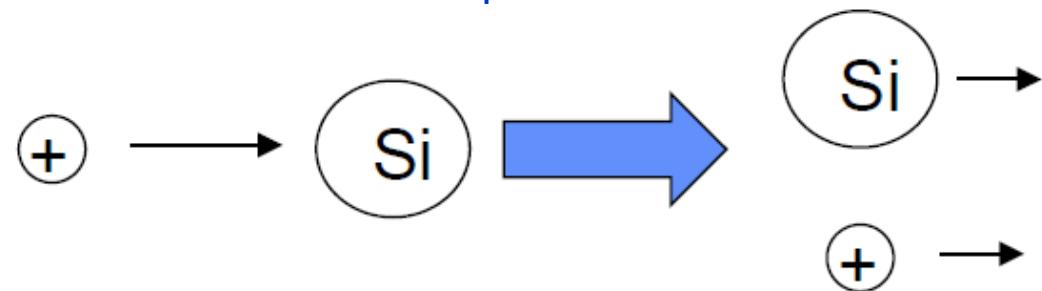
The ions are stopped at random positions, mostly not in crystalline sites, so not active as dopant (need anneal to activate them).

Ion implantation energy loss mechanisms

LSS theory: in 1963, Lindhard, Scharff and Schiott proposed that the energy loss of incident ion can be divided into two independent process, namely nuclear stopping and electronic stopping. Total energy loss is the sum of the two processes.

Nuclear stopping:

Nuclear energy loss tends to dominant towards the end of the range when ion has lost much of its energy, and this is where nuclear collisions produced most of the damage. Maximum when $m_1=m_2$.

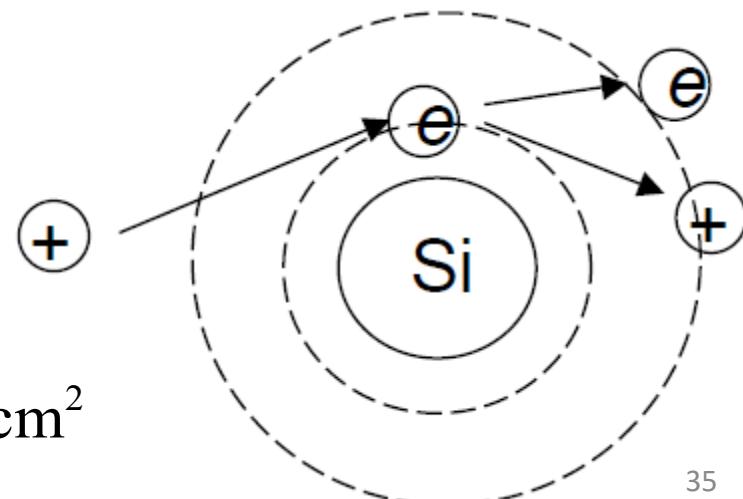


$$S_n(E) = 2.8 \times 10^{-15} \frac{Z_1 Z_2}{\sqrt{Z_1^{2/3} + Z_2^{2/3}}} \frac{m_1}{m_1 + m_2} \text{ eV cm}^2$$

Electronic stopping:

Because electrons can follow fields up to optical frequencies (velocities of 10^5 m/s - 100 times faster than phonons), electronic losses dominate at higher ion velocities.

$$S_e(E) = Cv_{ion} = kE^{1/2}, k \approx 0.2 \times 10^{15} \text{ eV}^{1/2} \text{cm}^2$$



Nuclear and electronic stopping power

Energy loss rate
(N is target atom density)

$$\frac{dE}{dx} = -N[S_n(E) + S_e(E)]$$

Nuclear stopping power
(unit: eV cm²)

$$S_n(E) = \frac{1}{N} \left(\frac{dE}{dx} \right)_n$$

Electronic stopping power

$$S_e(E) = \frac{1}{N} \left(\frac{dE}{dx} \right)_e$$

Ion range

$$R = \int_0^R dx = \frac{1}{N} \int_0^{E_0} \frac{dE}{S_n(E) + S_e(E)}$$

Lighter ions/at higher energy: more electronic stopping

Heavier ions/at lower energy: more nuclear stopping

H⁺: electronic stopping dominates

B⁺: electronic stopping dominates

As⁺: nuclear stopping dominates

Electronic and nuclear stopping power

	E1(keV)	E2(keV)
B into Si	3	17
P into Si	17	140
As into Si	73	800

E_2 : energy where electronic stopping power is equal to nuclear stopping power.

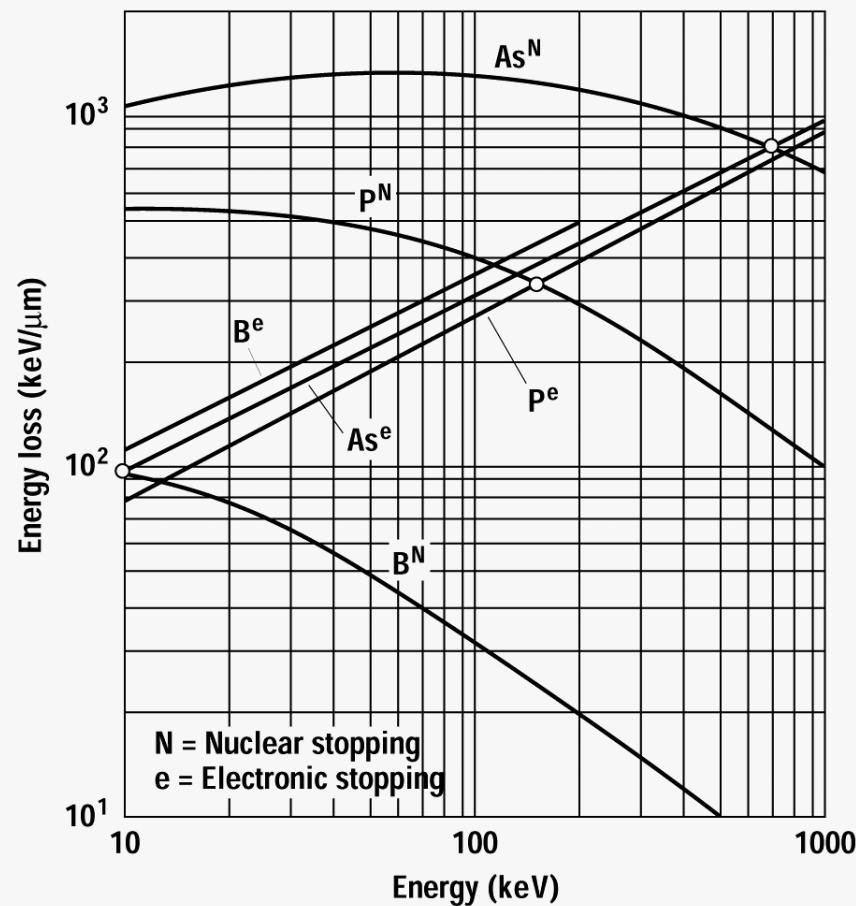
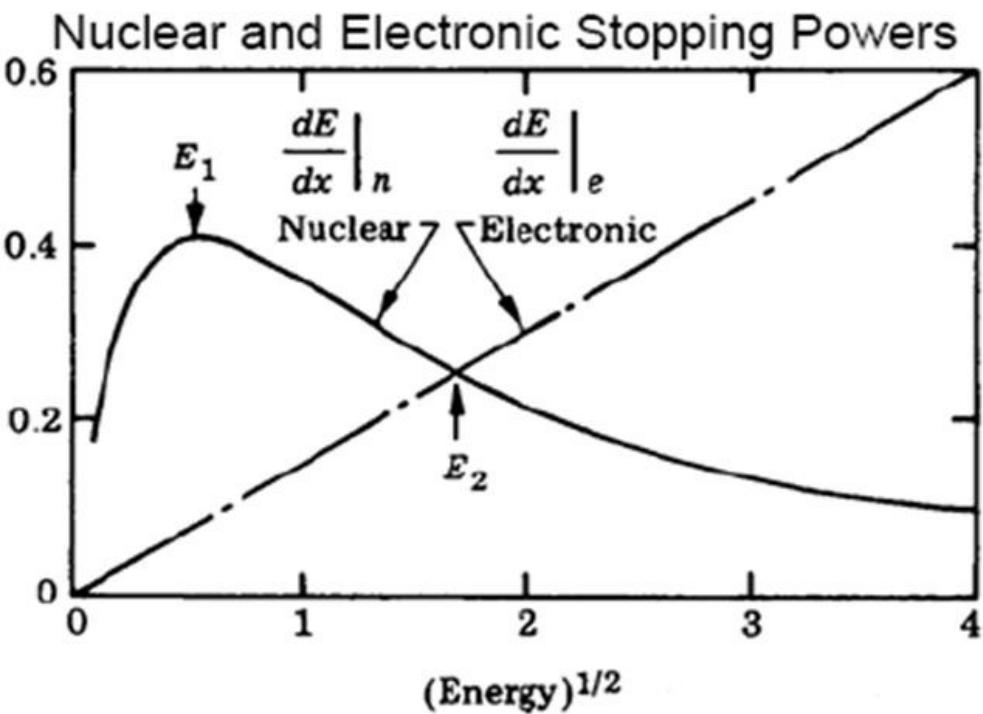


Figure 5.8 Nuclear and electronic components of $S(E)$ for several common silicon dopants as a function of energy (after Smith as redrawn by Seidel, "Ion Implantation," reproduced by permission, McGraw-Hill, 1983).

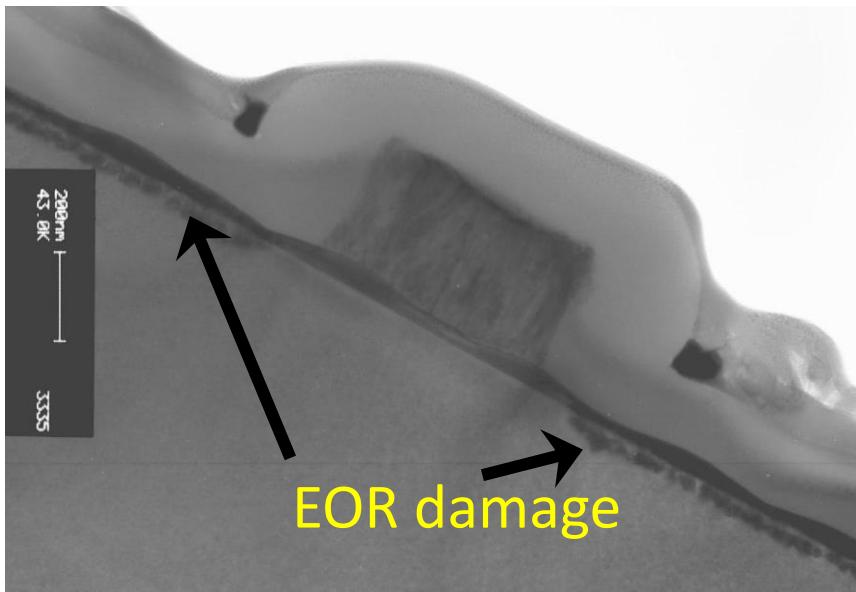
B^+ : light, electronic stopping dominates

As^+ : heavy, nuclear stopping dominates

Chapter 8 Ion implantation

1. Introduction and application.
2. Ion implantation tools.
3. Dopant distribution profile.
4. Mask thickness and lateral distribution.
5. Effect of channeling.
6. Modeling: nuclear and electronic stopping.
7. Damage caused by ion implantation and its repair.

Damage in ion implantation



- The threshold energy or displacement energy is E_d (for Si, $E_d \sim 15\text{eV}$).
- The number of displaced atoms created by an energetic particle is on the order of $N = E_n/E_d$, where E_n is the energy lost in nuclear collision (not electronic stopping).
- For example, 30keV As ion will lead to order of 1000 displaced atoms.
- The number will be less for 30keV light ions, whose energy is mainly lost by electronic stopping.

- ❑ The main disadvantage of ion implantation is the production of lattice damage (vacancies and interstitials, or V/I) which may evolve from simple point defects into complex dislocations or voids.
- ❑ Eventually, implantation leads to an amorphous silicon structure.
- ❑ Most damage can be repaired by annealing. However, annealing cycles of 30 min at close to 1000°C can cause considerable spreading of the implant by diffusion.

Amorphization

- If the dose is high enough, the implanted layer will become amorphous.
- The dose required to produce an amorphous silicon layer is called critical implant dose.
- The heavier the impurity, the lower the dose that is required to create an amorphous layer.

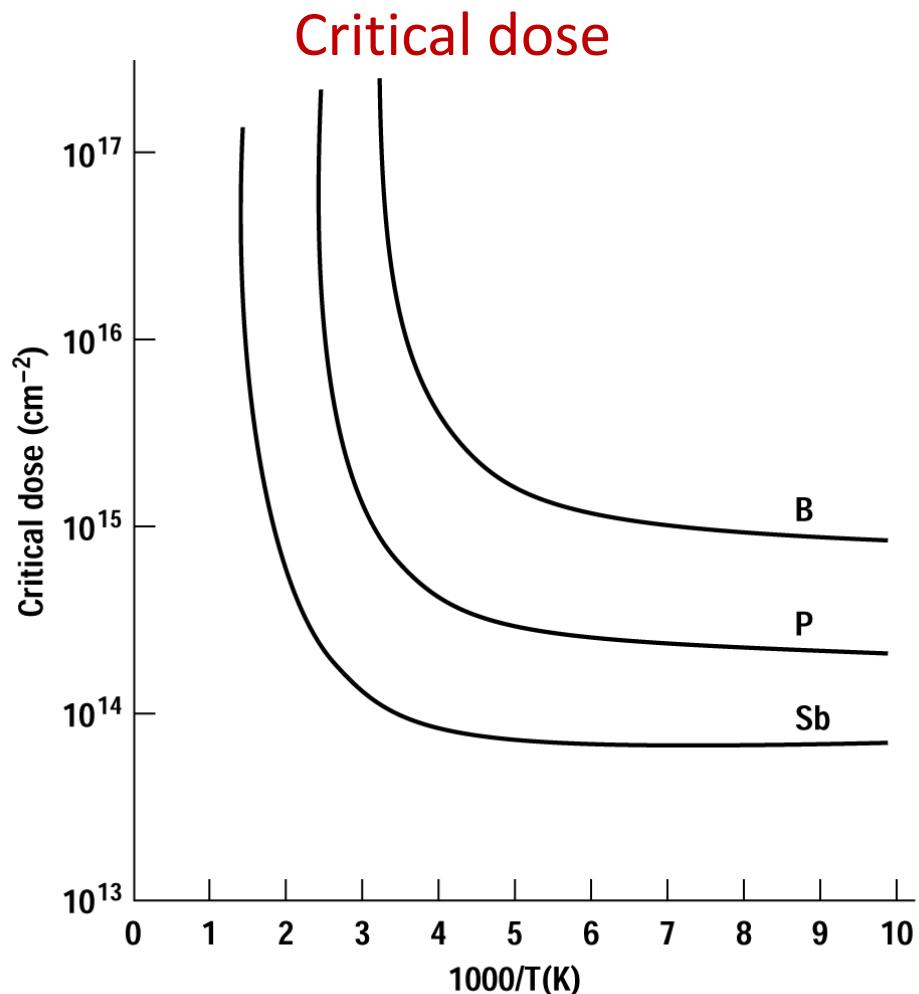


Figure 5.13 Critical implant dose required to amorphize a silicon substrate as a function of substrate temperature for several common silicon dopants (*after Morehead and Crowder*). 40

Annealing repair damage and activate dopants

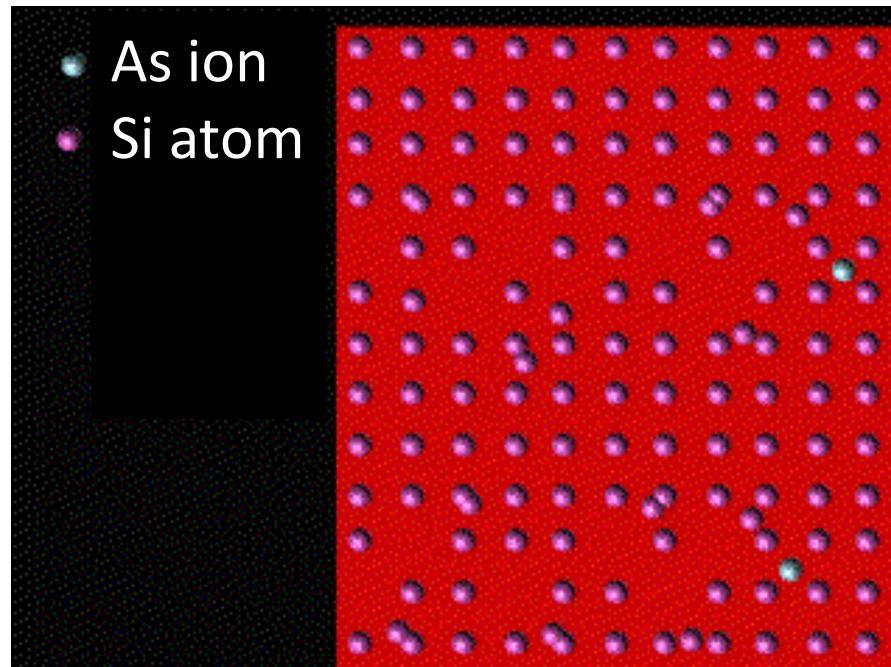
After implantation, we need an annealing step, usually under Ar, N₂ or vacuum.

A typical ~900°C, 30min will:

Restore silicon lattice to its perfect crystalline state - silicon atoms can move back into lattice sites at these temperatures.

Put dopants into Si substitution sites for electrical activation - nearly all of the implanted dose becomes electrically active except for impurity concentrations exceeding 10¹⁹/cm³.

Restore the electron and hole mobility – now that the lattice becomes perfect again.



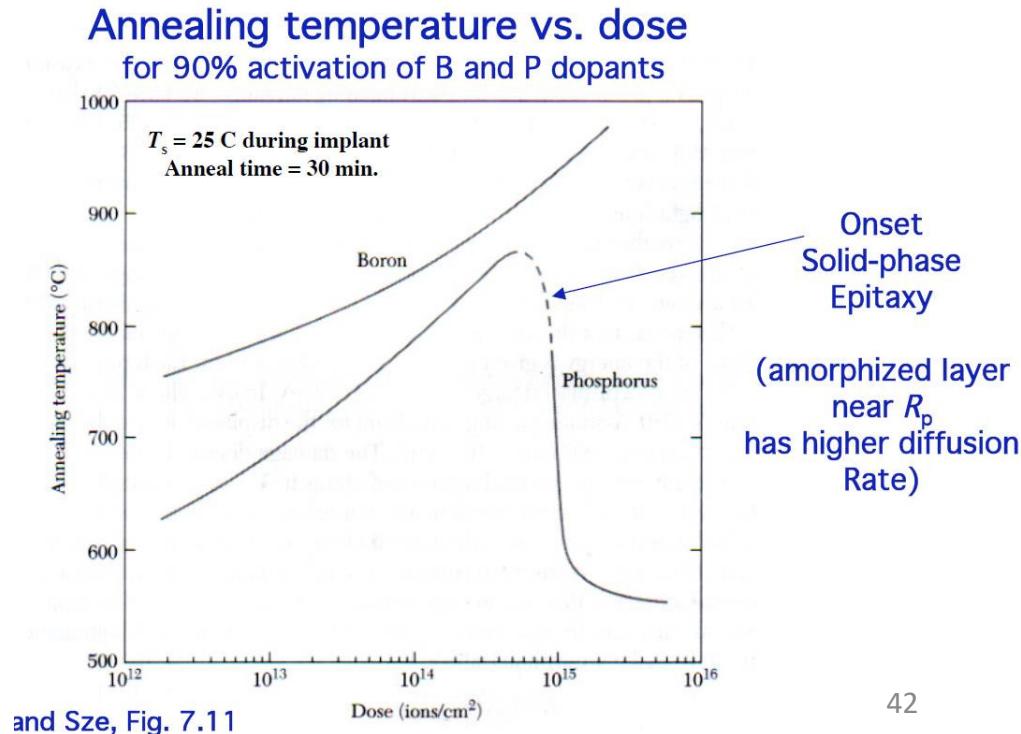
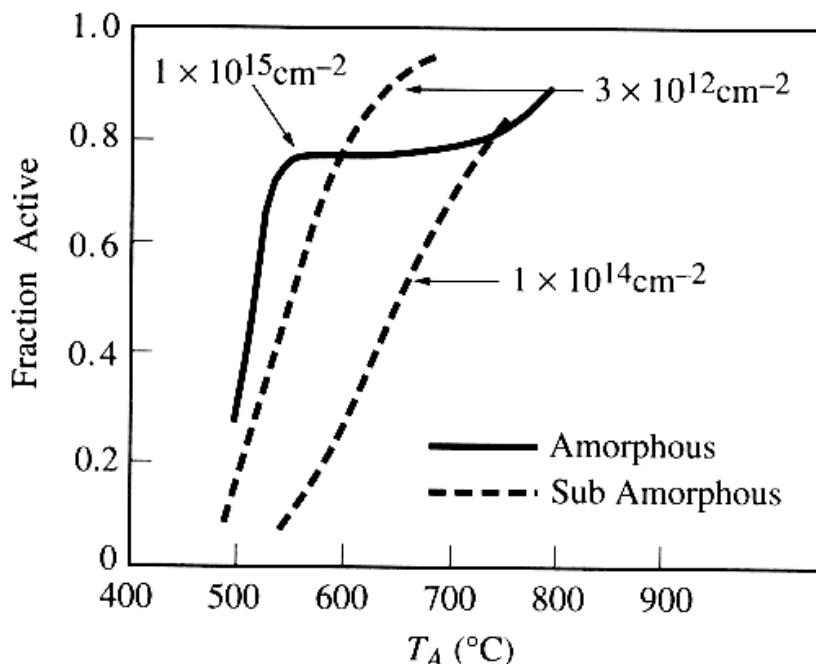
Solid phase epitaxy

Solid phase epitaxy (SPE): when substrate has been rendered amorphous, the crystallinity is repaired by SPE, where crystal reforms using the underlying undamaged substrate as a template. Most of impurities are incorporated into the growing lattice .

Due to the high activation energies required to annihilate defects ($\sim 5\text{eV}$), it is often easier to regrow the crystal from an amorphous layer via SPE (activation energy $\sim 2.3\text{eV}$ in Silicon) than it is to anneal out defects. Thus, two schemes for implants are used:

Implant above the critical dose and use low temperature anneal to regrow material.

Implant below the critical dose and use high temperature anneal to get rid of defects.



Rapid thermal processing/annealing

Dopants can diffuse during high temperature anneal (activation energy ~3-4eV)

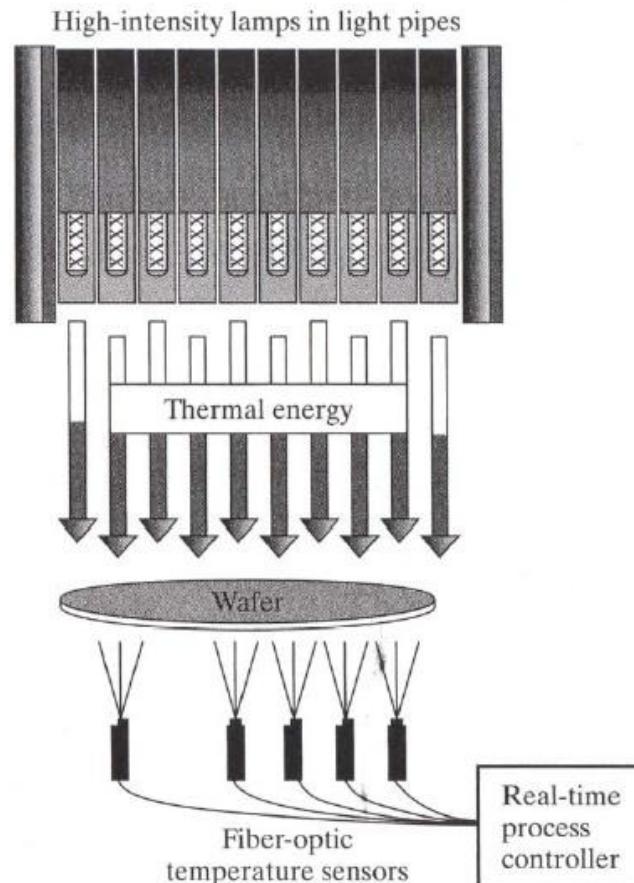
To minimize this unwanted diffusion, one can use Rapid Thermal Processing (RTP) or Rapid Thermal Anneal (RTA).

RTA is extremely important for shallow junction devices.

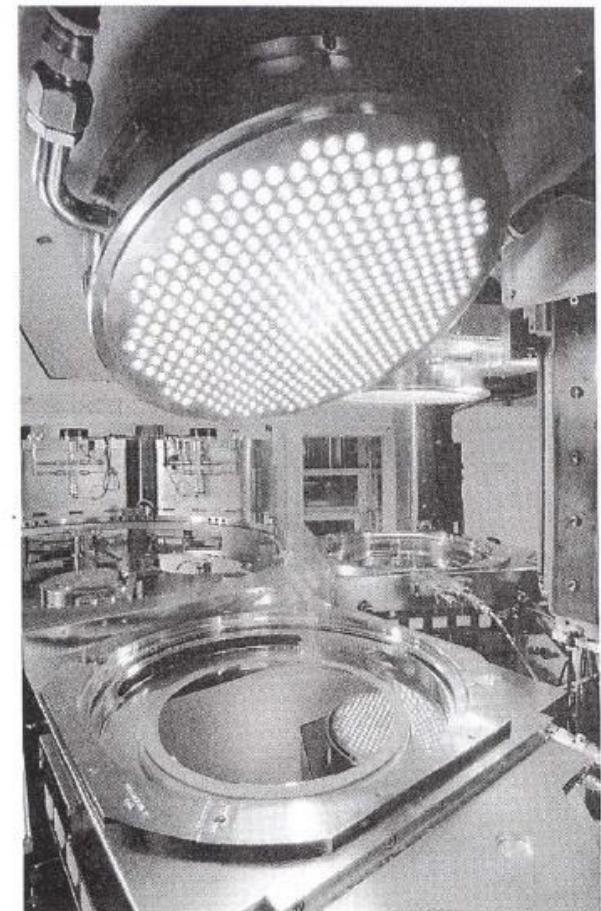
Rapid heating source:

- high power laser
- electron beam
- high intensity halogen lamp

Applied Materials 300mm RTP System



(a)



(b)

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

Thin film

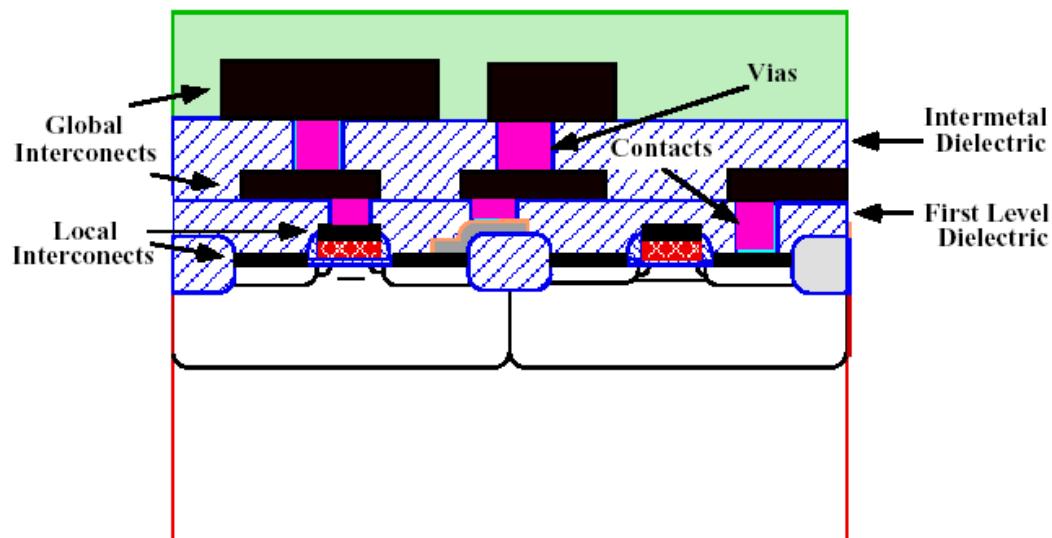
Thin film: thickness typically <1000nm.

Special properties of thin films: different from bulk materials, it may be –

- Not fully dense
- Under stress
- Quasi - two dimensional (very thin films)
- Strongly influenced by surface and interface effects

Typical steps in making thin films:

1. Emission of particles from source (heat, high voltage . . .)
2. Transport of particles to substrate
3. Condensation of particles on substrate



Lithography, thin film deposition and its etching are the three most important processes for micro-nano fabrication.

Thin film deposition methods

Two main deposition methods are used today:

Chemical Vapor Deposition (CVD)

Reactant gases introduced in the chamber, chemical reactions occur on wafer surface leading to the deposition of a solid film.

E.g. APCVD, LPCVD, PECVD, most commonly used for dielectrics and Si.

Physical Vapor Deposition (PVD) (no chemical reaction involved)

Vapors of constituent materials created inside the chamber, and condensation occurs on wafer surface leading to the deposition of a solid film.

E.g. evaporation, sputter deposition, most commonly used for metals.

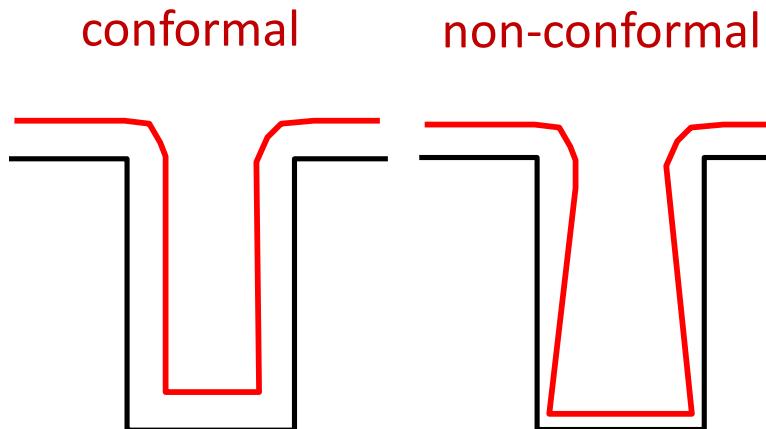
Other methods that are gaining importance in ULSI fabrication:

1. Coating with a liquid that becomes solid upon heating, e.g. spin-on-glass used for planarization.
2. Electro-deposition: coating from a solution that contains ions of the species to be coated. E.g. Cu electroplating for global interconnects.
3. Thermal oxidation.

General characteristics of thin film deposition

- Deposition rate
- Film uniformity:
 - Across wafer uniformity.
 - Run-to-run uniformity.
- Materials that can be deposited: metal, dielectric, polymer.
- Quality of film:
 - Physical and chemical properties
 - Electrical property, breakdown voltage
 - Mechanical properties, stress and adhesion to substrate
 - Optical properties, transparency, refractive index
 - Composition, stoichiometry
 - Film density, defect (pinhole...) density
 - Texture, grain size, boundary property, and orientation
 - Impurity level, doping
- Deposition directionality:
 - Directional - good for lift-off, trench filling
 - Non-directional - good for conformal step coverage
- Cost of ownership and operation.

Step coverage



Poor (non-conformal) step coverage is good for liftoff.
Conformal film is good for electrical connection...

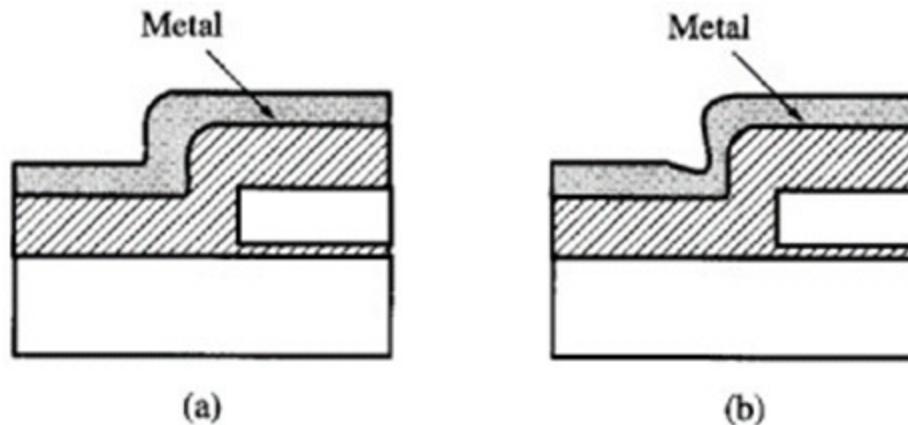


Figure 9-1 Step coverage of metal over non-planar topography.

- (a) Conformal step coverage, with constant thickness on horizontal and vertical surfaces.
- (b) Poor step coverage, here thinner for vertical surfaces.

Thin film filling of holes/trenches

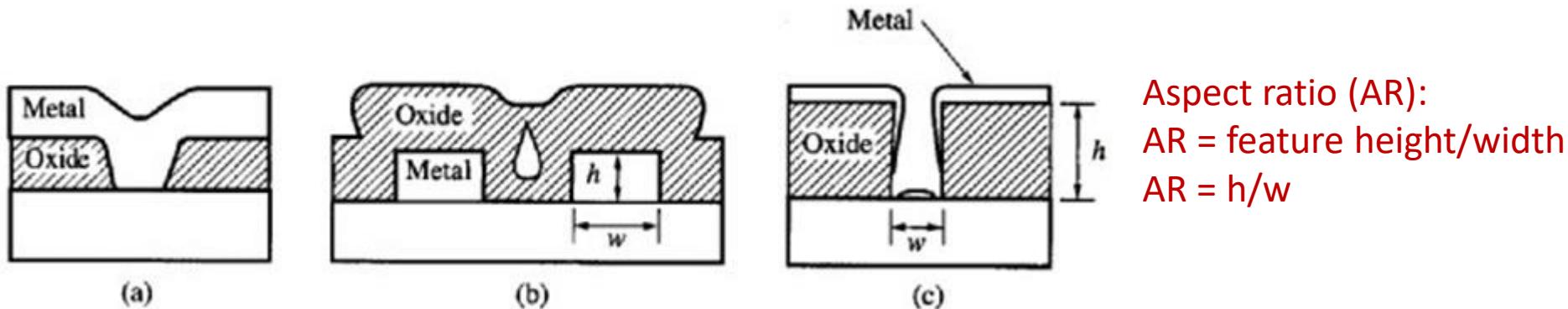


Figure 9-2 Thin film filling issues.

- (a) Good metal filling of a via or contact in a dielectric layer.
- (b) Silicon dioxide filling of the space between metal lines, with poor filling leading to void formation.
- (c) Poor filling of the bottom of a via hole with a barrier or contact metal.

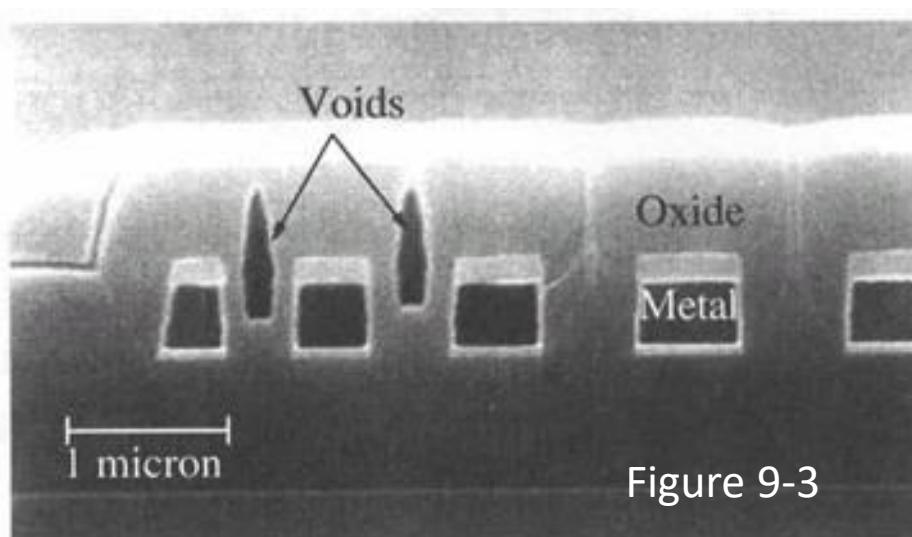
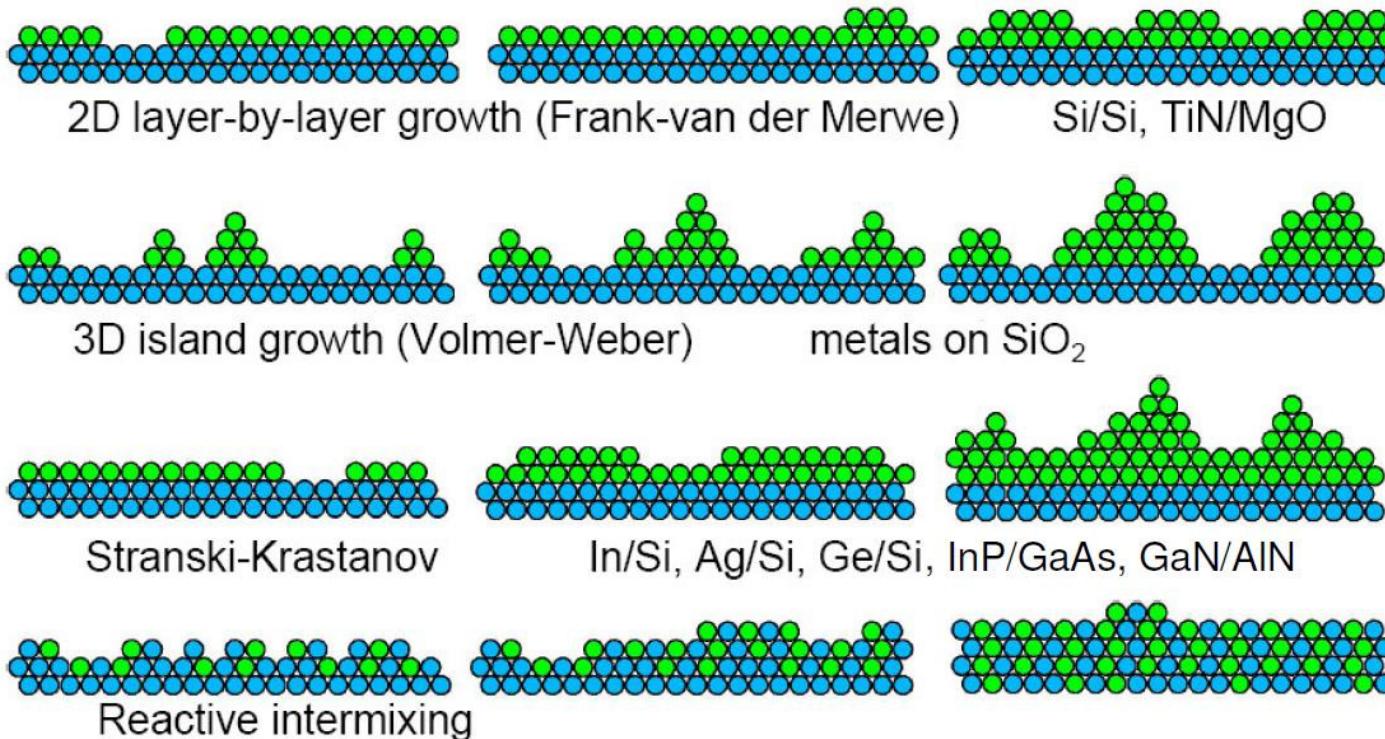


Figure 9-3

Voids in a chemical vapor deposition (CVD) oxide layer for narrow spaces between metal lines.
More difficult to fill without void for higher aspect ratio.

Four equilibrium growth/deposition modes

Epitaxy = crystal structure of film fits with the one of the substrate



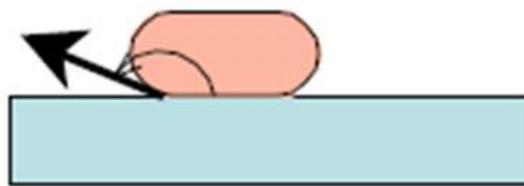
- Layer by layer growth: film atoms more strongly bound to substrate than to each other and/or fast diffusion
- Island growth: film atoms more strongly bound to each other than to substrate and/or slow diffusion.
- Mixed growth (Stranski - Krastanov): initially layer by layer then forms three dimensional islands.

Four growth modes

Thin film types based on crystallinity:

- Epitaxial (single-crystalline, formed layer-by-layer, lattice match to substrate): no grain boundaries, requires high temperatures and slow growth rate. High quality thin films such as III-V semiconductor films (e.g. GaAs) and complex oxides.
- Polycrystalline (island or mixed growth): lots of grain boundaries, e.g. most elemental metals grown near room temperatures.
- Amorphous (island or mixed growth): no-crystalline structures (yet with some short range atomic ordering), no crystalline defects, e.g. common insulators such as amorphous SiO₂.

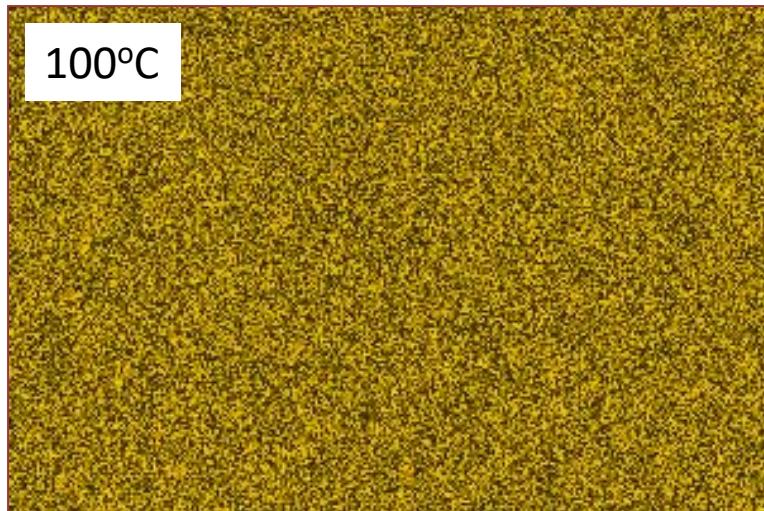
metals on dielectrics



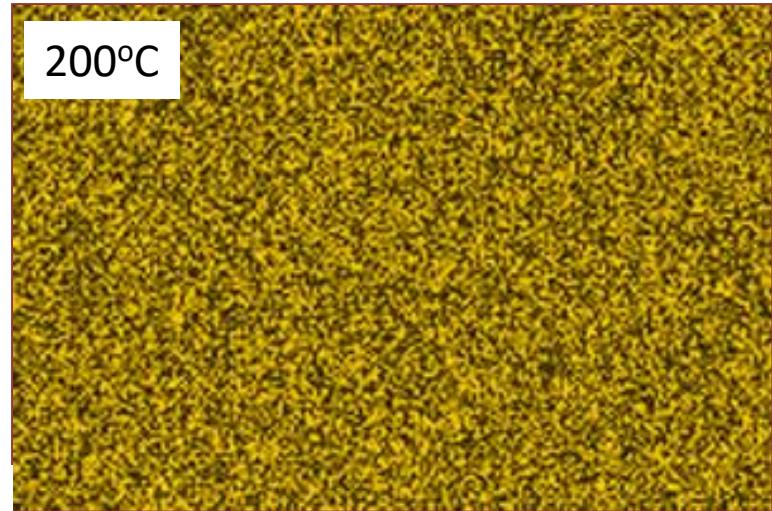
no wetting

- Noble metals don't bond ("wet") to Si/SiO₂ substrate, so tend to have island growth.
- Ag always form island (not continuous film); Au is better than Ag.
- Adhesion layer Ti or Cr can reduce island formation, but for Ag, surface is still very rough.
- Here, higher adhesion is because Ti or Cr bond chemically to O in SiO₂.

Effect of substrate temperature on the lateral grain size



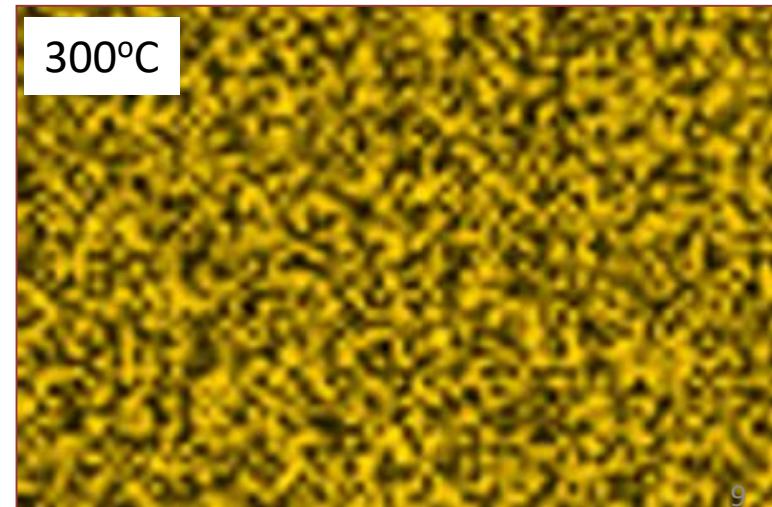
100 Å thick Au films deposited at 100, 200, and 300°C by vacuum evaporation



The small islands start coalescing with each other in an attempt to reduce the surface area.

This tendency to form bigger islands is termed agglomeration and is enhanced by increasing the surface mobility of the adsorbed species, such as by increasing the substrate temperature.

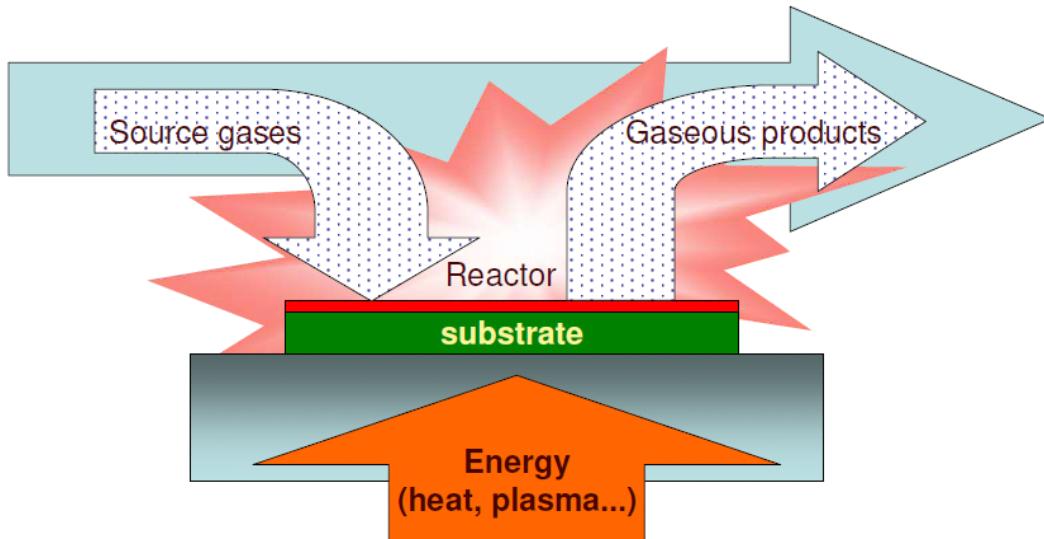
Except under special conditions, the crystallographic orientation and the topographical details of different islands are randomly distributed.



Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Chemical Vapor Deposition (CVD)

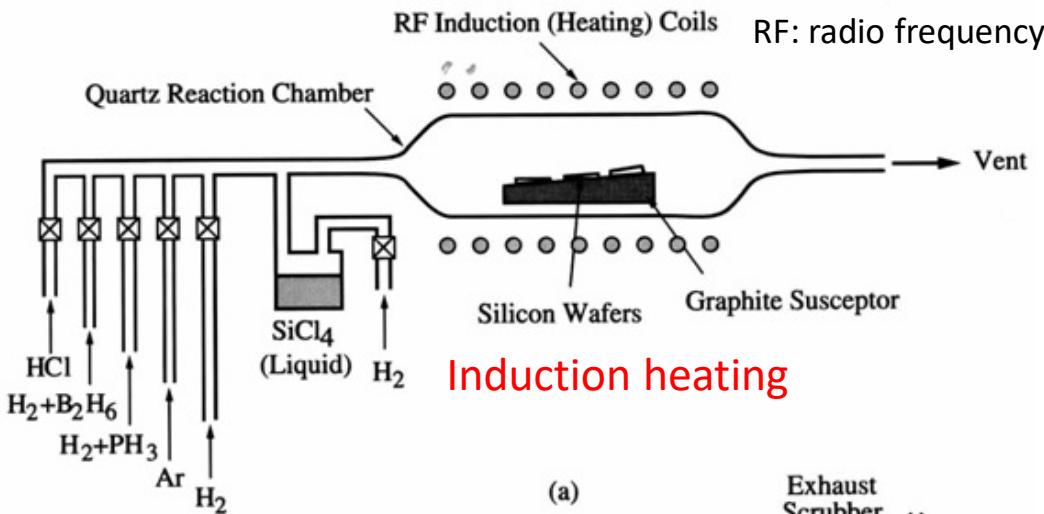


CVD : deposit film through chemical reaction and surface absorption.

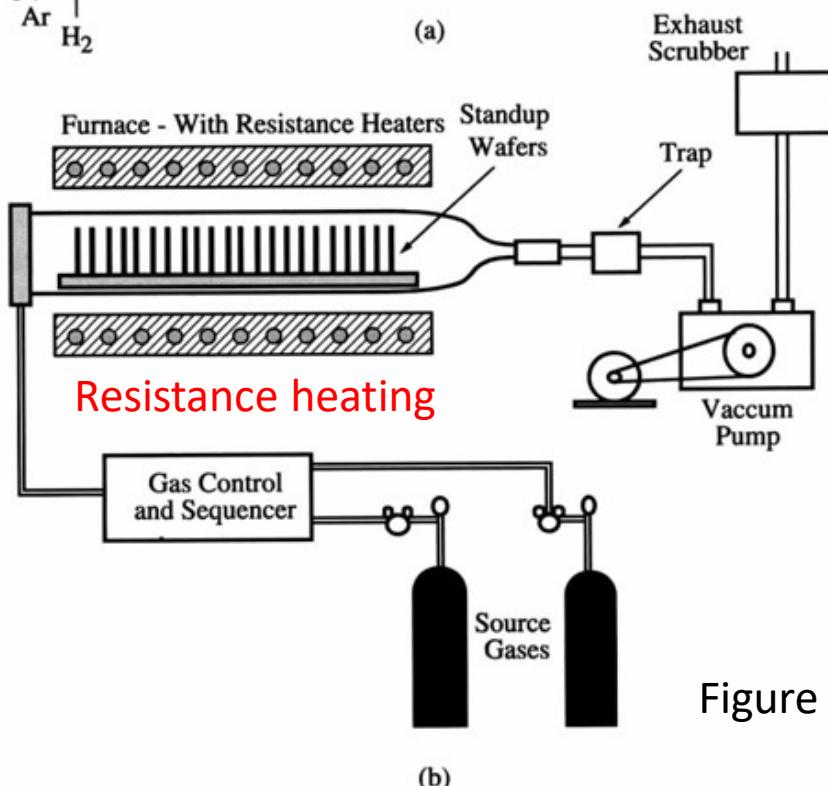
CVD steps:

- Introduce reactive gases to the chamber.
- Activate gases (decomposition) by heat or plasma.
- Gas absorption by substrate surface .
- Reaction take place on substrate surface, film formed.
- Transport of volatile byproducts away form substrate.
- Exhaust waste.

Chemical vapor deposition (CVD) systems



Atmospheric cold-wall system used for deposition of epitaxial silicon.
($\text{SiCl}_4 + 2\text{H}_2 \rightarrow \text{Si} + 4\text{HCl}$)



Low pressure hot-wall system used for deposition of polycrystalline and amorphous films, such as poly-silicon and silicon dioxide.

Figure 9-4

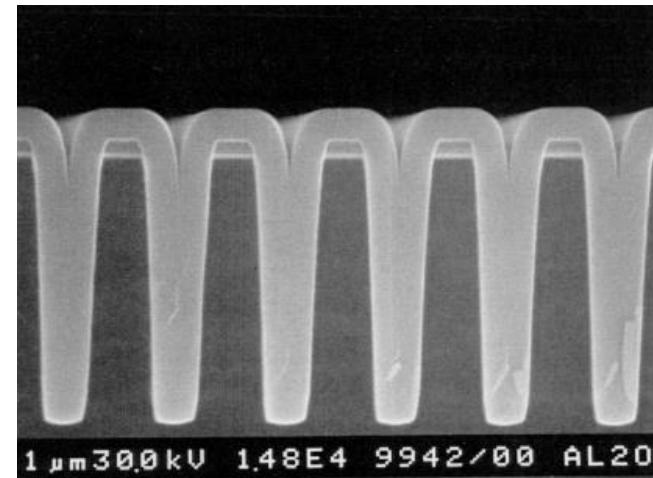
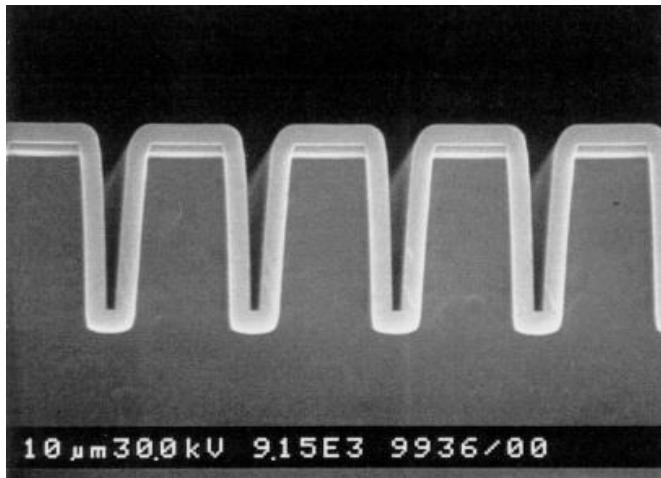
CVD advantages and disadvantages (as compared to physical vapor deposition)

Advantages:

- High growth rates possible, good reproducibility.
- Can deposit materials which are hard to evaporate.
- Can grow epitaxial films. In this case also termed as “vapor phase epitaxy (VPE)”. For instance, MOCVD (metal-organic CVD) is also called OMVPE (organo-metallic VPE).
- Generally better film quality, more conformal step coverage (see image below).

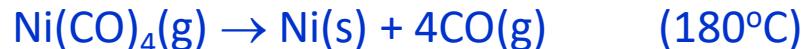
Disadvantages:

- High process temperatures.
- Complex processes, toxic and corrosive gasses.
- Film may not be pure (hydrogen incorporation...).



Types of CVD reactions

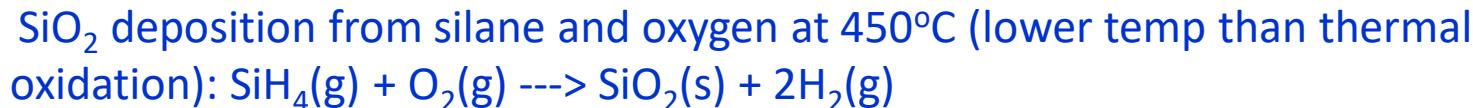
- **Thermal decomposition**



- **Reduction (using H₂)**

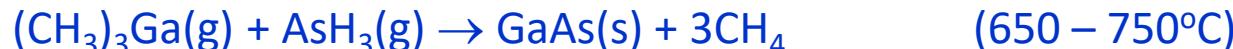


- **Oxidation (using O₂)**

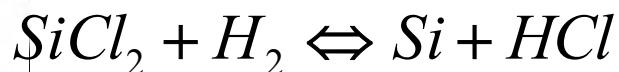
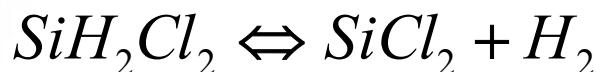
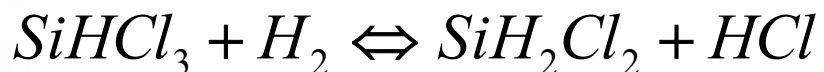
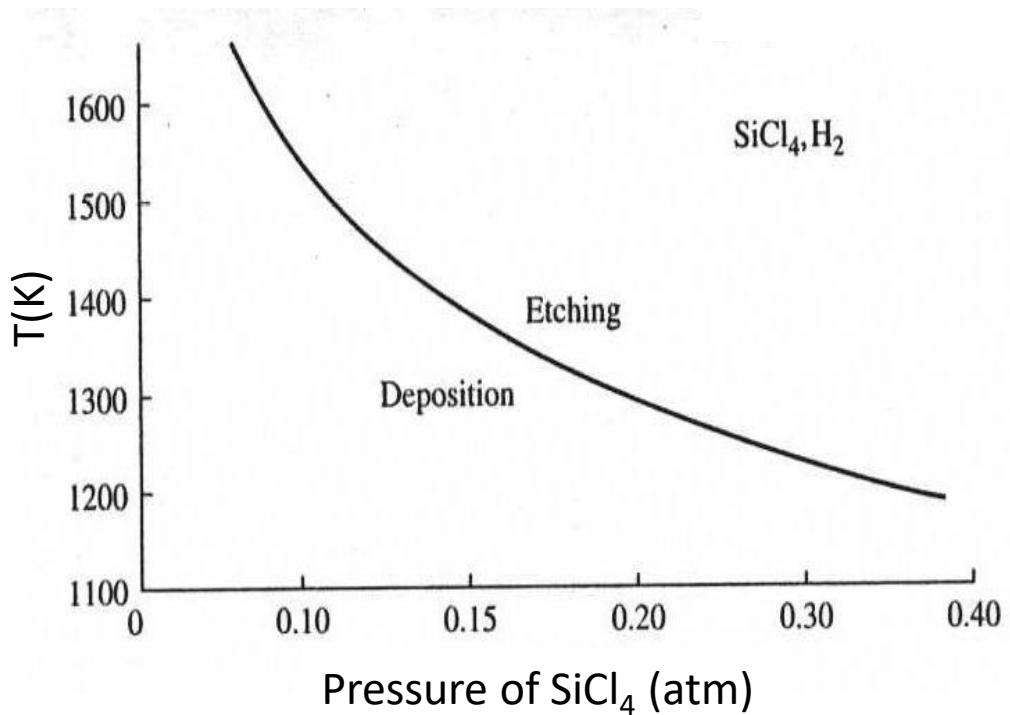


(O is more electronegative than Cl)

- **Compound formation (using NH₃ or H₂O)**



Chemical reactions for silicon epitaxial growth



(HCl etches Si at high T , which is used to prepare electronic grade Si)

Except SiH_4 decomposition, ALL other reactions are reversible.

Which direction (etching of Si or growth of Si) to go depends on the partial pressures of the reactants and temperature.

Thermal (not plasma-enhanced) CVD films

Thermal CVD Films and Coatings

Deposited material	Substrate	Input reactants	Deposition temperature (°C)	Crystallinity
Si	Single-crystal Si	SiCl_2H_2 , SiCl_3H , or $\text{SiCl}_4 + \text{H}_2$	1050–1200	E
Si		$\text{SiH}_4 + \text{H}_2$	600–700	P
Ge	Single-crystal Ge	GeCl_4 or $\text{GeH}_4 + \text{H}_2$	600–900	E
GaAs	Single-crystal GaAs	$(\text{CH}_3)_3\text{Ga} + \text{AsH}_3$	650–750	E
InP	Single-crystal InP	$(\text{CH}_3)_3\text{In} + \text{PH}_3$	725	E
SiC	Single-crystal Si	SiCl_4 , toluene, H_2	1100	P
AlN	Sapphire (Al_2O_3)	AlCl_3 , NH_3 , H_2	1000	E
$\text{In}_2\text{O}_3:\text{Sn}$	Glass	In-chelate, $(\text{C}_4\text{H}_9)_2\text{Sn}(\text{OOCH}_3)_2$, H_2O , O_2 , H_2	500	A
ZnS	GaAs, GaP	Zn, H_2S , H_2	825	E
CdS	GaAs, sapphire	Cd, H_2S , H_2	690	E
Al_2O_3	Si, cemented carbide	$\text{Al}(\text{CH}_3)_3 + \text{O}_2$, AlCl_3 , CO_2 , H_2	275–475 850–1100	A A
SiO_2	Si	$\text{SiH}_4 + \text{O}_2$, $\text{SiCl}_2\text{H}_2 + \text{N}_2\text{O}$	450	A
Si_3N_4	SiO_2	$\text{SiCl}_2\text{H}_2 + \text{NH}_3$	750	A
TiO_2	Quartz	$\text{Ti}(\text{OC}_2\text{H}_5)_4 + \text{O}_2$	450	A
TiC	Steel	TiCl_4 , CH_4 , H_2	1000	P
TiN	Steel	TiCl_4 , N_2 , H_2	1000	P
BN	Steel	BCl_3 , NH_3 , H_2	1000	P
TiB_2	Steel	TiCl_4 , BCl_3 , H_2	>800	P

Note: E = epitaxial; P = polycrystalline; A = amorphous.

Adapted from Refs. 1, 2, 3.

Types of sources:

- Gasses (easiest)
- Volatile liquids
- Sublimable solids

Types of CVD

APCVD (Atmospheric Pressure CVD), mass transport limited growth rate, leading to non-uniform film thickness.

LPCVD (Low Pressure CVD)

- Low deposition rate limited by surface reaction, so uniform film thickness (*many wafers stacked vertically* facing each other; in APCVD, wafers have to be laid horizontally side by side).
- Gas pressures around 1-1000mTorr (lower P => higher diffusivity of gas to substrate)
- Better film uniformity & step coverage and fewer defects
- Process temperature $\geq 500^{\circ}\text{C}$

PECVD (Plasma Enhanced CVD)

- Plasma helps to break up gas molecules: high reactivity, able to process at lower temperature and lower pressure (good for electronics on plastics).
- Pressure higher than in sputter deposition: more collision in gas phase, less ion bombardment on substrate
- Can run in RF plasma mode: avoid charge buildup for insulators
- Film quality is poorer than LPCVD.
- Process temperature around 100 - 400 $^{\circ}\text{C}$.

MOCVD (Metal-organic CVD, also called OMVPE - organo metallic VPE), epitaxial growth for many optoelectronic devices with III-V compounds for solar cells, lasers, LEDs, photo-cathodes and quantum wells.

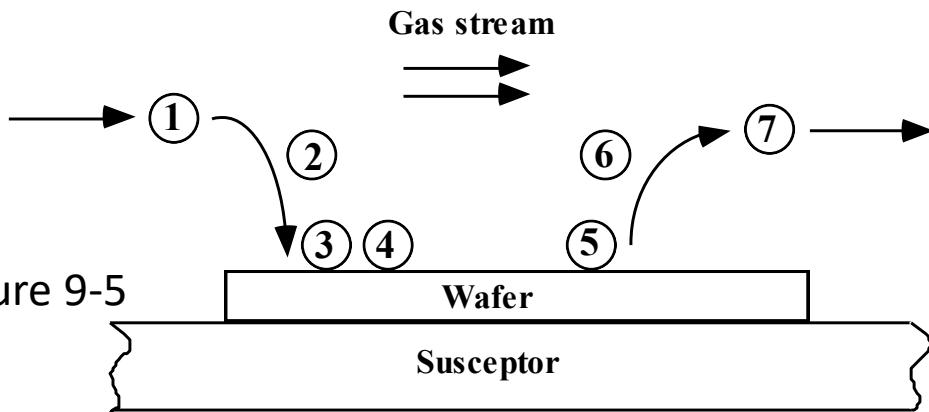
For R&D, PECVD is most popular (easy and low cost), followed by LPCVD.

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Steps involved in a CVD process

Figure 9-5



Reaction rate may be limited by:

- Gas transport to/from surface.
- Surface chemical reaction rate that depends strongly on temperature.

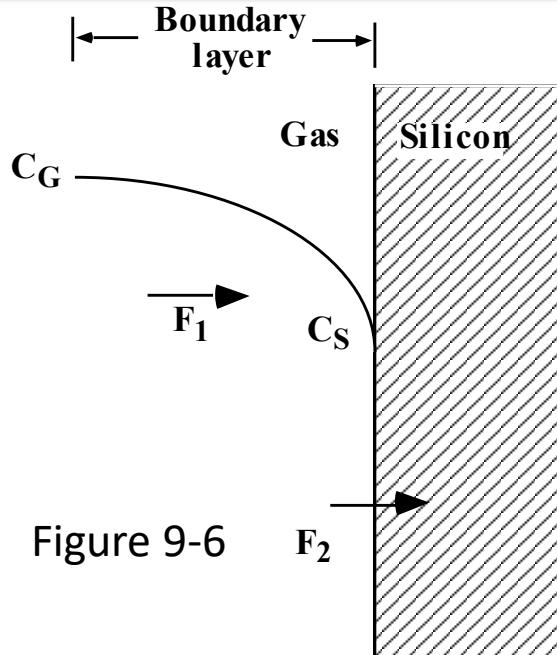
1. Transport of reactants to the deposition region.
2. Transport of reactants from the main gas stream through the boundary layer to the wafer surface.
3. Adsorption of reactants on the wafer surface.
4. Surface reactions, including: chemical decomposition or reaction, surface migration to attachment sites (kinks and edges); site incorporation; and other surface reactions.
5. Desorption of byproducts.
6. Transport of byproducts through boundary layer.
7. Transport of byproducts away from the deposition region.

Steps 2-5 are most important for growth rate.

Steps 3-5 are closely related and can be grouped together as “surface reaction” processes.

Derivation of film growth rate

(similar to/simpler than Deal-Grove model for thermal oxidation)



F_1 = diffusion flux of reactant species to the wafer through the boundary layer (step 2) = mass transfer flux

$$F_1 = h_G (C_G - C_S) \quad (1)$$

where h_G is the mass transfer coefficient (in cm/sec).

F_2 = flux of reactant consumed by the surface reaction (steps 3-5) = surface reaction flux,

$$F_2 = k_S C_S \quad (2)$$

where k_S is the surface reaction rate (in cm/sec).

In steady state:

$$F = F_1 = F_2 \quad (3)$$

Equating Equations (1) and (2) leads to

$$C_S = C_G \left(1 + \frac{k_S}{h_G} \right)^{-1} \quad (4)$$

The growth rate of the film is now given by

$$v = \frac{F}{N} = \frac{k_S h_G}{k_S + h_G} \frac{C_G}{N} = \frac{k_S h_G}{k_S + h_G} \frac{C_T}{N} Y \quad (5)$$

where N is the number of atoms per unit volume in the film and Y is the mole fraction (partial pressure/total pressure) of the incorporating species, C_T is total concentration of all molecules in the gas phase .

Derivation of film growth rate (continued)

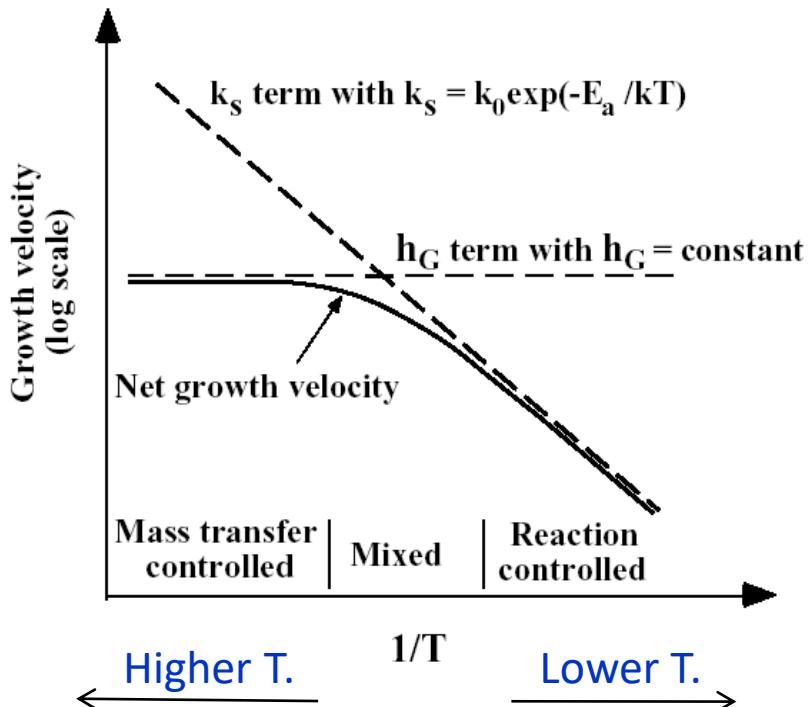
$$v = \frac{F}{N} = \frac{k_S h_G}{k_S + h_G} \frac{C_G}{N} = \frac{k_S h_G}{k_S + h_G} \frac{C_T}{N} Y \quad (5)$$

(a). If $k_S \ll h_G$, then we have the surface reaction controlled case:

$$v \approx \frac{C_T}{N} k_S Y \quad (6)$$

(b) If $h_G \ll k_S$, then we have the mass transfer, or gas phase diffusion, controlled case:

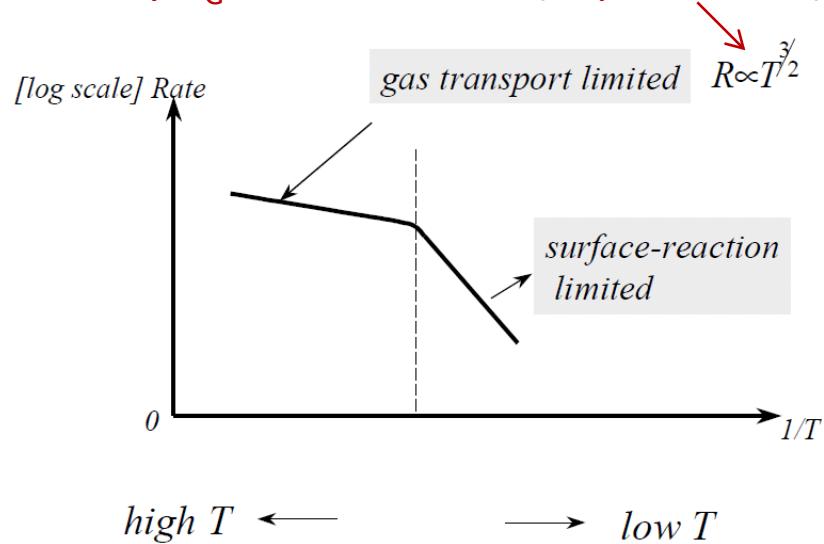
$$v \approx \frac{C_T}{N} h_G Y \quad (7)$$



- k_S increases with temperature.
(Arrhenius with E_A depending on the particular reaction, e.g. 1.6 eV for single crystal silicon deposition).
- $h_G \approx \text{constant}$
(diffusion through boundary layer is insensitive to temperature)

CVD film growth rate

Actually h_G is not constant (depends on T)



Deposition rate vs. gas glow rate

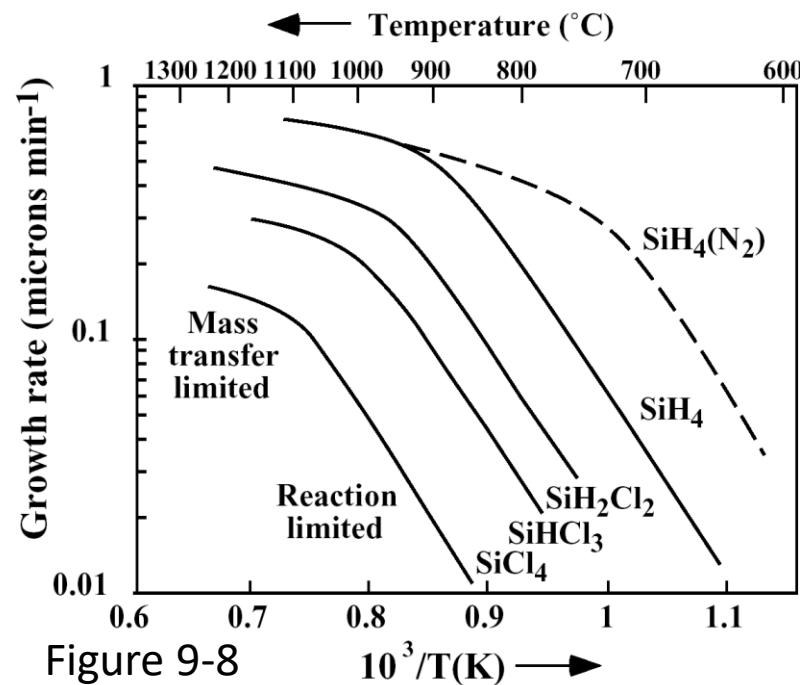
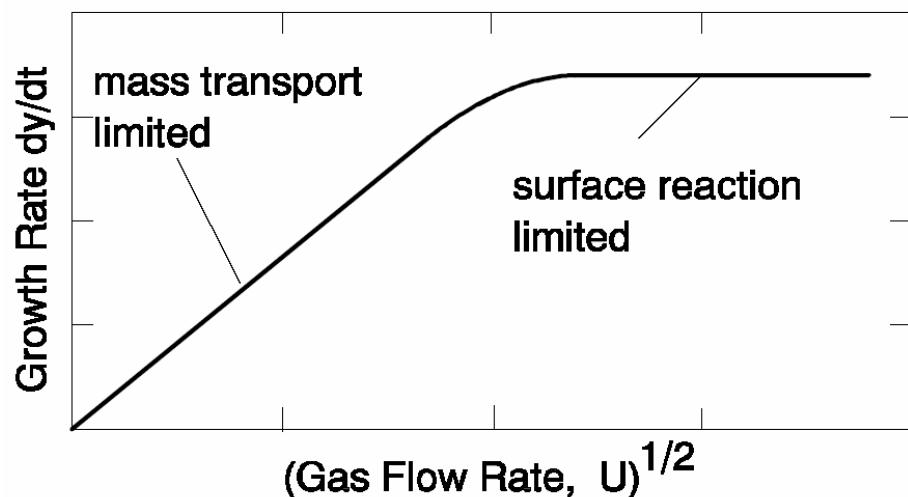
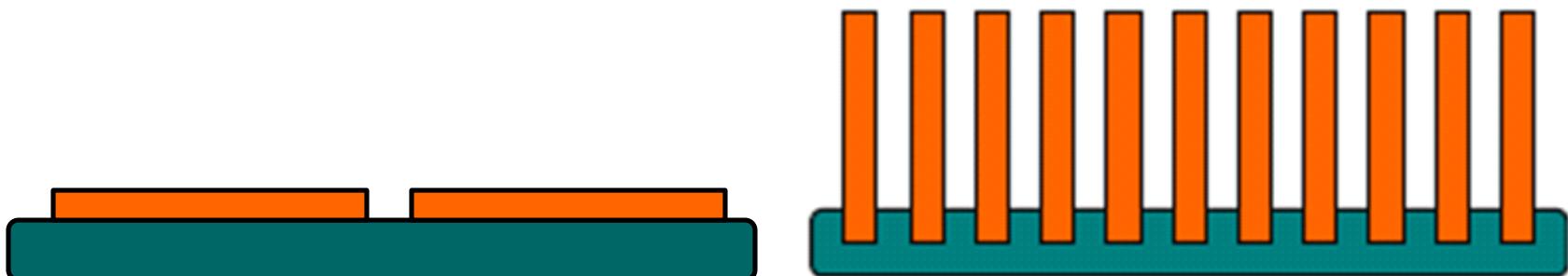


Figure 9-8 Growth or deposition rate for silicon by APCVD. The partial pressure of the reactant gas is 0.8Torr (1atm=760Torr!!). H_2 is used as the carrier or diluent gas for the solid curves. For SiH_4 , using N_2 carrier gas increases the growth rate, because the carrier gas H_2 is a reaction product of SiH_4 decomposition, thus slowing down the reaction.

Chemical Vapor Deposition (CVD) growth rate

- k_s limited deposition is VERY temperature sensitive.
$$k_s = k_0 \exp\left(-\frac{E_a}{kT}\right)$$
- h_G limited deposition is VERY geometry (boundary layer) sensitive.
- Si epitaxial deposition is often done at high T to get high quality single crystal growth. It is then h_G controlled, and horizontal reactor configuration is needed for uniform film thickness across the wafer.
- When a high film quality is less critical (e.g. SiO_2 for inter-connect dielectric), deposition is done in reaction rate controlled regime (lower temperature). Then one can greatly increase the throughput by stacking wafers vertically (for research, usually 25 wafers per run; 100-200 for industry).



Other factors affecting growth rate: thickness of boundary layer and source gas depletion

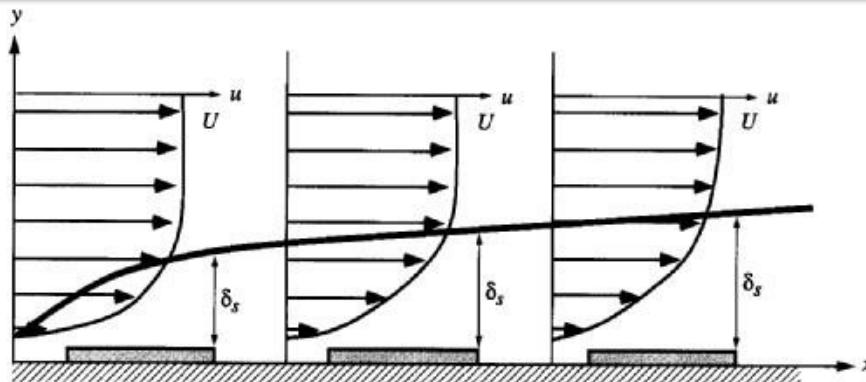


Figure 9-9 Boundary layer velocities along susceptor. δ_s is the thickness of the boundary layer. The boundary layer increases with distance in the direction of the gas flow from Newton's second law.

Gas moves with the constant velocity U .
Boundary layer (caused by friction) increases along the susceptor, so mass transfer coefficient h_G decreases.
Source gas also depletes (consumed by chemical reaction) along the reactor.
Both decrease growth rate along the chamber.

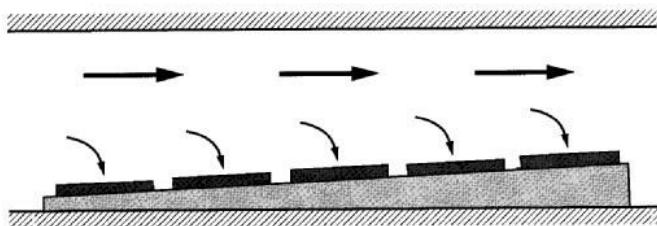


Figure 9-10 The susceptor in a horizontal epitaxial reactor is tilted so that the cross-sectional area of the chamber is decreased, increasing the gas velocity along the susceptor. This compensates for both the boundary layer and depletion effects.

To compensate for this, one can:

- Use tilted susceptor.
- Use temperature gradient 5-25°C.
- Gas injectors along the tube.
- Use moving belt.

Doping in CVD films

- Doping is usually done for epitaxial (thus single crystal) film during film growth.
- Dopant will be grown directly onto crystalline site (no need of dopant activation).
- Doping is realized by adding gas containing the dopant. Such as PH_3 , B_2H_6 , AsH_3 (all gas phase at room temperature (RT)); or PCl_3 , BCl_3 , AsCl_3 (all liquid at RT).
- They will go through: dissociation, lattice site incorporation, and burying of dopants by other atoms in the film.
- The dopant concentration C : (P is partial pressure of the dopant species, and v growth rate)

$$C \propto P_i \quad \text{for low growth rates}$$

$$C \propto \frac{P_i}{v} \quad \text{for high growth rates}$$

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Low Pressure Chemical Vapor Deposition (LPCVD)

Atmospheric pressure systems (APCVD): at high T, a horizontal configuration must be used (few wafers at a time).

Obviously, the solution is to operate at low pressure – LPCVD.

In the mass transfer limited regime,

$$h_G = \frac{D_G}{\delta_s} \quad \text{But diffusivity } D_G \propto \frac{1}{P_{total}}$$

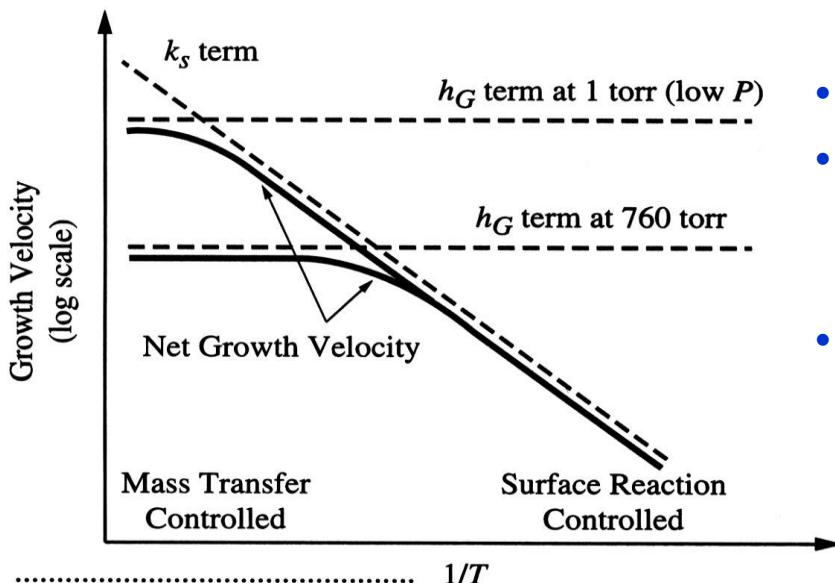
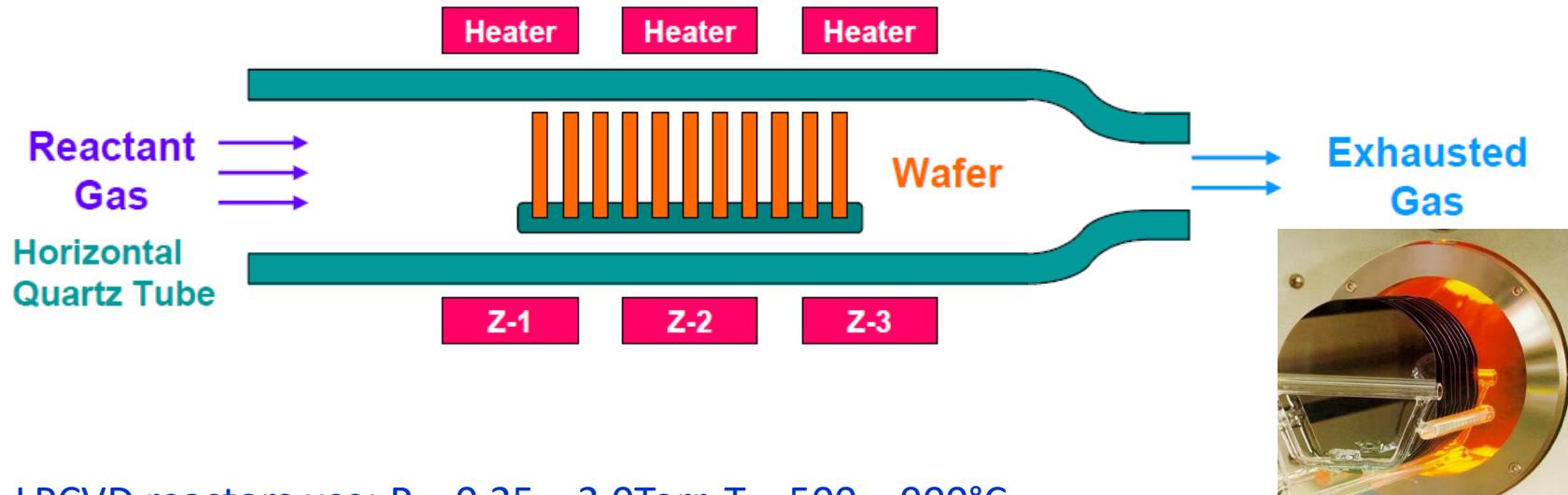


Figure 9-13 Growth velocity vs. $1/T$ for APCVD (760 torr) and LPCVD (1 torr) systems. The lower total pressure (with P_G and C_G remaining fixed) shifts the h_G curve upward, extending the surface reaction regime to higher temperatures.

- So as P_{total} goes down, D_G and hence h_G will go up.
- E.g. when pressure reduced from 1 atmosphere to 1 Torr ($760\times$), h_G increases by $\sim 100\times$ (because δ_s increases by only $5-7\times$).
- Higher h_G means higher T can be used while still $k_s < h_G$ (i.e. still in surface reaction controlled regime).

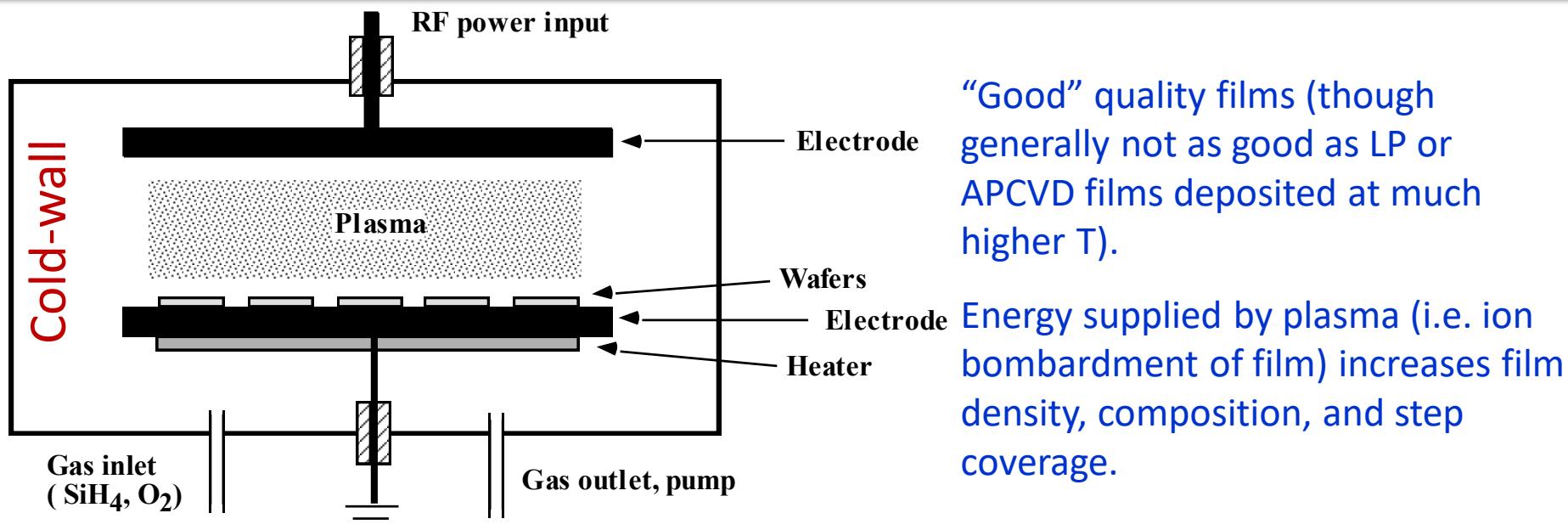
This is not one expects: lower pressure means less reactants, thus lower rate. But for APCVD, the reactant gas is only a small portion of the total gas.

Low Pressure Chemical Vapor Deposition (LPCVD)



- LPCVD reactors use: $P = 0.25 - 2.0 \text{ Torr}$, $T = 500 - 900^\circ\text{C}$.
- Transport of reactants from gas phase to surface through boundary layer is not rate limiting, so wafers can be stacked vertically for high throughput (100-200 wafers per run). That is, as long as temperature is uniform, deposition is uniform even though gas transport to surface is very non-uniform (e.g. faster transport to wafer edge than center).
- Because LPCVD operates in reaction limited regime, it is VERY sensitive to temperature and so temperature needs to be controlled closely (within $\pm 1^\circ\text{C}$), hence it uses hot walled reactor (resistance heating, like oxidation furnace) for this precise control.
- Requires no carrier gas, since low gas pressure drastically reduces gas-phase reaction, which causes particle cluster that contaminates the wafer and system. That is, we want reaction/deposition only on wafer surface, not in free space (gas phase).

Plasma Enhanced CVD (PECVD)



“Good” quality films (though generally not as good as LP or APCVD films deposited at much higher T).

Energy supplied by plasma (i.e. ion bombardment of film) increases film density, composition, and step coverage.

- Use RF-induced plasma to transfer energy into the reactant gases, forming radicals that is very reactive. (RF: radio-frequency, typically 13.56MHz for PECVD)
- Low temperature process ($<300^\circ\text{C}$), as thermal energy is less critical when RF energy exists.
- Used for depositing film on metals (Al...) and other materials that cannot tolerate high temperatures. (APCVD/LPCVD at such low temperatures gives poor film)
- Surface reaction limited deposition, thus substrate temperature control is important to ensure uniformity.
- At low T, surface diffusion is slow, so one must supply kinetic energy for surface diffusion – plasma (ion bombardment) provides that energy and enhances step coverage.
- Disadvantages: plasma damage, not pure film (often lots of H incorporated into film).

PECVD process parameter

Substrate temperature (100-300°C, up to 1000°C PECVD available)

Gas flow (10s to 100s sccm – standard cubic centimeter per minute)

- Higher flow rates can increase deposition rate and uniformity

Pressure ($P \approx 50\text{mTorr} - 5\text{Torr}$)

- Changes the energy of ions reaching electrodes
- Can change deposition rate
- Increases pressure may lead to chemical reaction in the gas (bad)

Power (10s to 100s watts)

- Increased power increases deposition rate
- Increased power may lead to chemical reaction in gas

Frequency (mostly 13.56MHz, same for plasma etching and sputter deposition)

- Changes plasma characteristics
- Changes ion bombardment characteristics

Examples of PECVD systems and applications

Material Deposited	Common Precursors	Deposition Temp. °C	Applications	Status
a-Si	SiH ₄ -H ₂	250	semiconductor production photovoltaic	
Epitaxial silicon	SiH ₄	750	semiconductor	R & D
Si ₃ N ₄	SiH ₄ -N ₂ -NH ₃	300	passivation	production
SiO ₂	SiH ₄ -N ₂ O	300	passivation optical fiber decorative	production
Boro-phospho-silicate	SiH ₄ -TEOS-B ₂ H ₆ -PH ₃	355	passivation	semi-production
W	WF ₆	250–400	conductor in IC's	R & D
WSi ₂	WF ₆ -SiH ₄	230	conductor in IC's	semi-production
TiSi ₂	TiCl ₄ -SiH ₄	380–450	conductor in IC's	semi-production
TiC	TiCl ₄ -C ₂ H ₂	500	abrasion cutting tools	R & D
TiN	TiCl ₄ -NH ₃	500	abrasion cutting tools	R & D
Diamond-like carbon	CH ₄ -H ₂ hydrocarbon	300	wear, erosion, optical	semi-production

PECVD Films, Source Gases, and Deposition Temperatures

Film	Source gases	Deposition temperature (°C)
Elemental		
Al	AlCl ₃ -H ₂	100–250
a-B	BCl ₃ -H ₂	400
a-C	C _n H _m -H ₂ /Ar	25–250
a-Si	SiH ₄ -H ₂	300
c-Si	SiH ₄ -H ₂	400
Oxides		
Al ₂ O ₃	AlCl ₃ -O ₂	100–400
SiO ₂	SiCl ₄ -O ₂	100–400
TiO ₂	TiCl ₄ -O ₂	100–500
Nitrides		
AlN	AlCl ₃ -N ₂	<1000
BN	B ₂ H ₆ -NH ₃	300–700
	BCl ₃ -NH ₃ /Ar	300–700
Si ₃ N ₄	SiH ₄ -NH ₃ -N ₂	25–500
TiN	TiCl ₄ -N ₂ -H ₂	100–500
Carbides		
B ₄ C	B ₂ H ₆ -CH ₄	400
BCN	B ₂ H ₆ -CH ₄ -N ₂	~25
	C ₈ H ₁₈ BN	250
SiC	SiH ₄ -C _n H _m	140–600
TiC	TiCl ₄ -CH ₄ -H ₂	400–900
Borides		
TiB ₂	TiCl ₄ -BCl ₃ -H ₂	480–650

Miscellaneous: selective deposition and laser CVD

Selective deposition:

- Especially important in microelectronics, surface patterning and 3D-growth.
- Reaction rate of precursor is limited on a non-growth surface. E.g. deposition of Cu from (hfac)Cu(PMe₃) occur on Cu, Pt... but not on SiO₂.
- Growth surface acts as co-reactant, and is selectively consumed. E.g. Si reacts with WF₆ or MoF₆, while reaction at SiO₂ or Si₃N₄ is slower.
- A chemical reaction of a gaseous co-reactant occur on the growth surface. E.g. H₂ dissociation on a metal surface, but not on SiO₂ or metal oxide surfaces.

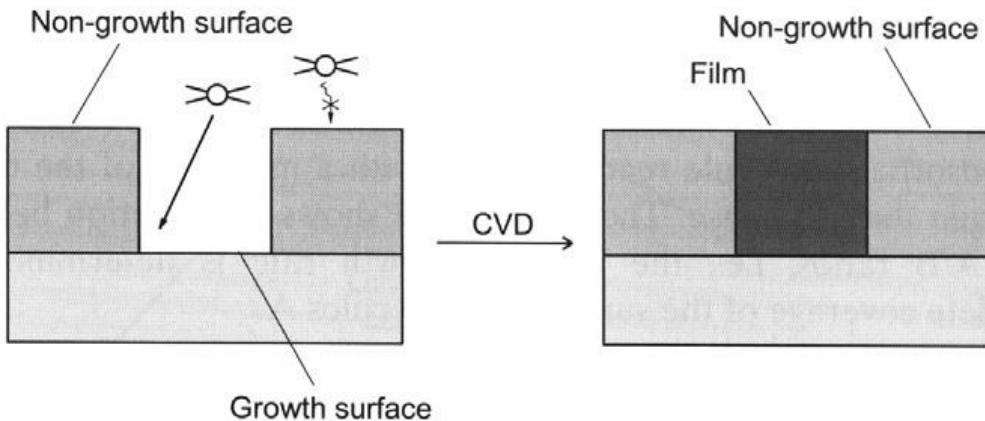
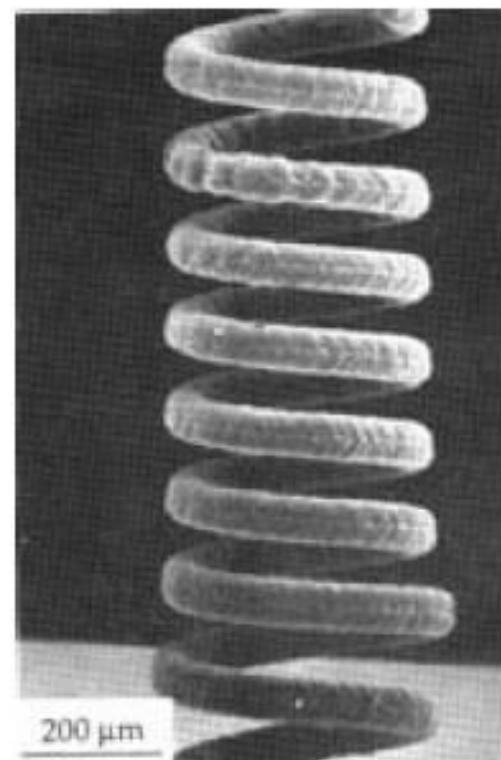


Figure 3-11. Selective deposition by CVD.

Laser CVD

(energy provided by laser)



Tungsten spring grown by laser CVD.

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

Physical vapor deposition (PVD): evaporation and sputtering

In PVD, chemical reactions are not involved, except for reactive (add reactive gases into chamber) evaporation or reactive sputter deposition, which are not widely used.

Evaporation:

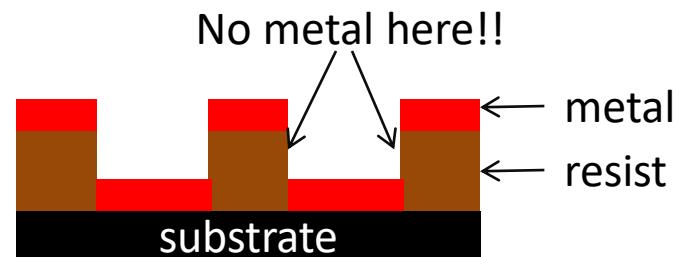
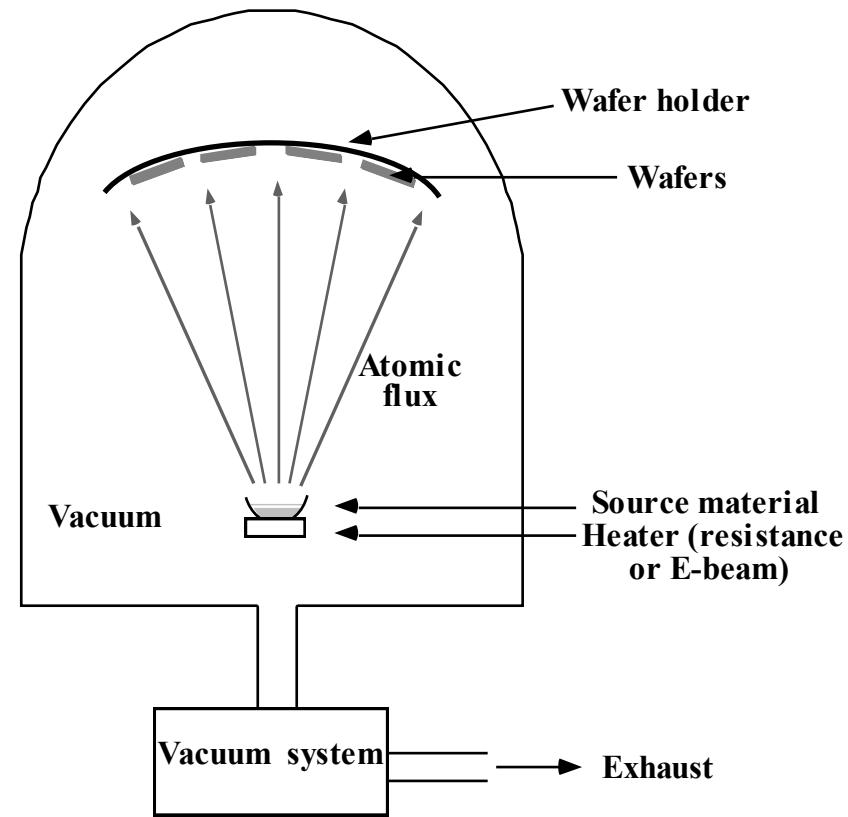
- Material source is heated to high temperature in vacuum either by thermal or e-beam methods.
- Material is vapor transported to target in vacuum.
- Film quality is often not as good as sputtered film (that involves energetic bombardment of ions to the as-deposited film, which makes the film denser).
- The film thickness can be monitored precisely using a quartz balance – this is necessary as the deposition is not reproducible (tiny change in T leads to large change of deposition rate. T is not monitored, power is).

Sputter deposition: (there is also sputter etching)

- Material is removed from target by momentum transfer.
- Gas molecules (Ar) are ionized in a glow discharge (plasma), ions (Ar^+) strike target and remove the atoms.
- Sputtered atoms condense on the substrate.
- Not in vacuum, gas (Ar) pressure 5-50mTorr.

Evaporation (also called vacuum deposition)

- In evaporation, source material is heated in high vacuum chamber ($P < 10^{-5}$ Torr), hence the name vacuum deposition.
- High vacuum is required to minimize collisions of source atoms with background species (no collision, line
- of sight deposition)
- Heating is done by resistive or e-beam sources.
- Surface interactions are physical, can be very fast ($>1\mu\text{m}/\text{min}$ possible, but film quality may suffer. For R&D typical $0.1\text{-}1\text{nm/sec}$).
- High sticking coefficient (i.e. adatom stays wherever it hits, with little surface migration), leading to poor conformal coverage/significant shadow. **But this poor coverage also makes evaporation the most popular thin film deposition for nanofabrication using liftoff process.**
 - For microfabrication R&D, evaporation is as important as sputter deposition.



Evaporation: vacuum pressure and mean free path

Assume a particle of diameter σ , moving in speed v .

Collision cross-section = $\pi\sigma^2$

Collision volume swept during dt = $\pi\sigma^2 v dt$

of collisions during dt = $(n/V) \pi\sigma^2 v dt$

Here n/V is number density.

Average mean free path $\lambda = vdt/\# = [(n/V) \pi\sigma^2]^{-1}$

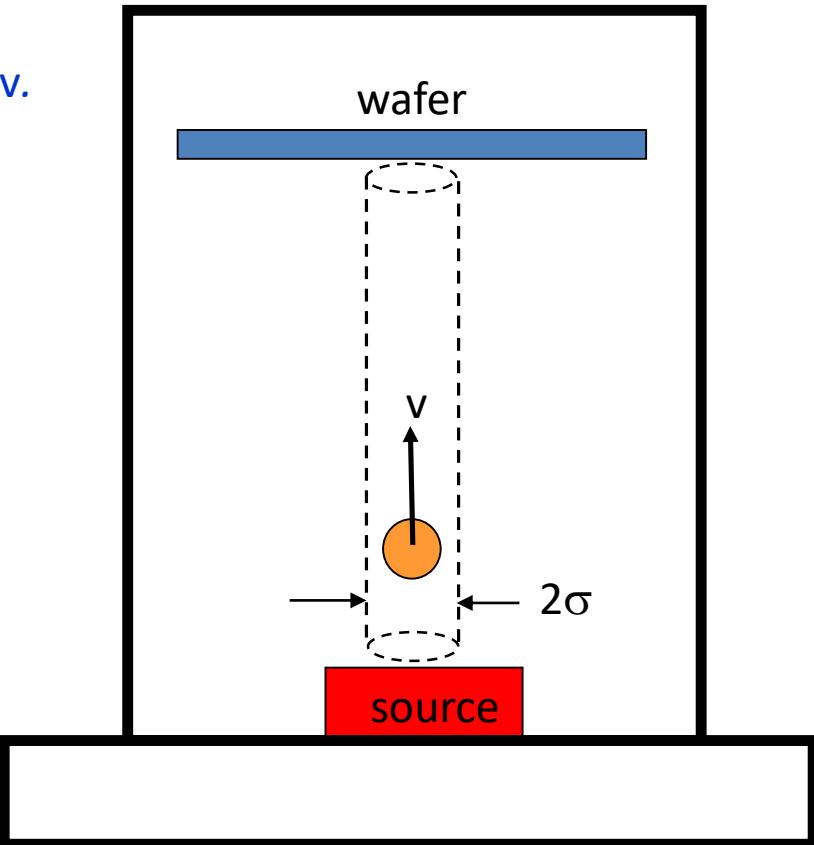
In more rigorous treatment, ($PV=nkT$)

$$\lambda = \left(\frac{n}{V} \pi \sigma^2 \sqrt{2} \right)^{-1} = \frac{kT}{\sqrt{2} \pi \sigma^2 P}$$

Assume $\sigma = 3\text{\AA}$, $T= 300\text{K}$

$$P(\text{Torr}) \times \lambda(\text{cm}) = 7.8 \times 10^{-3}$$

$$1\text{Torr} = 1\text{mmHg} = 1/760\text{atm}$$



If $\lambda=30\text{ cm}$, then a pressure $<2.6 \times 10^{-4}\text{ Torr}$ is required.

Typical vacuum for evaporation $<5 \times 10^{-6}\text{ Torr}$

Mean-free path of varied gases

Gas	L·P (cm·torr) at 20°C	
O ₂	4.9×10^{-3}	Atmospheric Pressure (760 torr)
N ₂	4.6×10^{-3}	~ 40 nm (small than most chip features)
H ₂	9×10^{-3}	Low Vacuum (0.76 torr)
He	13.6×10^{-3}	~ 40 μm (larger than most chip features)
Ar	4.8×10^{-3}	Medium Vacuum (7.6 mtorr)
Ne	9.3×10^{-3}	~ 4.0 mm (larger than working gap)
Xe	2.7×10^{-3}	High Vacuum (7.6 μtorr)
CO ₂	3.0×10^{-3}	~ 4.0 m (larger than chamber)
H ₂ O	3.0×10^{-3}	
Cl	2.3×10^{-3}	
NH ₃	3.5×10^{-3}	<i>Evaporated material travels straight to wafer (line of sight) – thus is very directional</i>

Arrival ratio

Pressure [Torr]	Mean Free Path	Arrival Ratio*
10^{-1}	0.5 mm	10^{-4}
10^{-3}	5 cm	10^{-2}
10^{-5}	5 m	1
10^{-6}	50 m	10
10^{-7}	500 m	10^2
10^{-8}	5 km	10^3
10^{-9}	50 km	10^4

Arrival ratio:

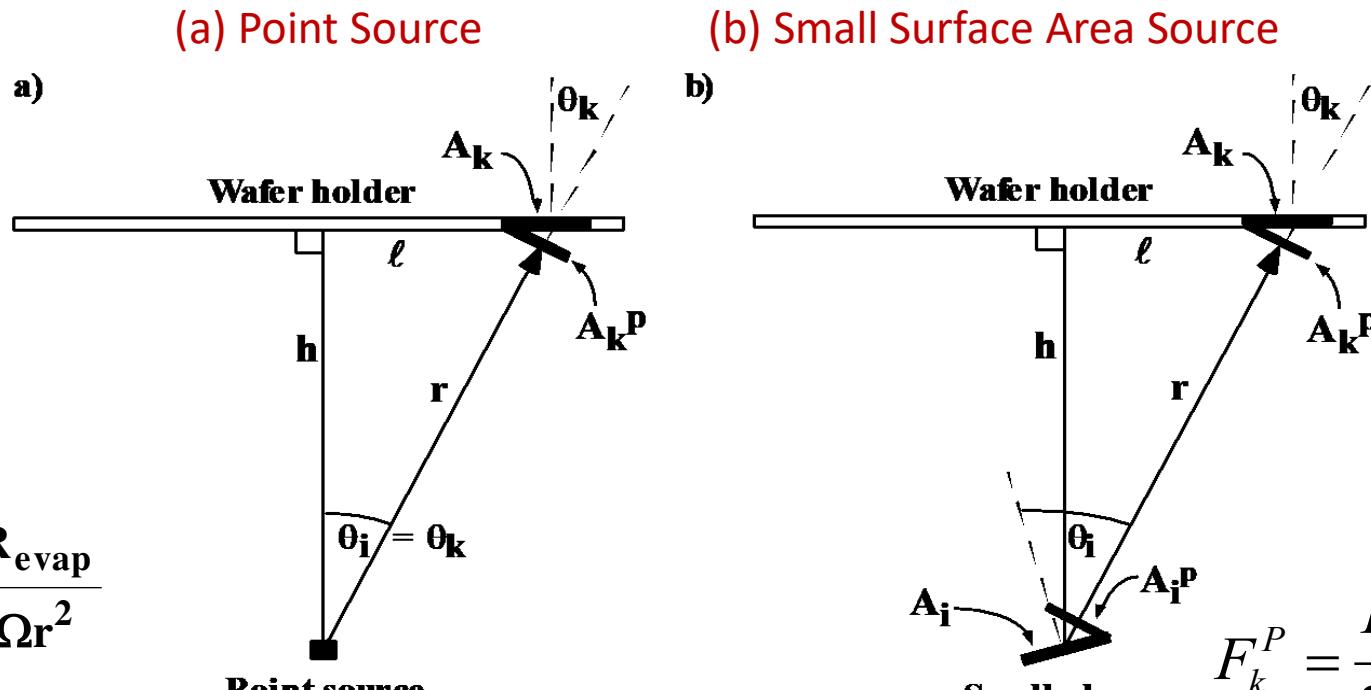
Ratio of molecular vapor (e.g. Al atom) arrival at 1nm/sec deposition rate to molecular impact of residual gas (e.g. O₂).

Arrival ratio = 1 means the number of film molecules hitting the surface per second is the same as the number of gas molecules.

For materials that is very reactive with the gas (such as Ti to O₂), the film can be very impure when arrival ratio ~1.

In fact, Ti is used as a pump for ultrahigh vacuum (called Ti pump, where Ti vapor is produced to trap gas in vacuum chamber).

Point source and surface source model



$$F_k^P = \frac{R_{evap}}{\Omega r^2}$$

$$v = \frac{R_{evap}}{\Omega N r^2} \cos \theta_k$$

Figure 9-16

F: flux that strikes area A (atoms or grams per cm^{-2}sec).

v: deposition rate (nm/sec).

N: density of the material (atoms or grams per cm^3).

R_{evp} : evaporation rate from the source, in atoms or grams per second.

Ω : solid angle, $=4\pi$ if source emits in all directions; $=2\pi$ if emits only upward.

$$F_k^P = \frac{R_{evap}}{2\pi r^2} \cos^n \theta_i$$

$$v = \frac{R_{evap}}{2\pi N r^2} \cos^n \theta_i \cdot \cos \theta_k$$

Deposition rate across a wafer

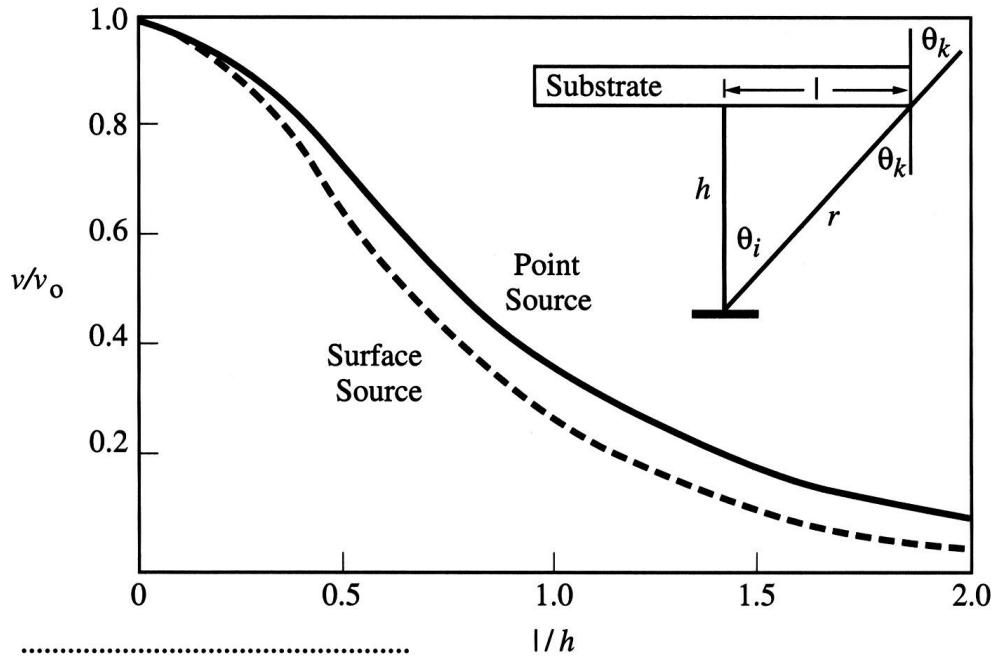
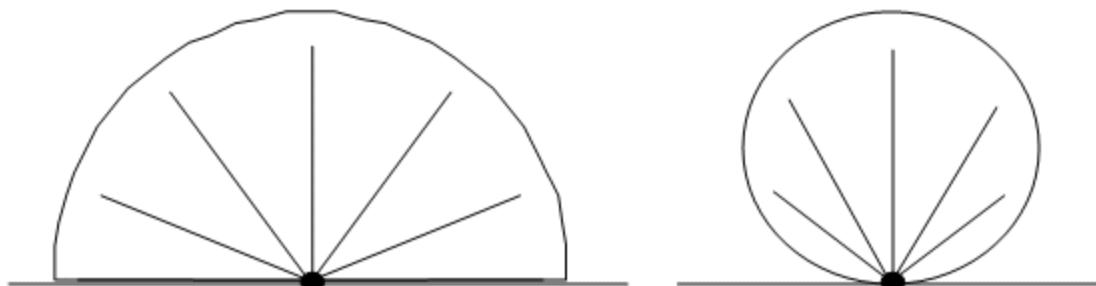


Figure 9-17 Deposition rate of evaporated film as a function of position on substrate for point and surface sources. $\theta_i = \theta_k$ in this configuration for both point and surface sources. (After [9.6].)

For typical evaporator: $h=40\text{cm}$.
Then for a 4" (10cm) wafer, wafer edge $l/h=5/40=0.12$

Larger h means: more uniform deposition, but

- Bigger chamber
- Higher capacity vacuum pump
- Lower deposition rate
- High evaporant waste



a. Uniform (isotropic emission from a point source).

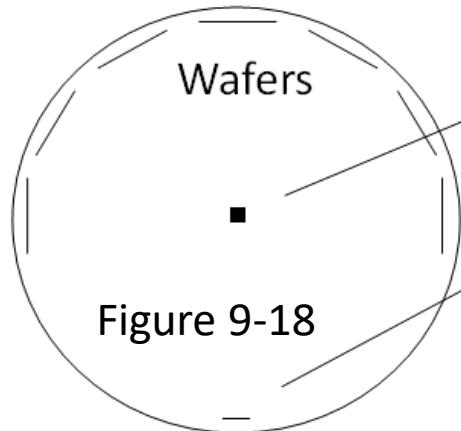
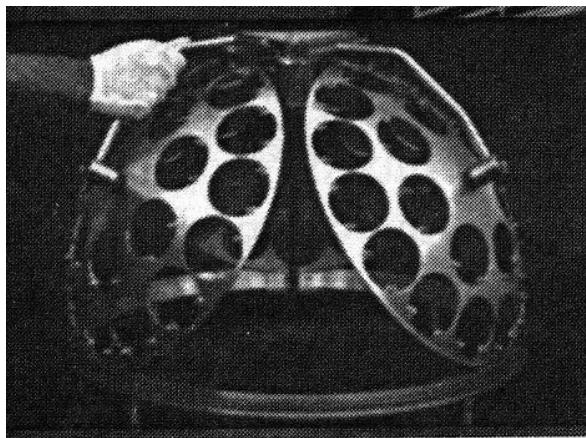
b. Ideal cosine emission from a small planar surface source. ($n=1$ in $\cos^n\theta_i$ distribution)

c. Non-ideal, more anisotropic emission from a small surface source. ($n>1$ in $\cos^n\theta_i$ distribution)

How to achieve uniform deposition

Use spherical wafer holder:

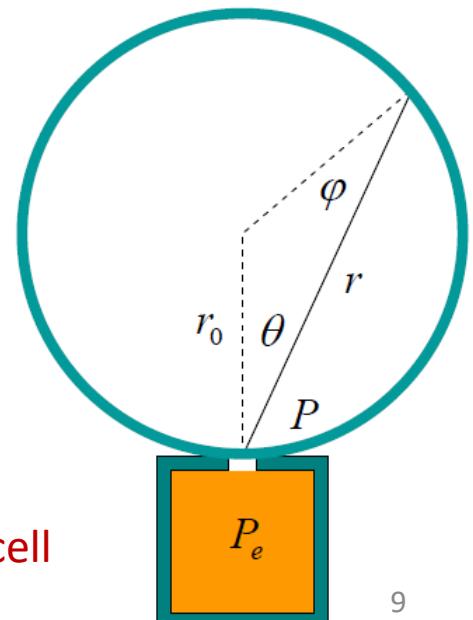
- Point source: put source at center of sphere.
- Small surface source: put source on inside surface of sphere (compensates for $\cos^n\theta$).



Location of source which behaves like an ideal point source

Location of source which behaves like an ideal small area surface source

It is shown that Knudsen-cell behavior (i.e. $\cos\theta$ emission from a small surface), rather than point source emission or source with $\cos^n\theta$ $n>1$, is close to *experimental* results.



Spherical surface with source on its edge:

$$\cos\theta = \cos\varphi = \frac{r}{2r_0}$$

$$v = \frac{R_{evap}}{\pi N r^2} \cos\theta \cdot \cos\varphi = \frac{R_{evap}}{4\pi N r_0^2}$$

Angle independent, uniform coating!

Knudsen-cell

Langmuir-Knudsen theory for R_{evp}

$$R_{\text{evp}} = 5.83 \times 10^{-2} A_s \left(\frac{m}{T} \right)^{1/2} P_e$$

R_{evp} : evaporation rate (g/sec)

A_s : source area (cm^2)

m : gram-molecular mass (g/mol)

T : temperature (K)

P_e : equilibrium vapor pressure (Torr)

To see whether this makes sense or not, consider the “Knudsen-cell” (i.e. a box of gas with pressure P_e emitting from a small opening into vacuum).

Assume all gas molecules move only along $\pm x, y, z$ -direction. Then $1/6$ moves along $+z$ direction with velocity v .

For opening area A , the emitted number of molecules during time Δt is: $1/6n \times A(v \times \Delta t)$, here n is molecular number density.

So per unit emitting area, emitting/evaporation rate is: $1/6nv$.

A more strict derivation gives flux $1/4nv$.

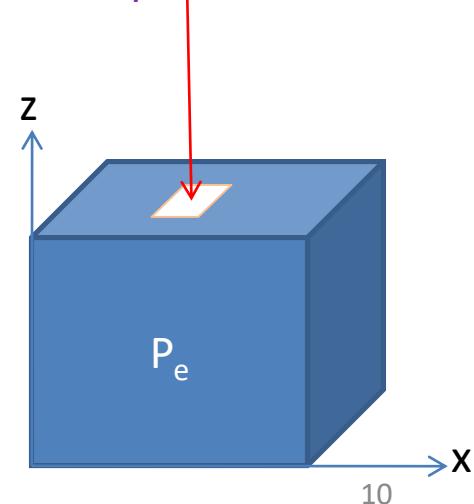
$$R = \frac{1}{4} nv = \frac{1}{4} \frac{P}{kT} \sqrt{\frac{8kT}{\pi m}} = \frac{P}{\sqrt{2\pi mkT}}$$

To change unit from (Pa, kg/mol, m^2) into (Torr, g/mol, cm^2):

$$R = \frac{133P}{\sqrt{2\pi kT} \frac{M}{1000N_A}} \times 10^{-4} = 3.51 \times 10^{22} \frac{P}{\sqrt{MT}} \frac{\text{molecules}}{\text{cm}^2 \cdot \text{sec}}$$

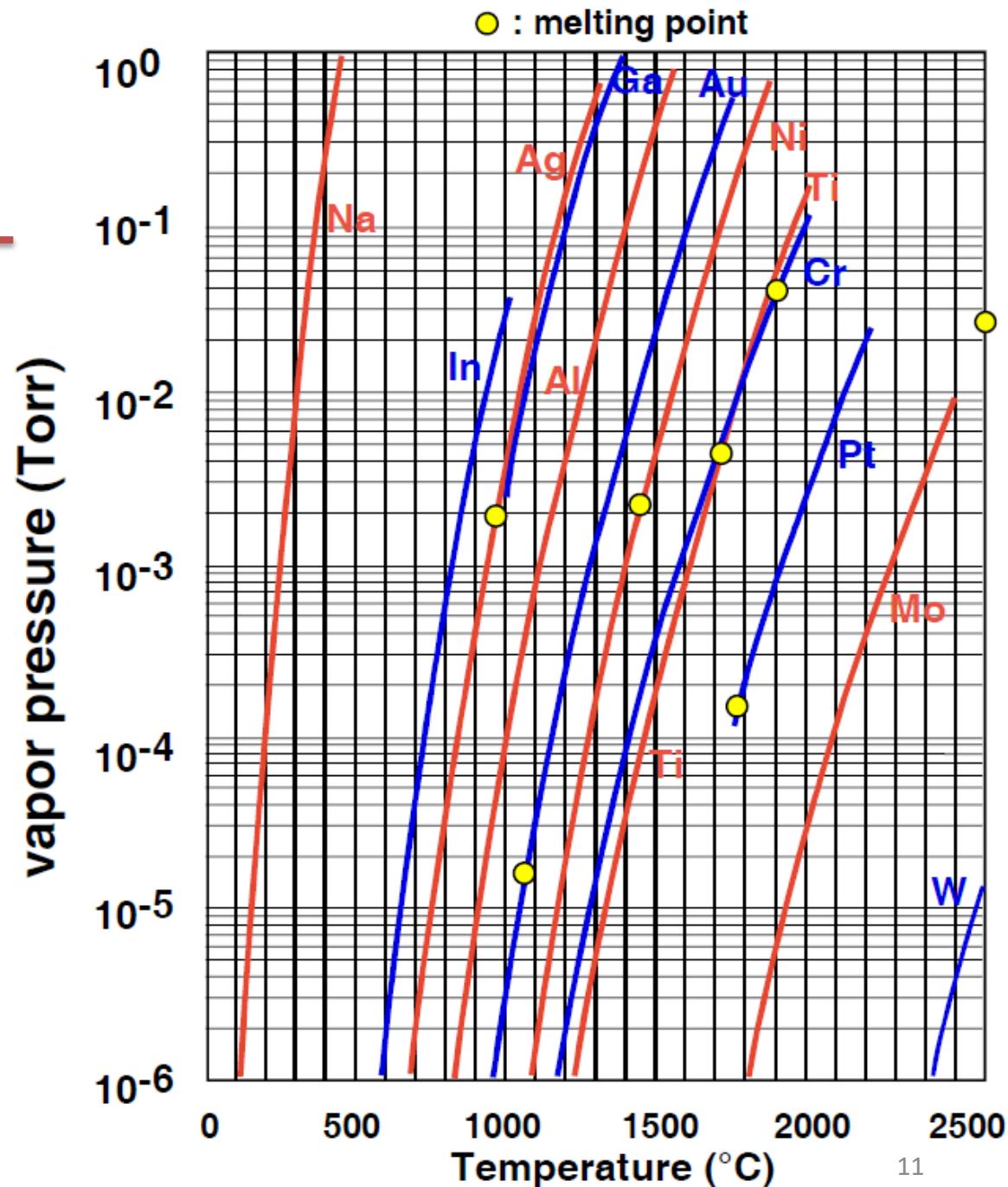
$$R \rightarrow R \times \frac{M}{N_A} = 5.83 \times 10^{-2} \sqrt{\frac{M}{T}} P \frac{\text{gram}}{\text{cm}^2 \cdot \text{sec}}$$

The orifice/opening appears as an evaporation source.



Vapor pressure of common materials

Vapor pressure of 0.1-10mTorr is typically used.



Evaporation characteristics of materials

Table 3-3
Evaporation Characteristics of Materials

Material	Minimum evaporation temperature ^a	State of evaporation	Recommended crucible material	Deposition rate (Å/s)	Power (kW) (e-beam ^b)
Aluminum	1010	Melts	BN	20	5
Al_2O_3	1325	Semimelts	BN, Al_2O_3	10	0.5
Antimony	425	Melts	Al_2O_3	50	0.5
Arsenic	210	Sublimes	Graphite, BeO	100	0.1
Beryllium	1000	Melts	Graphite, BeO	100	1.5
BeO		Melts		40	1.0
Boron	1800	Melts	Graphite, WC	10	1.5
BC		Semimelts		35	1.0
Cadmium	180	Melts	Al_2O_3 , quartz	30	0.3
CdS	250	Sublimes	Graphite	10	0.2
CaF_2		Semimelts		30	0.05
Carbon	2140	Sublimes		30	1.0
Chromium	1157	Sublimes		15	0.3
Cobalt	1200	Melts	Al_2O_3 , BeO	20	2.0
Copper	1017	Melts	Graphite, Al_2O_3	50	0.2
Gallium	907	Melts	Al_2O_3 , graphite		
Germanium	1167	Melts	Graphite	25	3.0
Gold	1132	Melts	BN	30	6.0
Indium	742	Melts	Al_2O_3	100	0.1
Iron	1180	Melts	Al_2O_3 , BeO	50	2.5
Lead	497	Melts	Al_2O_3	30	0.1
LiF	1180	Melts	Mo, W	10	0.15
Magnesium	327	Sublimes	Graphite	100	0.04
MgF_2	1540	Semimelts	Al_2O_3	30	0.01
Molybdenum	2117	Melts		4.0	
Nickel	1262	Melts	Al_2O_3	25	2.0
Permalloy	1300	Melts	Al_2O_3	30	2.0
Platinum	1747	Melts	Graphite	20	4.0
Silicon	1337	Melts	BeO	15	0.15
SiO_2	850	Semimelts	Ta	20	0.7
SiO	600	Sublimes	Ta	20	0.1
Tantalum	2590	Semimelts		100	5.0
Tin	997	Melts	Al_2O_3 , graphite	10	2.0
Titanium	1453	Melts		20	1.5
TiO_2	1300	Melts	W	10	1.0
Tungsten	2757	Sublimes	Al_2O_3	20	5.5
Zinc	250	Sublimes	Quartz	50	0.25
ZnSe	660	Sublimes	Mo		
ZnS	300	Sublimes	W		
Zirconium	1987	Melts		20	5.0

. From "Materials science of thin films" by Ohring, 2002.

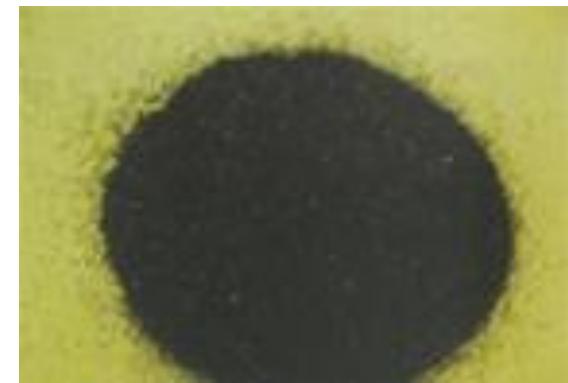
12^a Temperature (°C) at which vapor pressure is 10^{-4} torr.

^b For 10 kV, copper herth, source–substrate distance of 40 cm.

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Photos of source material for evaporation



Types of evaporation according to heating method

Three types:

Thermal evaporator – resistive heating, the only choice for evaporation of organic material.

Electron beam evaporator – heated by electron beam, most popular, more expensive than thermal evaporator.

Inductive heating (must be unpopular – I have never seen one).

Inductive heating:

Metal element is wound around crucible and RF power is run through coil.

RF induces eddy currents in the charge causing it to heat.

Eddy current is caused when a conductor is exposed to a changing magnetic field, that induces (changing) electric field.

In return, these circulating current create induced magnetic fields that oppose the change of the original magnetic field due to Lenz's law, causing repulsive or drag forces between the conductor and the magnet.

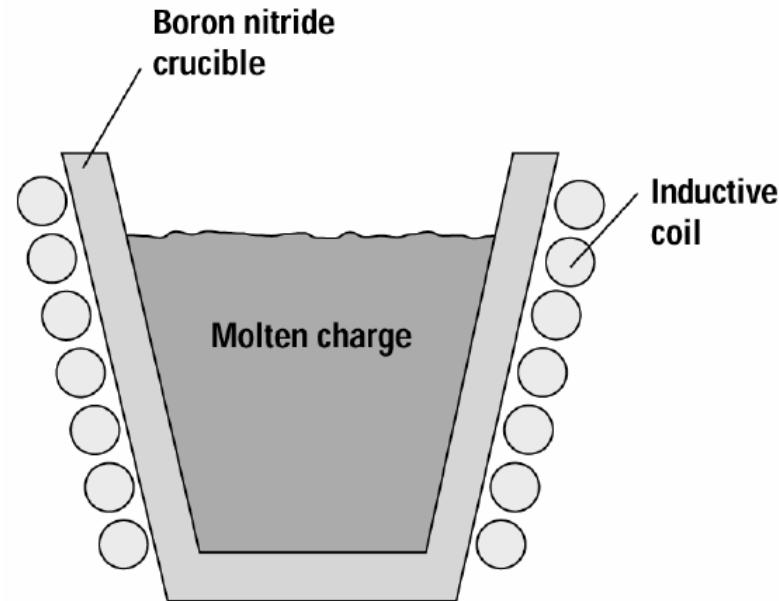
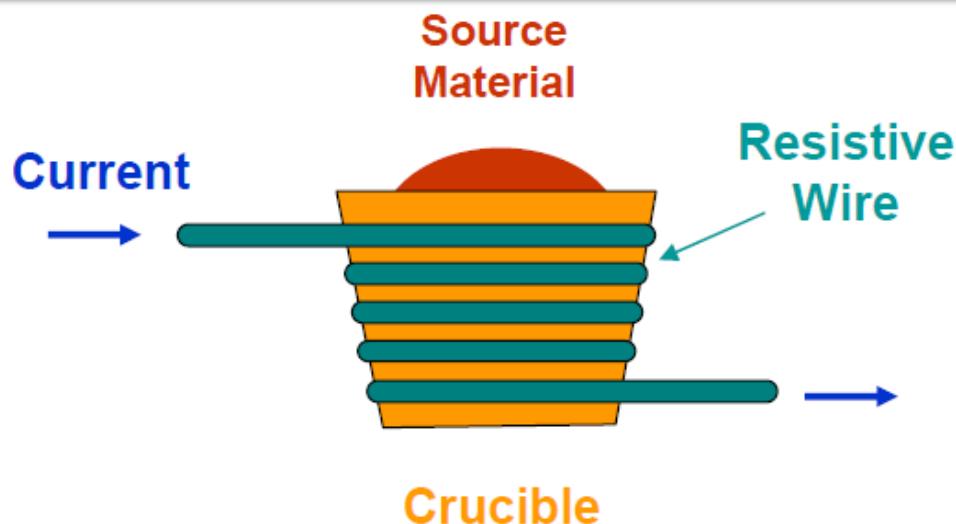


Figure 12.7 Example of an inductively heated crucible used to create moderately charged temperatures.

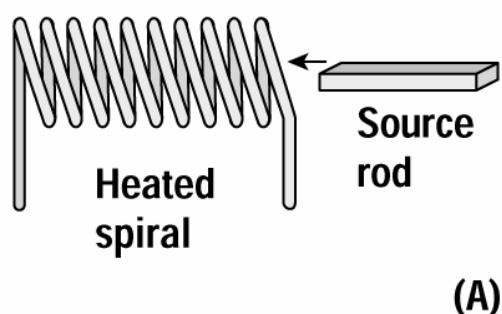
Thermal evaporation



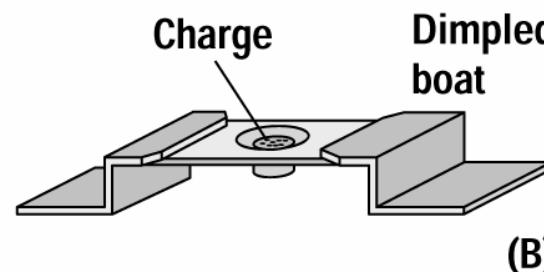
Widespread use for materials whose vapor pressure can be reasonable at 1600°C or below.

Common evaporant materials:

Au, Ag, Al, Sn, Cr, Sb, Ge, In, Mg, Ga;
CdS, PbS, CdSe, NaCl, KCl, AgCl, MgF₂,
CaF₂, PbCl₂.



(A)



(B)

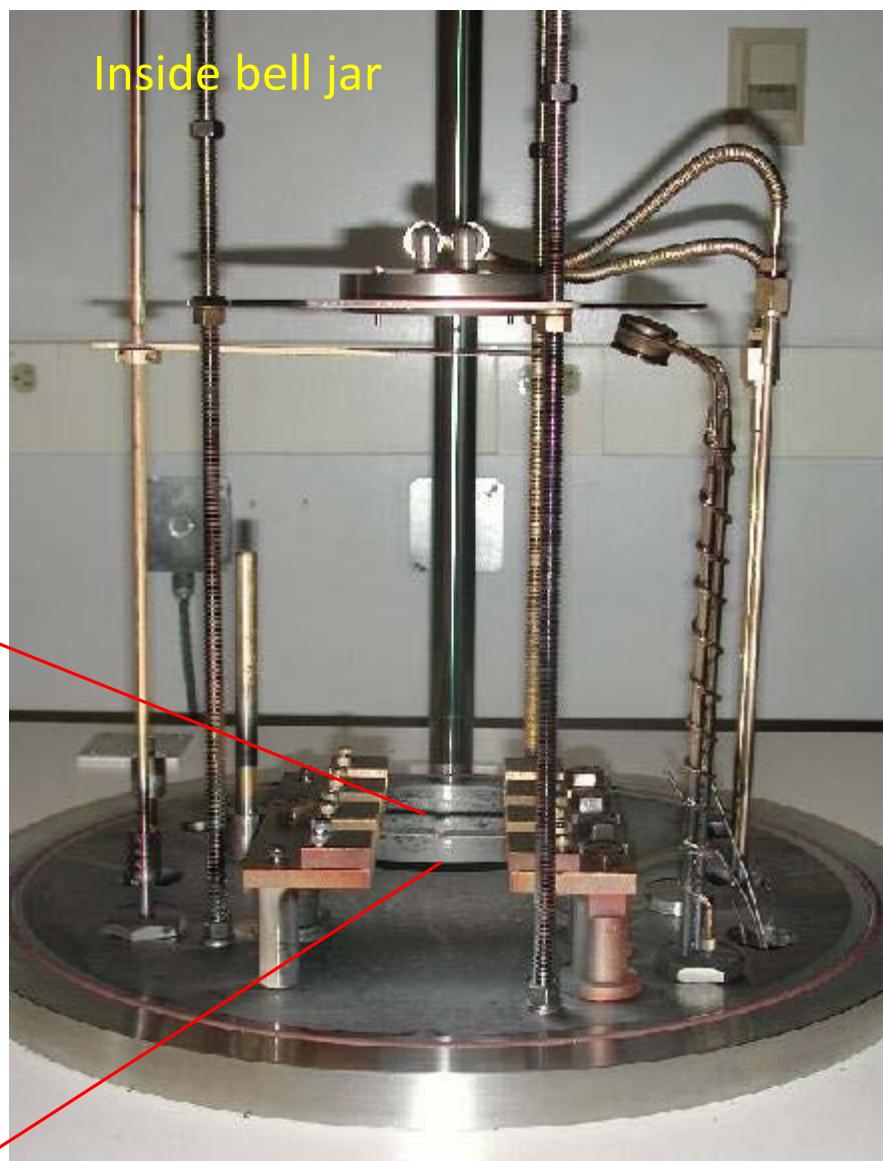
Figure 12.6 Resistive evaporator sources. (A) Simple sources including heating the charge itself and using a coil of refractory metal heater coil and a charge rod. (B) More standard thermal sources including a dimpled boat in a resistive media.

Photos of Sharon thermal evaporator

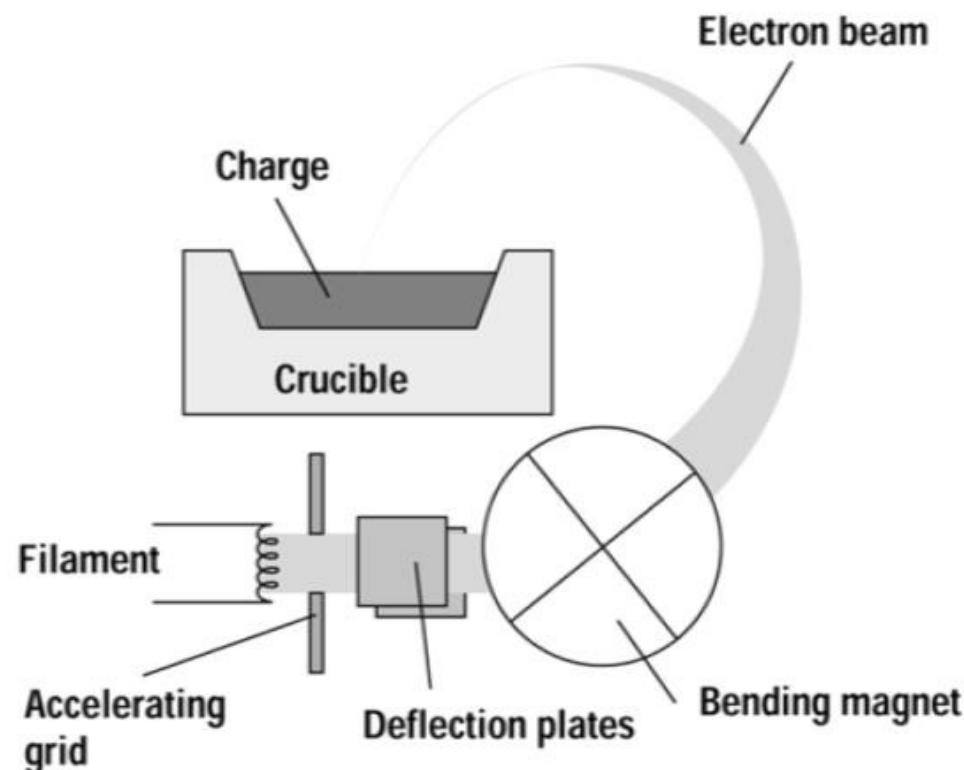
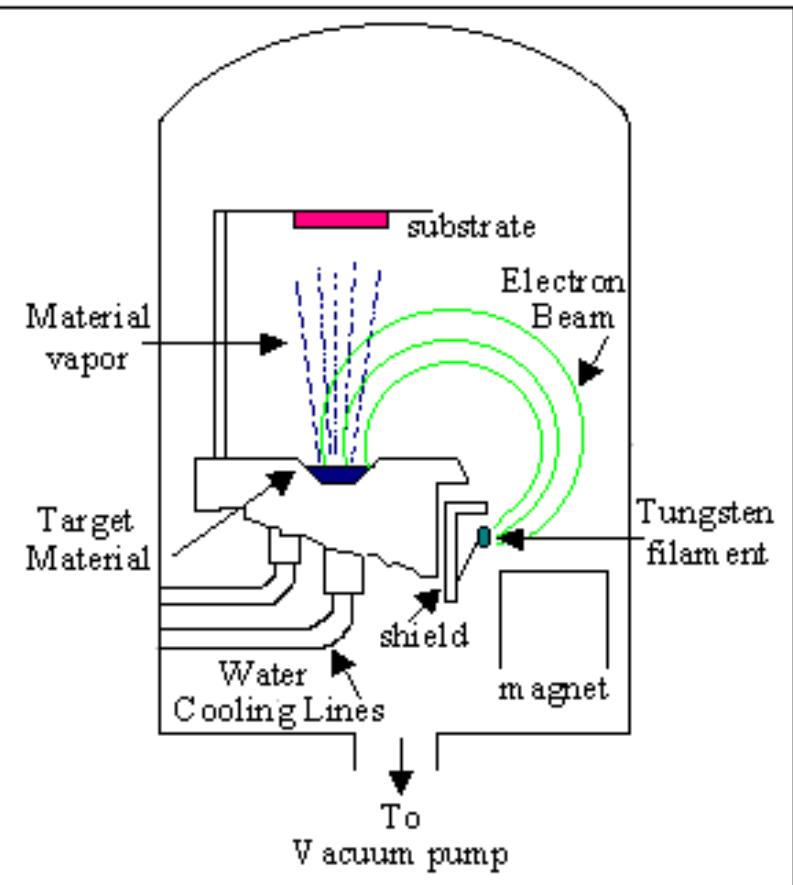
Bell jar



Inside bell jar



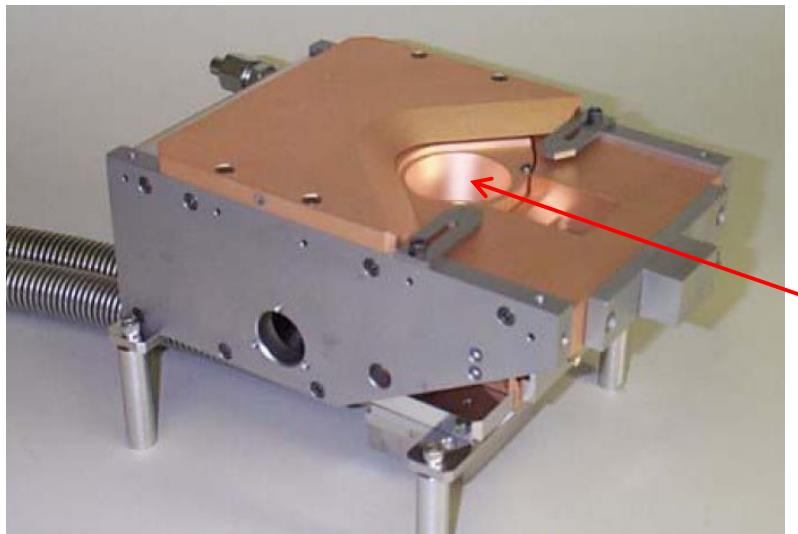
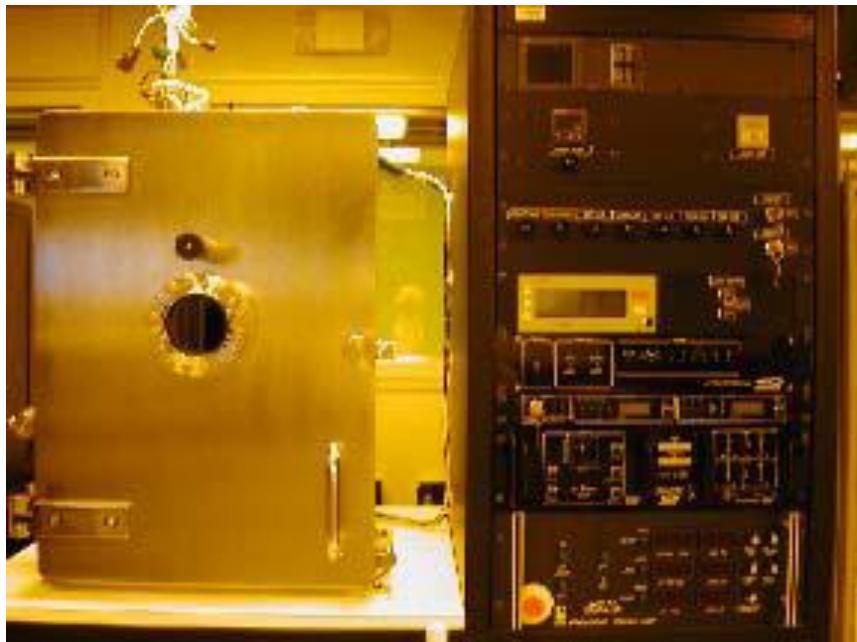
Electron beam evaporation



Deflection plates are to raster scan the beam across charge surface.

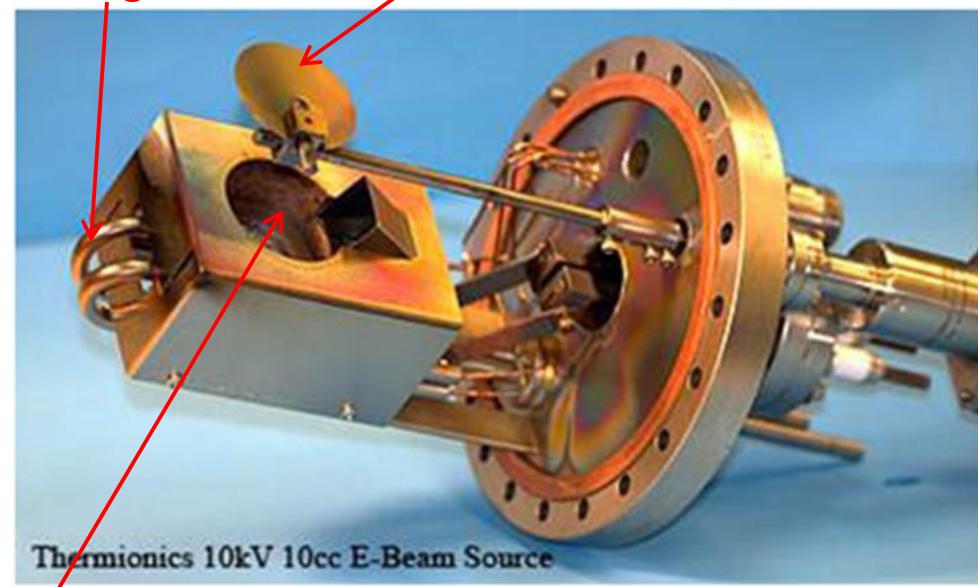
- Using a focused electron beam to heat and evaporate metals, electron temperature can be as high as 10,000 K. Electrons are accelerated by DC 10kV, and current 10s-100s of mA.
- Suitable for high T_{melt} metals like W, Ta, ...
- Evaporation occurs at a highly localized point near the beam bombardment spot on the source surface , so little contamination from the crucible (not hot, water cooled).

Photos of e-beam evaporator



Mechanical shutter:
Evaporation rate is set by temperature of source, but this cannot be turned on and off rapidly.

Cooling water Shutter



Put crucible
here

Heat conduction of the hearth
limits achievable temperature.

Power density: 10kV, up to 1.5A,
 $0.2\text{-}1\text{cm}^2 \rightarrow 15\text{-}75\text{kW/cm}^2$.

Comparison of thermal and e-beam evaporation

Deposition	Material	Typical Evaporant	Impurity	Deposition Rate	Temperature Range	Cost
Thermal	Metal or low melt-point materials	Au, Ag, Al, Cr, Sn, Sb, Ge, In, Mg, Ga CdS, PbS, CdSe, NaCl, KCl, AgCl, MgF ₂ , CaF ₂ , PbCl ₂	High	1 ~ 20 A/s	~ 1800 °C	Low
E-Beam	Both metal and dielectrics	Everything above, plus: Ni, Pt, Ir, Rh, Ti, V, Zr, W, Ta, Mo Al ₂ O ₃ , SiO, SiO ₂ , SnO ₂ , TiO ₂ , ZrO ₂	Low	10 ~ 100 A/s	~ 3000 °C	High

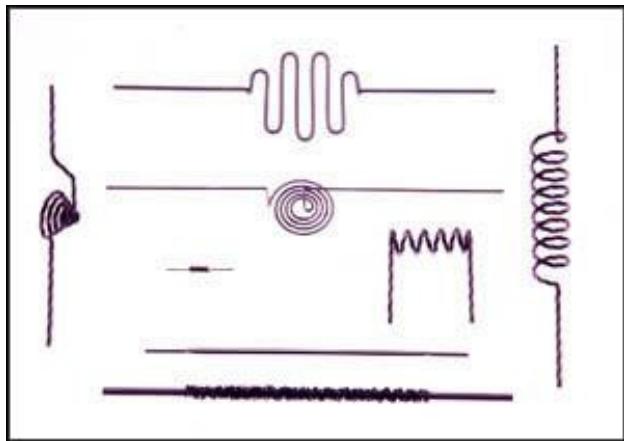
Thermal evaporation:

- Simple, robust, and in widespread use.
- Use W, Ta, or Mo filaments to heat evaporation source.
- Typical filament currents are 200-300 Amperes.
- Exposes substrates to visible and IR radiation.
- Contamination from heated boat/crucible.

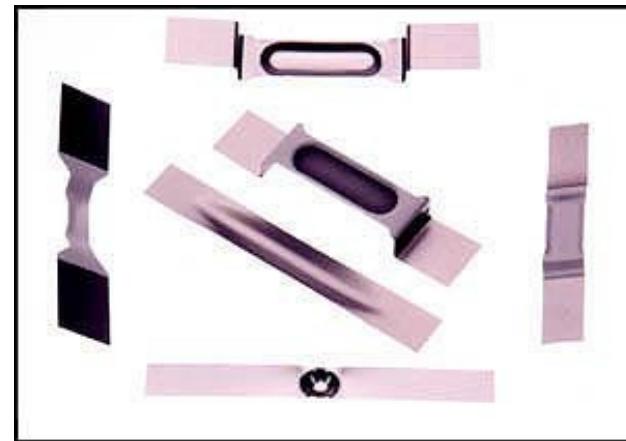
Electron beam evaporation:

- More complex, but extremely versatile, virtually any material.
- Less contamination, less heating to wafer (as only small source area heated to very high T).
- Exposes substrates to secondary electron radiation.
- X-rays can also be generated by high voltage electron beam.
- Since x-rays will damage substrate and dielectrics (leads to trapped charge), e-beam evaporators cannot be used in MOSFET.

Popular heating “containers” for evaporation source



Resistors (put source rod inside coil)

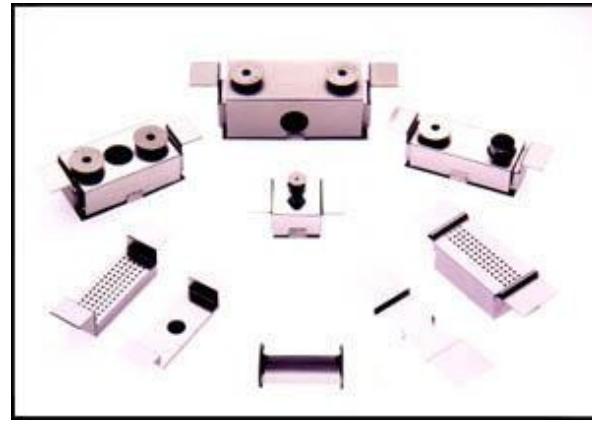


Heating boat (open top)



Crucibles

(Crucible is the only choice for e-beam evaporator.
Can also be used for thermal evaporator)



Box with small opening
(Knudsen cell!)

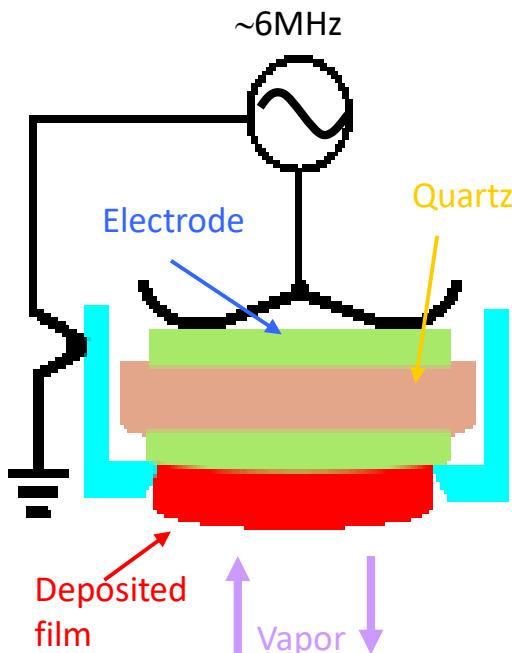
Typical boat/crucible material

Refractory Metals		
Material	Melting Point (°C)	Temperature for 10-mtorr Vapor Pressure (P _e) (°C)
Tungsten (W)	3380	3230
Tantalum (Ta)	3000	3060
Molybdenum (Mo)	2620	2530
Refractory Ceramics		
Graphitic Carbon (C)	3799	2600
Alumina (Al ₂ O ₃)	2030	1900
Boron Nitride (BN)	2500	1600

Considerations: thermal conductivity, thermal expansion, electrical conductivity, wetting and reactivity.

Graphite crucible is most popular (cheapest), but should avoid cracking the crucible due to stress/temperature gradients.

How to monitor film thickness during evaporation?



Quartz Crystal Micro-balance (QCM): (similar idea to quartz clock)

- Quartz is a piezoelectric material.
- With a high frequency AC voltage activation, the amplitude of vibration is maximum at resonance frequency.
- This resonance frequency will shift when film is deposited on its surface.
- Thus by measuring frequency shift Δf , one can measure film thickness with sub-Å accuracy.

Evaporation issues: step coverage

Step coverage is poor (line of sight and sticking coefficient $S_c \approx 1$).

Heating can decrease S_c , but may change film properties (structure – large grain).

Therefore, due to poor step coverage (and x-ray damage), rarely used in IC fabrication.

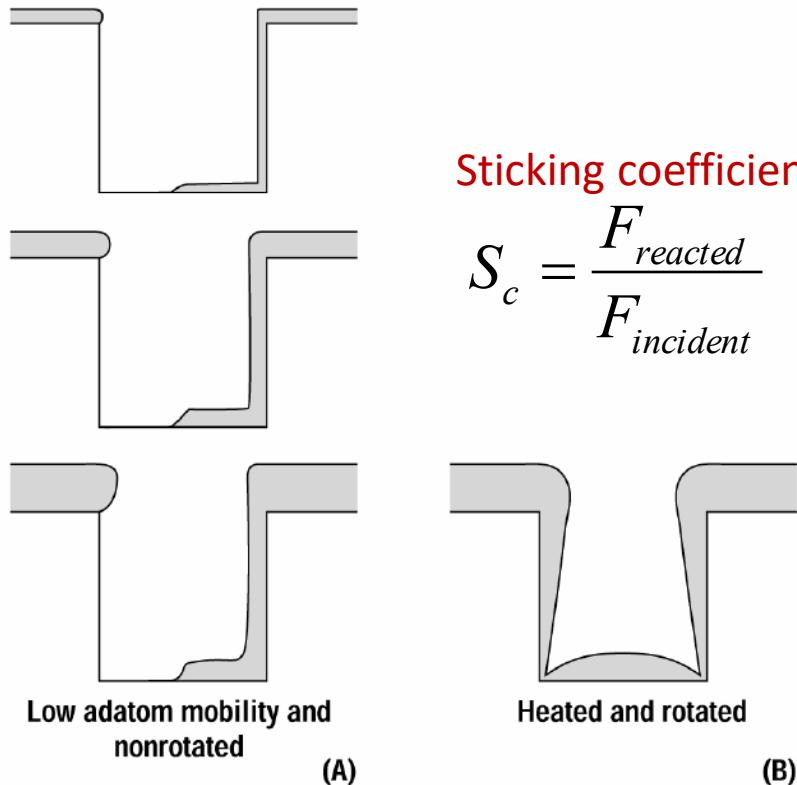
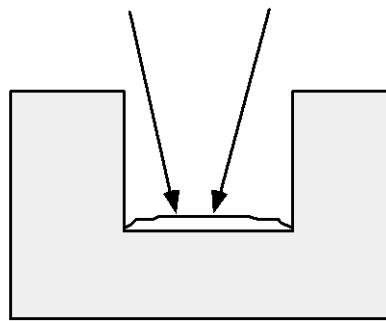


Figure 12.5 (A) Time evolution of the evaporative coating of a feature with aspect ratio of 1.0, with little surface atom mobility (i.e., low substrate temperature) and no rotation. (B) Final profile of deposition on rotated and heated substrates.

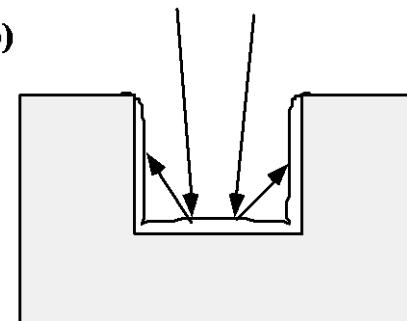
Sticking coefficient:

$$S_c = \frac{F_{reacted}}{F_{incident}}$$

a)



b)



$$S_c = 1$$

$$S_c < 1$$

- The depositing species have a high S_c so that they are deposited where they first strike.
- The depositing species have a low S_c , so that many are re-emitted and re-deposited (or migrated) elsewhere on the topography, such as on the sidewalls.

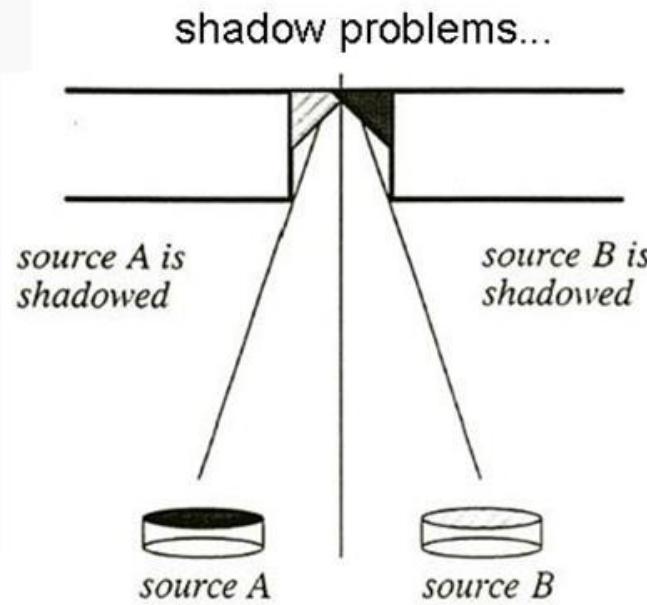
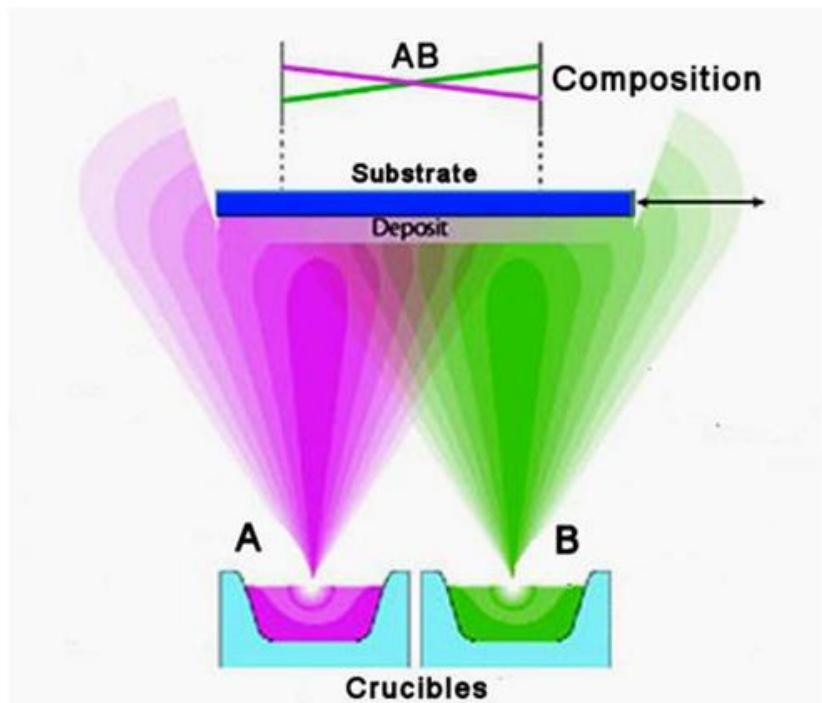
Evaporation issues: alloy evaporation

Stoichiometrical problem of evaporation:

- Compound material breaks down at high temperature.
- Each component has different vapor pressure, therefore different deposition rate, resulting in a film with different stoichiometry compared to the source.

One solution is co-evaporation (use two e-beam guns), see below.

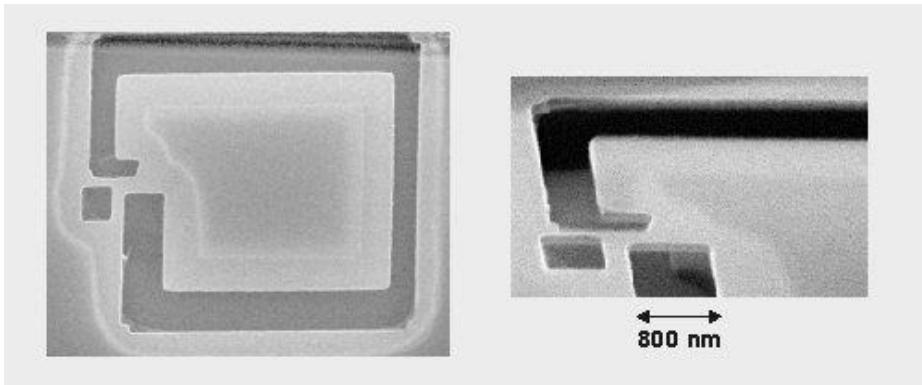
Lateral composition change



Shadow evaporation

Poor step coverage could be useful.

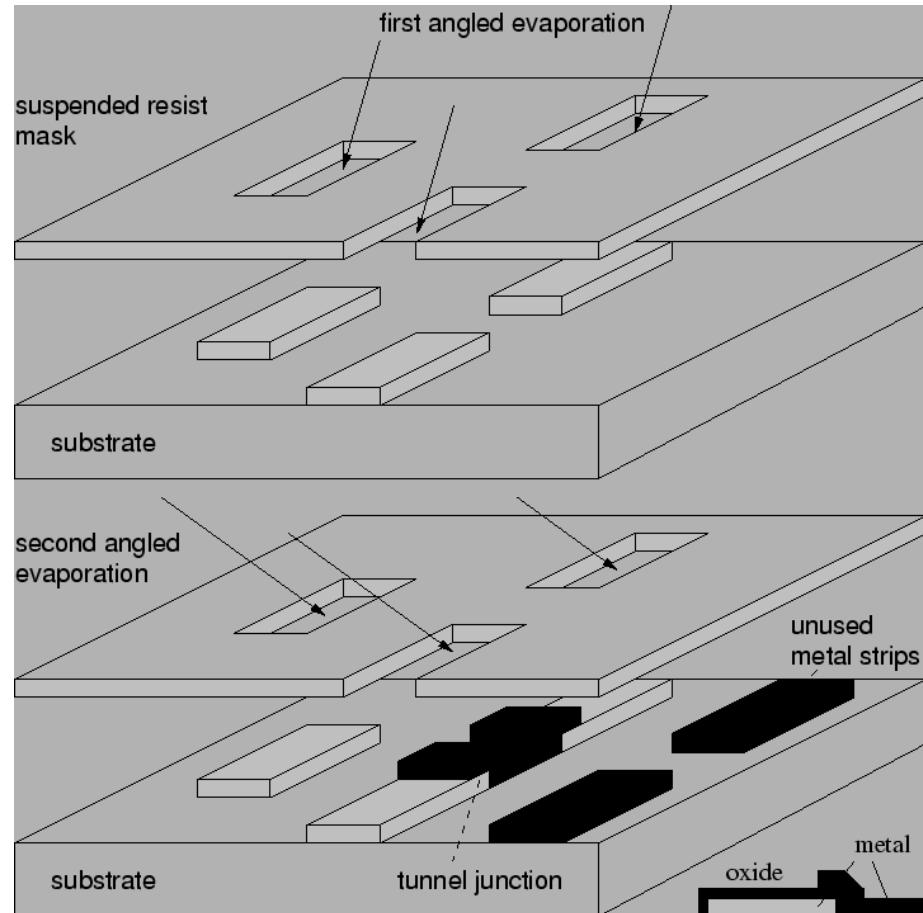
Shadow/angle evaporation is routinely used for nanofabrication.



The pictures above show shadow-masks for niobium rings containing a Josephson junction, prior to evaporation.

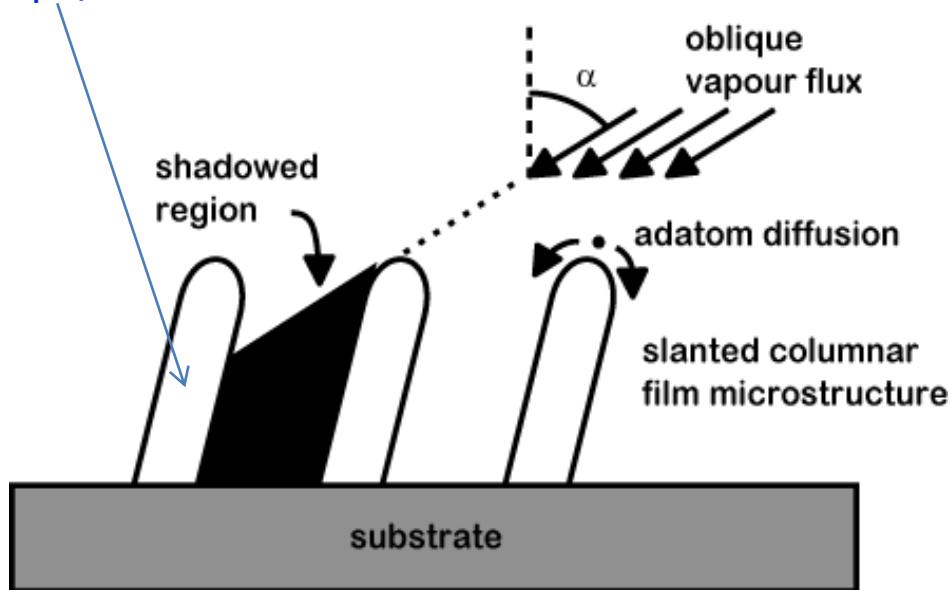
The metals are evaporated under different angles without breaking the vacuum.

The mask consists of Germanium while the sacrificial layer below the mask is made of high temperature capable plastic (polyether sulphone).



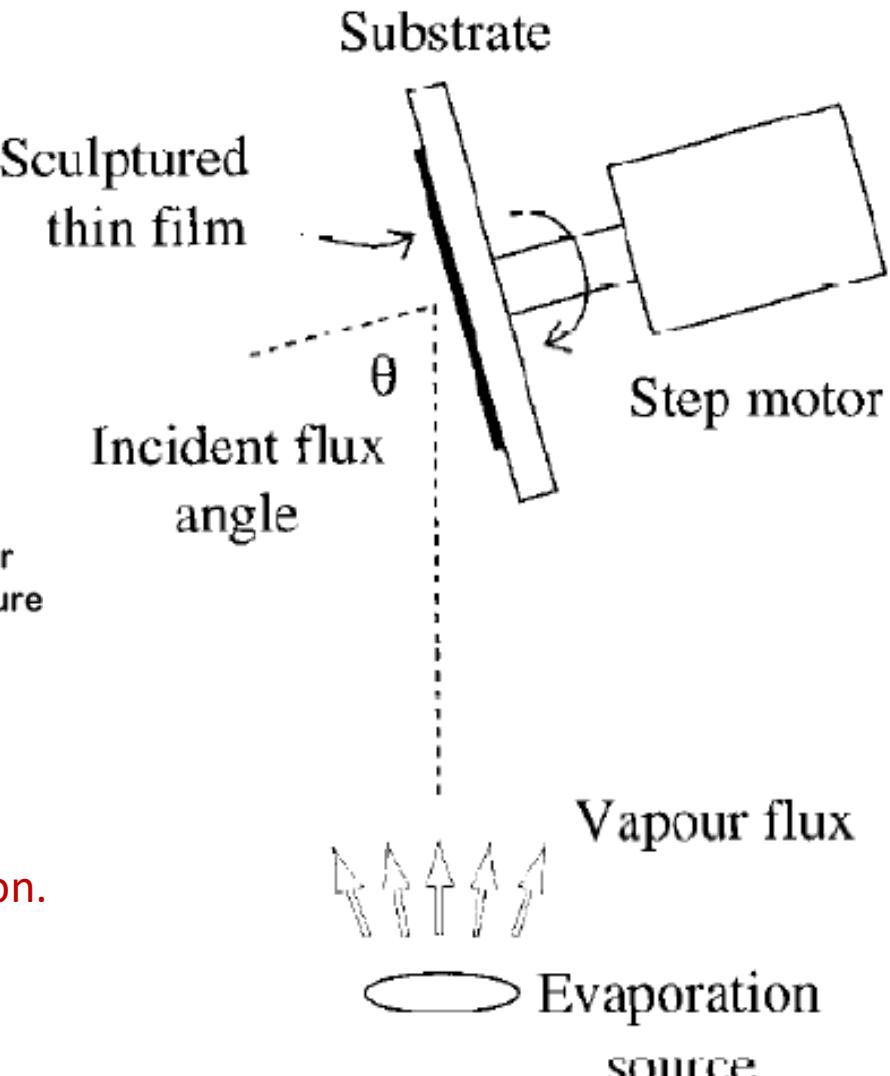
GLAD (glancing angle deposition): self assembly of film

This is actually a tiny grain (cluster of atoms) formed (nucleated) at the first moment of film deposition, at random position with random shape/size.

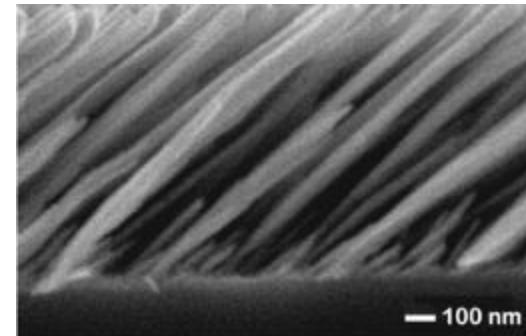
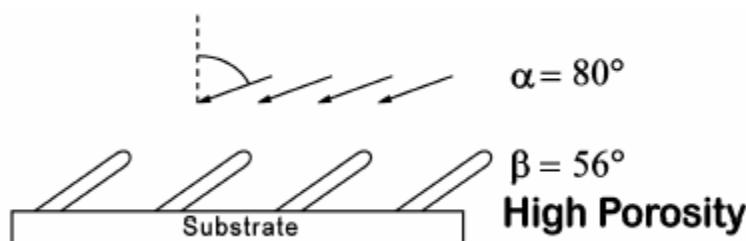
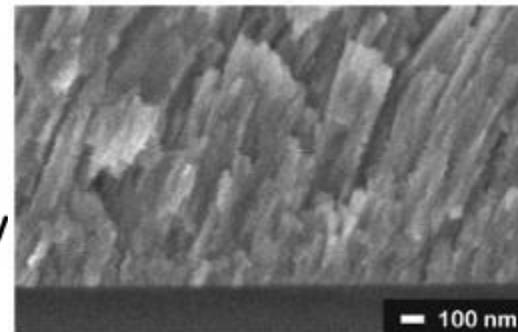
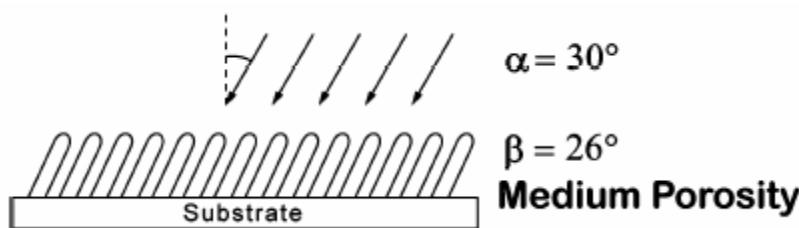
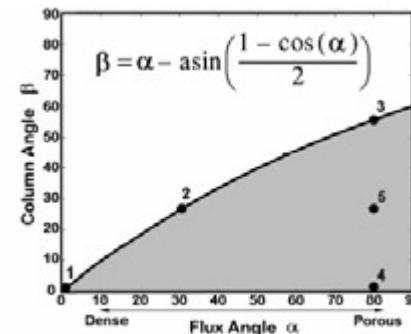
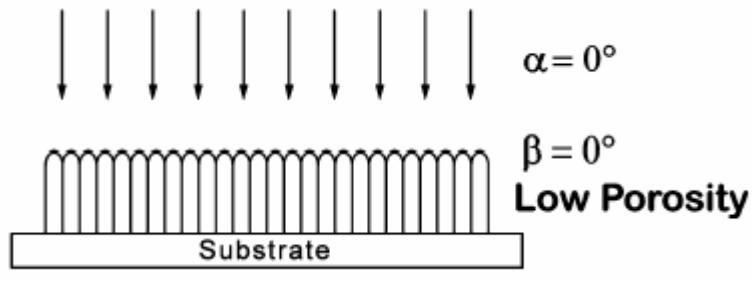


Angle $>80^\circ$, with (or without) substrate rotation.
Self assembly mainly due to shadowing effect
that magnifies the otherwise grain structures.

Invented by Michael Brett from University of Alberta
<http://www.ece.ualberta.ca/~glad/lab.html>

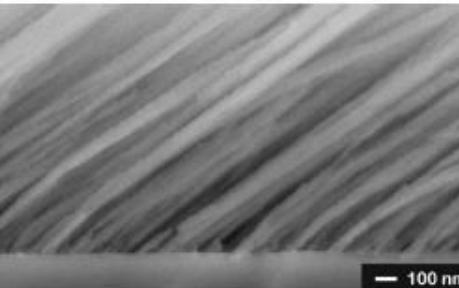
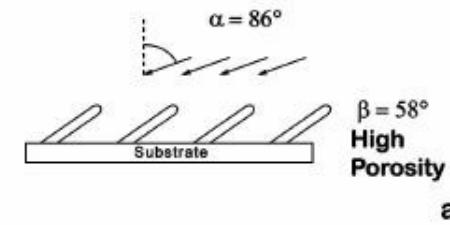


Glancing angle deposition (GLAD)

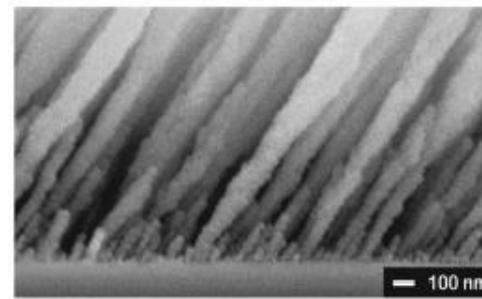
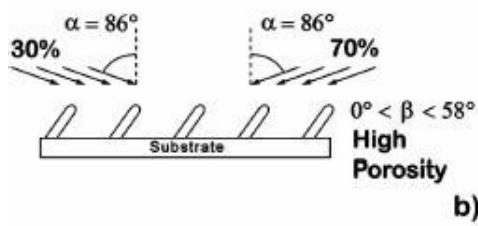


Stationary substrate (i.e. no rotation), one evaporation source, Cr films.

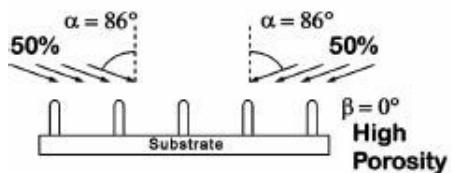
Glancing angle deposition (GLAD)



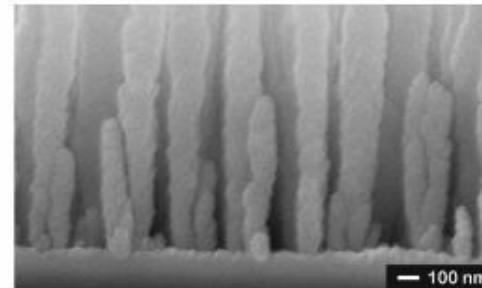
a)



b)



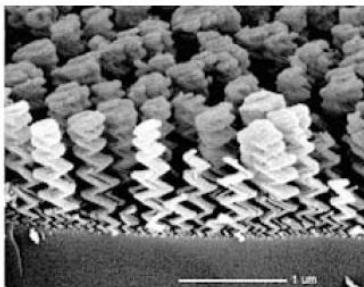
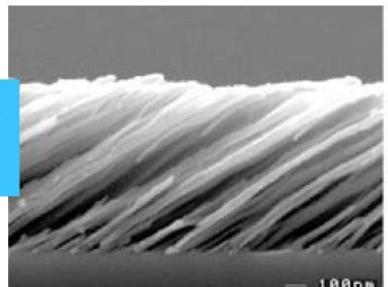
c)



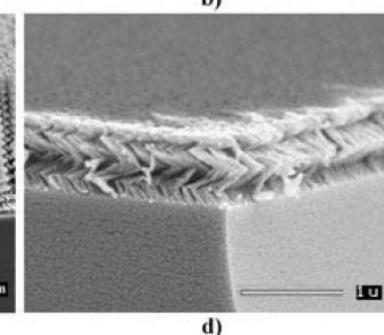
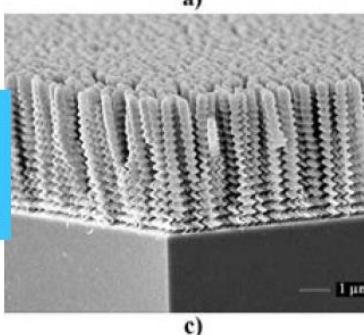
Stationary substrate, two evaporation sources, SiO_2 films.
Independent control of column angle and film porosity.
The porosity is constant, the column angle β is controlled between the inclined and vertical angles.

Various nanostructures obtained in GLAD thin films

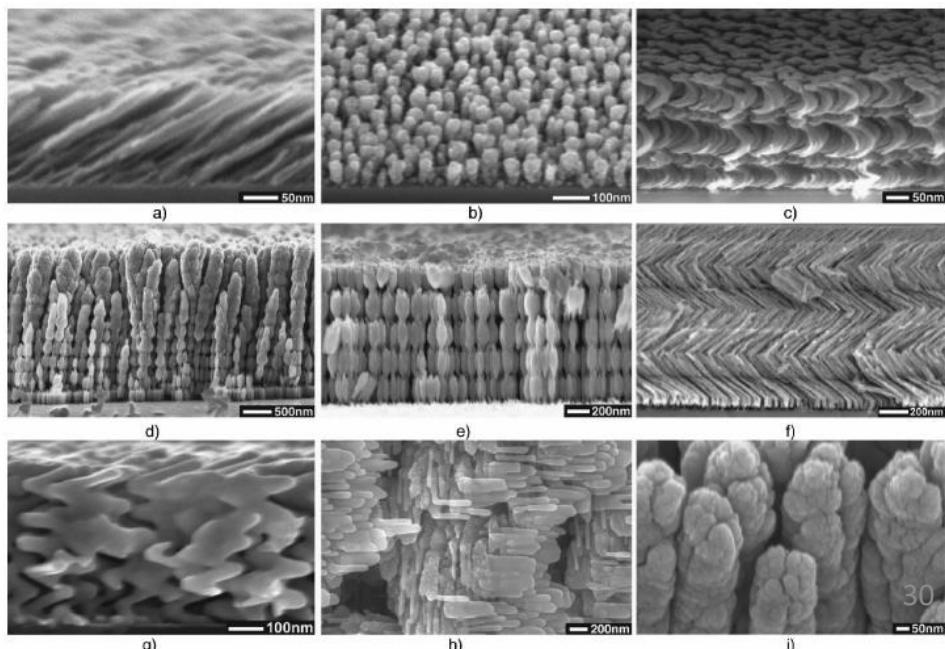
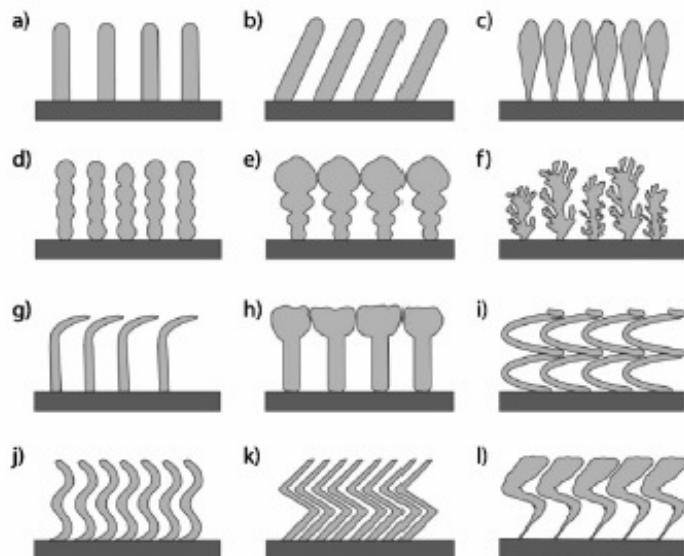
flux arrived from the right for the entire deposition



the direction of arrival of the flux was alternated from the left and right 12 times during deposition



a combination of the techniques used in b) and c). The substrate was rotated in 90 degree steps during deposition.



Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Sputter deposition

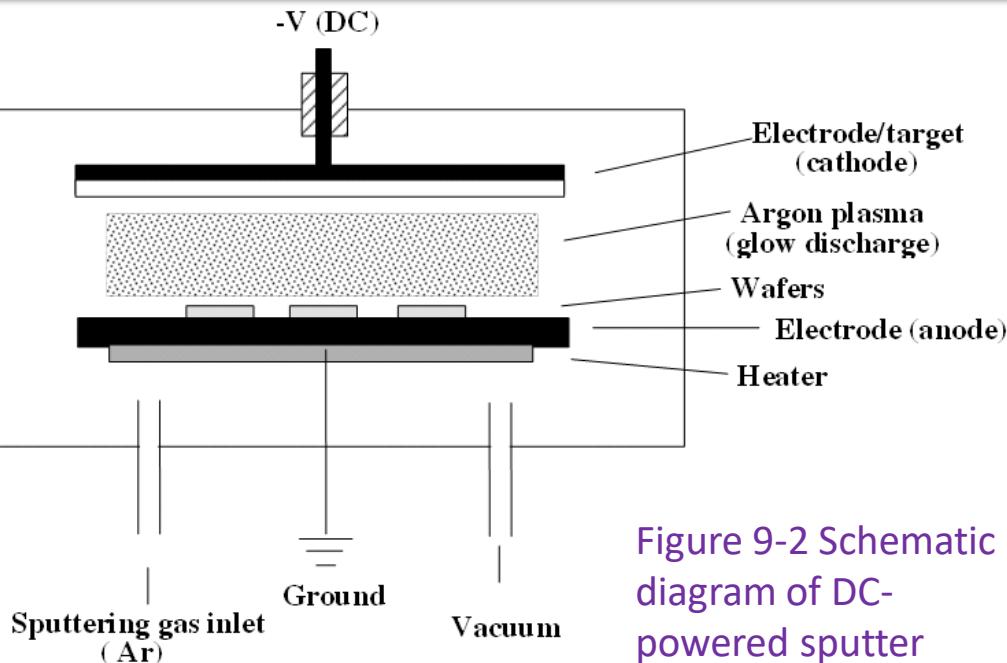
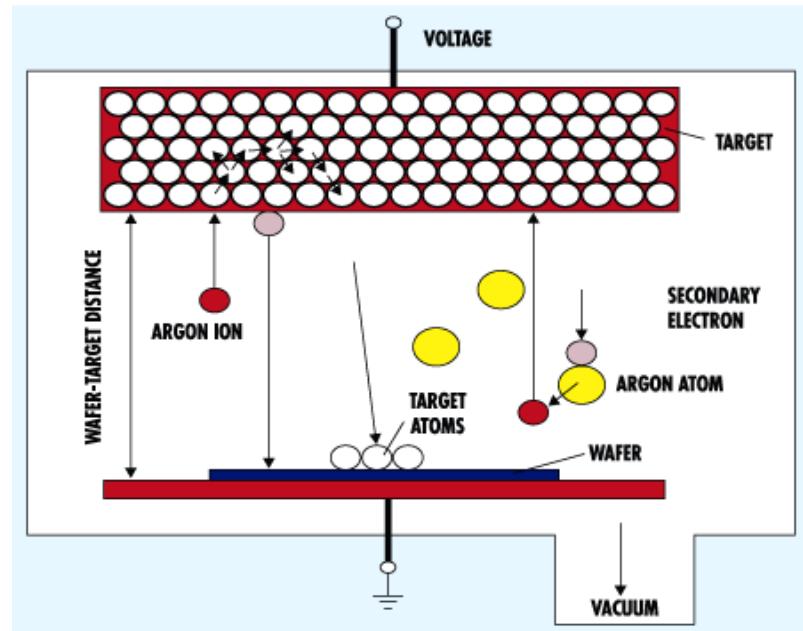


Figure 9-2 Schematic diagram of DC-powered sputter deposition equipment.



- Plasma is needed to make the gas conductive, and generated ions can then be accelerated to strike the target.
- Higher pressures than evaporation: 1-100 mTorr.
- Better at depositing alloys and compounds than evaporation.
- The plasma contains \approx equal numbers of positive argon ions and electrons as well as neutral argon atoms. Typically only <0.01% atoms are ionized!

Sputtering process

- Sputtering process can be run in DC or RF mode (insulator must be run in RF mode)
- Major process parameters:
 - Operation pressure (~1-100mTorr)
 - Power (few 100W)
 - For DC sputtering, voltage -2 to -5kV.
 - Additional substrate bias voltage.
 - Substrate temperature (20-700°C)

In addition to IC industry, a wide range of industrial products use sputtering: LCD, computer hard drives, hard coatings for tools, metals on plastics.

It is more widely used for industry than evaporator, partly because that, for evaporation:

- There are very few things (rate and substrate temperature) one can do to tailor film property.
- The step coverage is poor.
- It is not suitable for compound or alloy deposition.
- Considerable materials are deposited on chamber walls and wasted.



Targets for sputter deposition.

Sputter deposition advantages

Advantages:

- Capable of in-situ cleaning prior to film deposition by reversing the potential on the electrodes .
- Better film quality and step coverage than evaporation.
- This is partly because adatoms are more energetic, and film is ‘densified’ by in-situ ion bombardment, and it is easier to heat up to high T than evaporation that is in vacuum.
- More reproducible deposition control – same deposition rate for same process parameters (not true for evaporation), so easy film thickness control via time.
- Can use large area targets for uniform thickness over large substrates.
- Sufficient target material for many depositions.
- No x-ray damage.

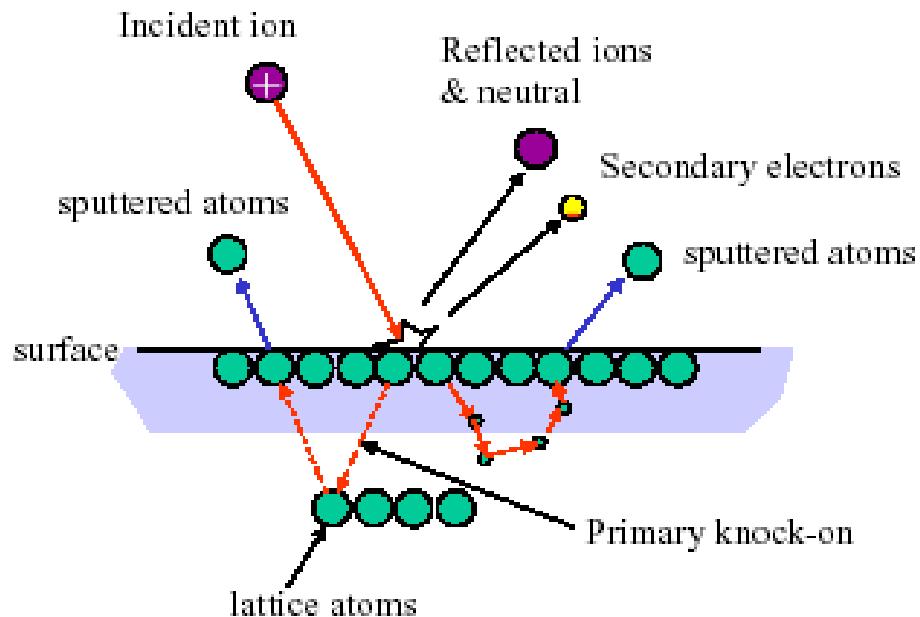
Disadvantages:

- Substrate damage due to ion bombardment or UV generated by plasma.
- Higher pressures 1 –100 mtorr ($< 10^{-5}$ torr in evaporation), more contaminations unless using ultra clean gasses and ultra clean targets.
- Deposition rate of some materials quite low.
- Some materials (e.g., organics) degrade due to ionic bombardment.
- Most of the energy incident on the target becomes heat, which must be removed.

Mechanisms of sputtering and alloy sputtering

The ion impact may set up a series of collisions between atoms of the target, possibly leading to the ejection of some of these atoms. This ejection process is known as sputtering.

Here we are interested in sputter deposition. Of course sputter can also be used as an etching method (the substrate to be etched will be the 'target'), which is called sputter etching.



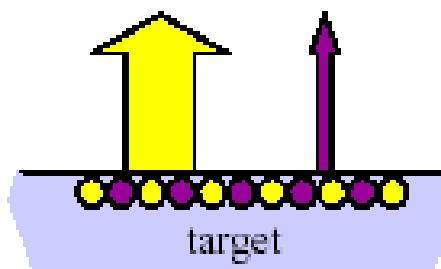
Alloys

Unlike evaporation, composition of alloy film is approximately the same as target.

Target NOT melted, slow diffusion (no material flow) mixing.

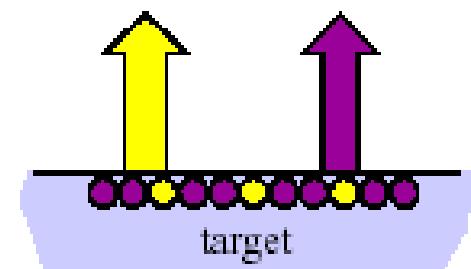
When target reaches steady state, surface composition balances sputter yield.

High yield material low yield material



Before surface equilibrium

High yield material Low yield material



After surface equilibrium

DC plasma

Plasma is ionized gas, with nearly equal number of ions and electrons, plus neutrals (un-ionized molecules including those at ground state and excited state; free radicals such as atomic O, H, F – but no free radicals for Ar plasma).

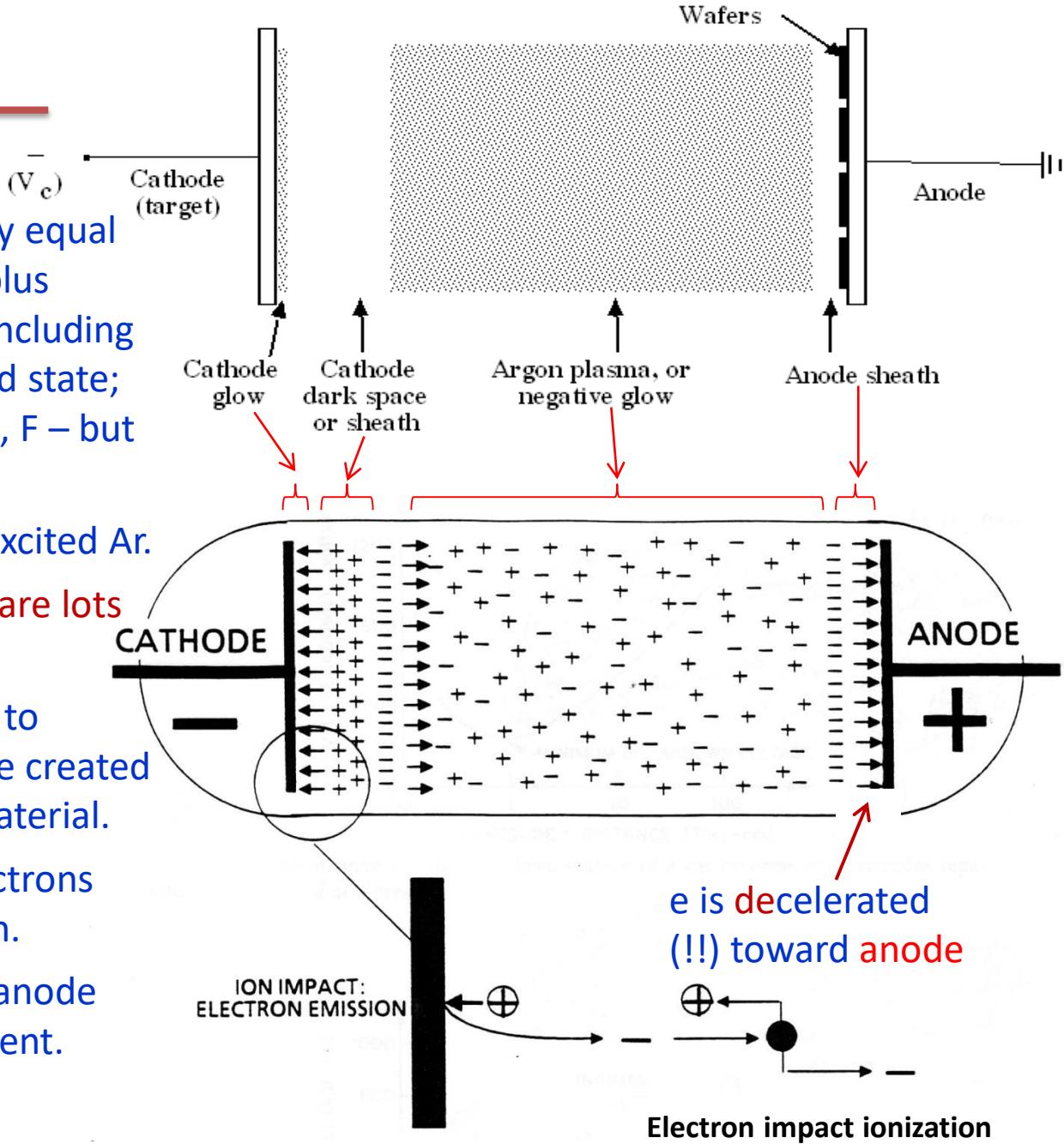
Glow is due to de-excitation of excited Ar.

So glow only exists where there are lots of electrons to excite Ar.

Cathode glow region: very close to cathode, secondary electrons are created by Ar bombardment of target material.

Cathode dark space/sheath: electrons pass too fast with little excitation.

Anode sheath: electrons lost to anode due to its faster *random* movement.



Explanation of DC plasma structure

Different velocities in a plasma:

Thermal energy random movement of Ar – 400 m/sec, order $(k_B T/m_{Ar})^{1/2}$.

Thermal energy random movement of electron – 10000 m/sec.

Velocity of Ar with energy 100eV – 20000 m/sec.

Velocity of electrons with energy 100eV – 600000 m/sec.

Thus plasma is highly conducting due to fast electrons – very little voltage drop in the plasma area where electrons are rich.

Voltage drop is only possible near the electrodes where electrons may lost to the electrode.

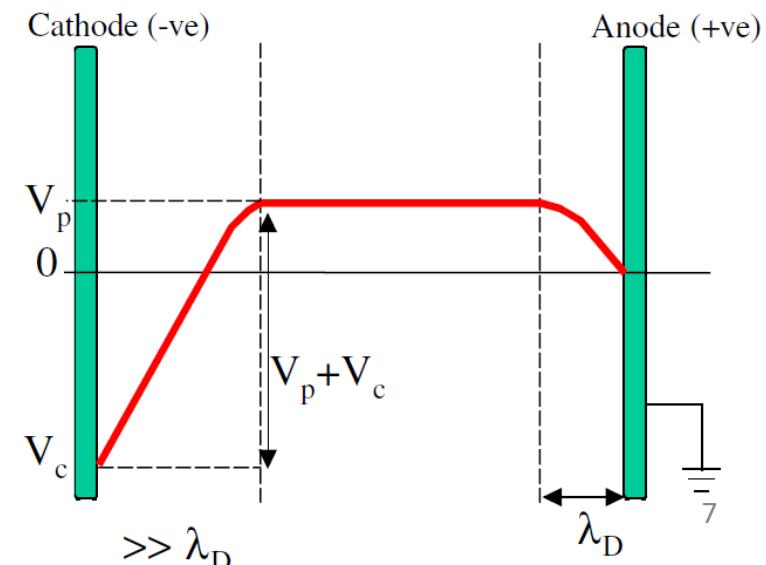
Even without applied voltage (assume plasma still exist), voltage drop may still exist due to faster *random* electrons movement that leads to their lost to electrode.

Therefore, the plasma is always positively biased relative to any electrode or anything (floating or not) inside the plasma.

This positive bias will accelerate positive Ar ion to strike the electrode.

But the bias V_p near the anode is very small (~10V), so no significant sputtering of the substrate.

The total bias (V_p plus applied voltage) is very high, leading to sputtering of cathode (target).

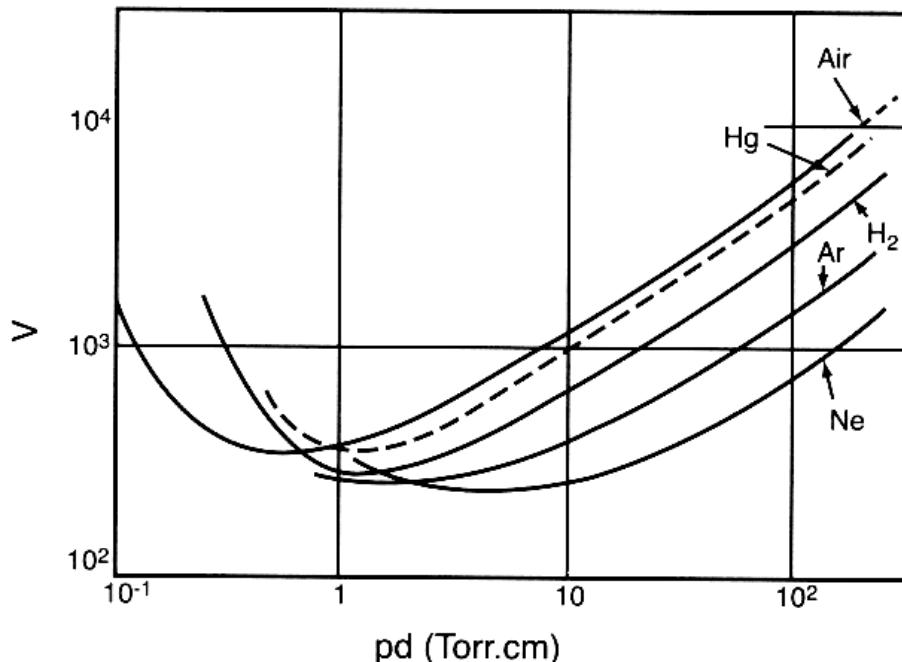


Requirement for self-sustained discharge (plasma)

Ions make (secondary) electrons when they bombard the target, and electrons make ions when they collide with Ar → self sustained discharge.

Condition for *sustaining* plasma: $pd > 0.5$ (cm·Torr).

For instance, typical target-substrate spacing $d \sim 10\text{cm}$, need $p > 50\text{mTorr}$ (actually sputter deposition is usually conducted at <10mTorr, due to magnetron...).



Condition for *igniting* the plasma.

Too large $P \times d$ leads to too many collisions that prevent electron energy buildup.

Too small $P \times d$, there will be too few collisions (electron just goes to the wall without ionizing a molecule or atom), and too few ions to bombard and generate secondary electrons.

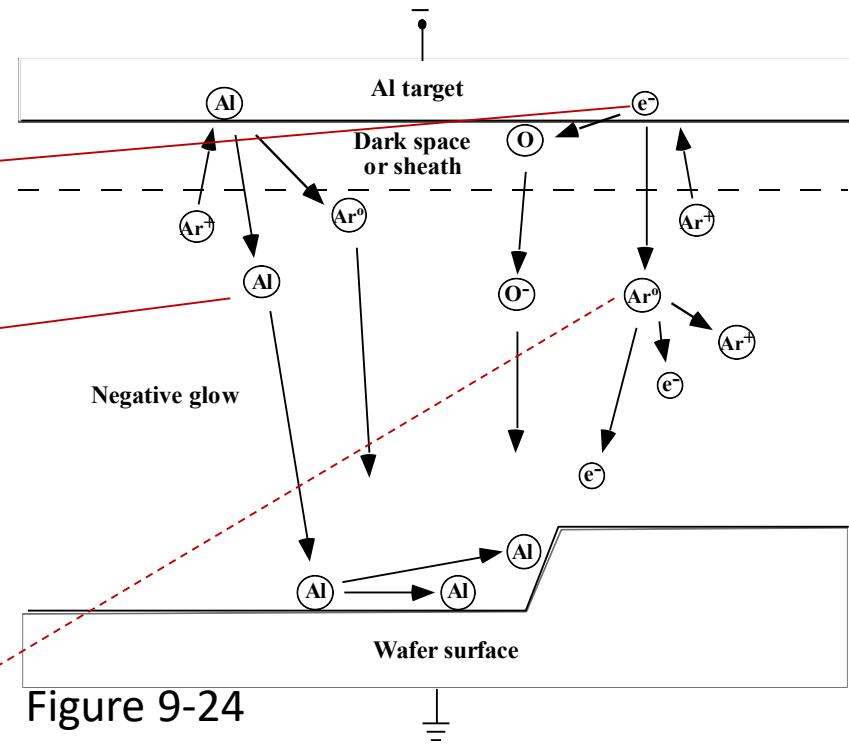
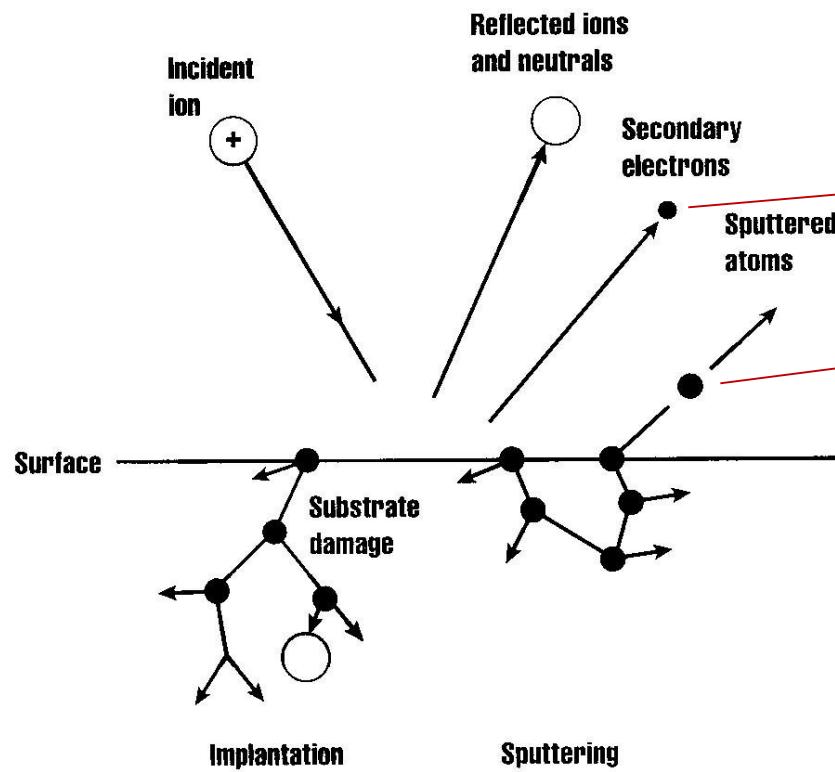
Once the plasma is ignited, it is very conductive, thus voltage drops to order 100 V only.

Figure 2.7 The dc breakdown voltage as a function of gas pressure P and electrode spacing d for plane parallel electrodes in air and some other gases. Such curves are determined experimentally and are known as *Paschen curves*.

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film property.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.

Sputtering process



After collision ionization, there are now TWO free electrons. This doubles the available electrons for ionization. This ongoing doubling process is called "impact ionization", which sustains a plasma.

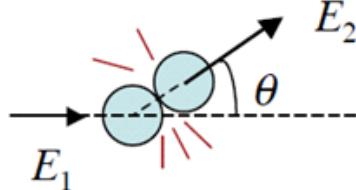
On the left side, sputter off an Al atom. On the right side, generate secondary electrons (e^-), which are accelerated across the sheath region and 1) ionize/excite an Ar; or 2) ionize an impurity atom, here O, to generate O⁻. This O⁻ is accelerated toward substrate and may go into the film (bad).

Sputtering process

- Energy of each incoming ion is 500-1000eV. Energy of sputtered atoms is 3-10eV.
- Thus, the sputtering process is very inefficient from the energy point of view, 95% of incoming energy goes to target heating & secondary electron.
- High rate sputter processes need efficient cooling techniques to avoid target damage from overheating (serious problem).
- The sputtered species, in general, are predominantly neutral.
- The energies of the atoms or molecules sputtered at a given rate are about one order of magnitude higher than those thermally evaporated at the same rate, which often lead to better film quality.
- However, since sputtering yields are low and the ion currents are limited, sputter-deposition rates are invariably one to two orders of magnitude lower compared to thermal evaporation rates under normal conditions.

Sputtering yield

Elastic energy transfer



E_2 is greatest for $M_1=M_2$.

There is also inelastic energy transfer, which leads to secondary electrons emission...

$$\frac{E_2}{E_1} \propto \frac{4M_1M_2}{(M_1 + M_2)^2} \cos^2 \theta$$

$$Y = \frac{\text{sputtered atoms}}{\text{bombing ions}} = \alpha \frac{Mm}{(M+m)^2} \frac{E_m}{U_M}$$

M : mass of target atom

m : mass of bombing ion

E_m : kinetic energy of bombing ion

U_M : Bonding energy of target metal

α : depends on striking /incident angle

- Sputter yield Y : the number of sputtered atoms per impinging ion.
- Obviously, the higher yield, the higher sputter deposition rate.
- Sputter yield is typically 1-10.
- The sputter yield depends on: (a) the energy of the incident ions; (b) the masses of the ions and target atoms; (c) the binding energy of atoms in the solid; and (d) the incident angle of ions.
- The yield is rather insensitive to the target temperature.

Dependence of sputter yield on ion energy

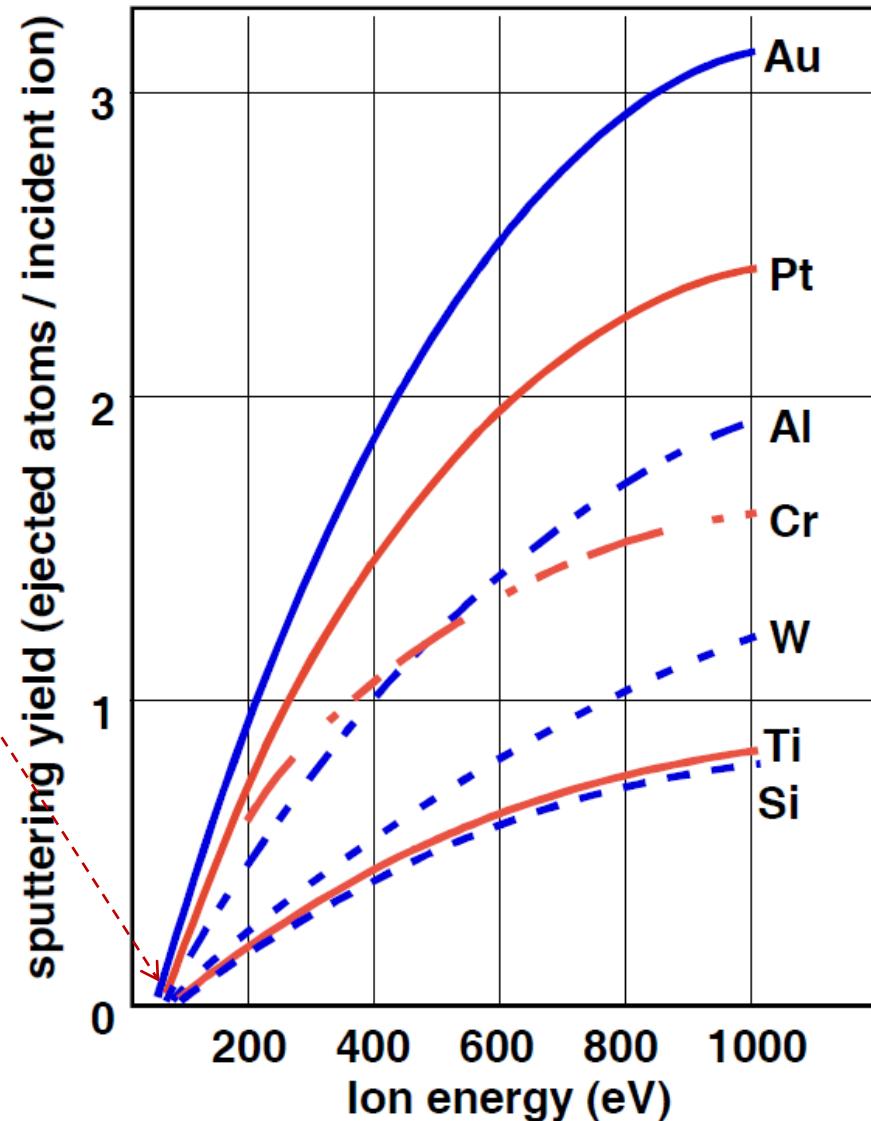
A threshold energy for the release of an atom from the target exists, below which the atom is not “sputtered”.

The yield increases with the energy.

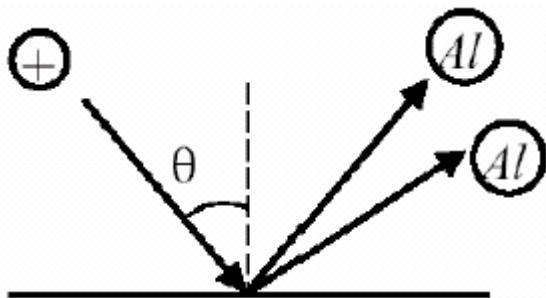
For higher energies, the yield approaches saturation, which occurs at higher energies for heavier bombarding particles.

e.g.: $\text{Xe}^+ \sim 100\text{keV}$ and $\text{Ar}^+ \sim 20\text{KeV}$ for saturation.

Sometimes, at very high energies, the yield decreases because of the increasing penetration depth and hence increasing energy loss below the surface.



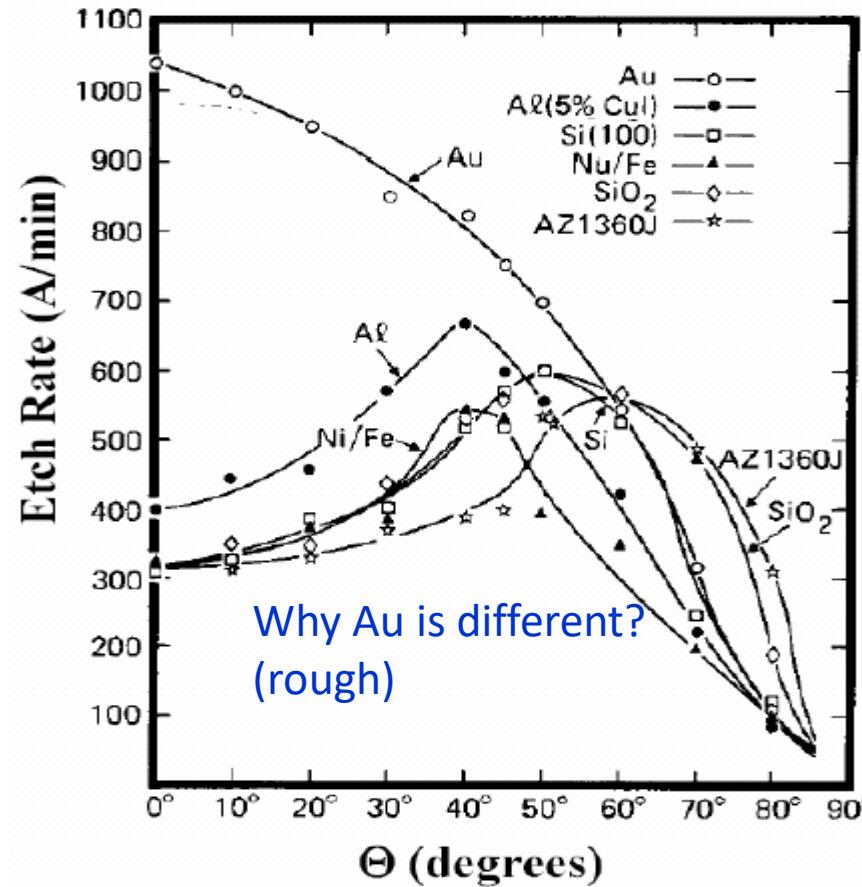
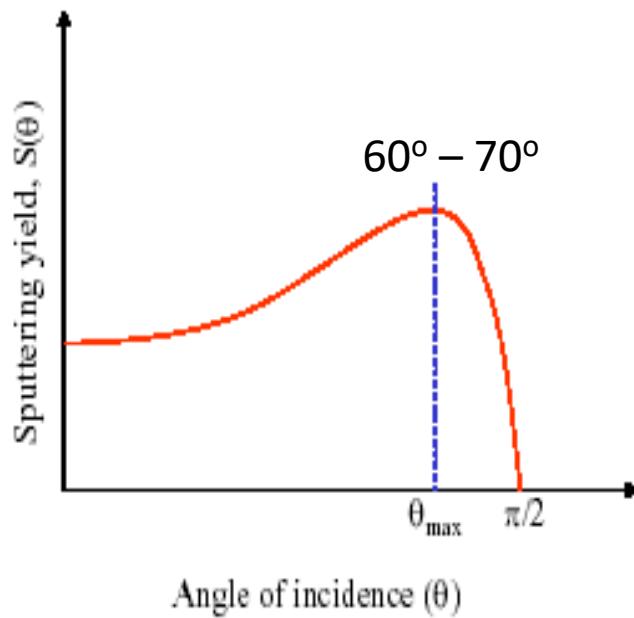
Dependence of sputter yield on ion incident angle



The yield increases as $(\cos\theta)^{-1}$ with increasing obliqueness (θ) of the incident ions.

However, at larger angle of incidence the (lack of) surface penetration effect decrease the yield drastically.

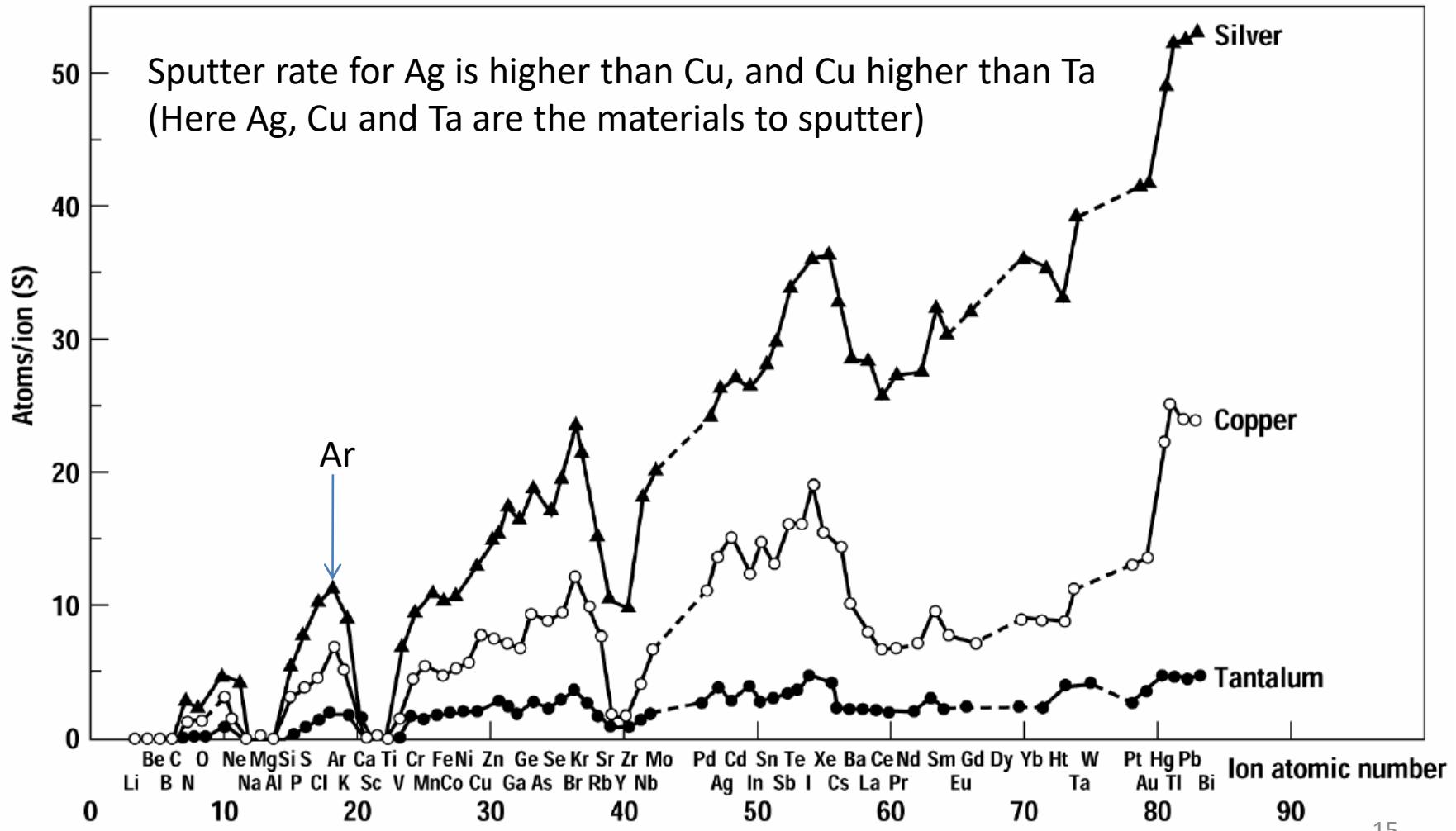
Obviously, at 90 deg (extreme glancing incidence), there should be no sputtering (yield ~ 0).



Dependence of sputter yield on ion mass

Sputter increases with ion mass.

Sputter yield is a maximum for ions with full valence shells: noble gasses such as Ar, Kr, Xe have large yields.



Sputter yield of elements at 500eV

Gas	He	Ne	Ar	Kr	Xe
Element					
Be	0.24	0.42	0.51	0.48	0.35
C	0.07	—	0.12	0.13	0.17
Al	0.16	0.73	1.05	0.96	0.82
Si	0.13	0.48	0.50	0.50	0.42
Ti	0.07	0.43	0.51	0.48	0.43
V	0.06	0.48	0.65	0.62	0.63
Cr	0.17	0.99	1.18	1.39	1.55
Cu	0.24	1.80	2.35	2.35	2.05
Fe	0.15	0.88	1.10	1.07	1.00
Ni	0.16	1.10	1.45	1.30	1.22
Nb	0.03	0.33	0.60	0.55	0.53
Mo	0.03	0.48	0.80	0.87	0.87
Pd	0.13	1.15	2.08	2.22	2.23
Ag	0.20	1.77	3.12	3.27	3.32
Ta	0.01	0.28	0.57	0.87	0.88
W	0.01	0.28	0.57	0.91	1.01
Re	0.01	0.37	0.87	1.25	—
Os	0.01	0.37	0.87	1.27	1.33
Ir	0.01	0.43	1.01	1.35	1.56
Pt	0.03	0.63	1.40	1.82	1.93
Au	0.07	1.08	2.40	3.06	3.01
Au	0.10	1.3	2.5	—	7.7
Pb	1.1	—	2.7	—	—

Step coverage of sputtering

Sputtering targets are generally large and provide a wide range of arrival angles in contrast to a point source. Step coverage is mainly determined by arrival angle distribution.

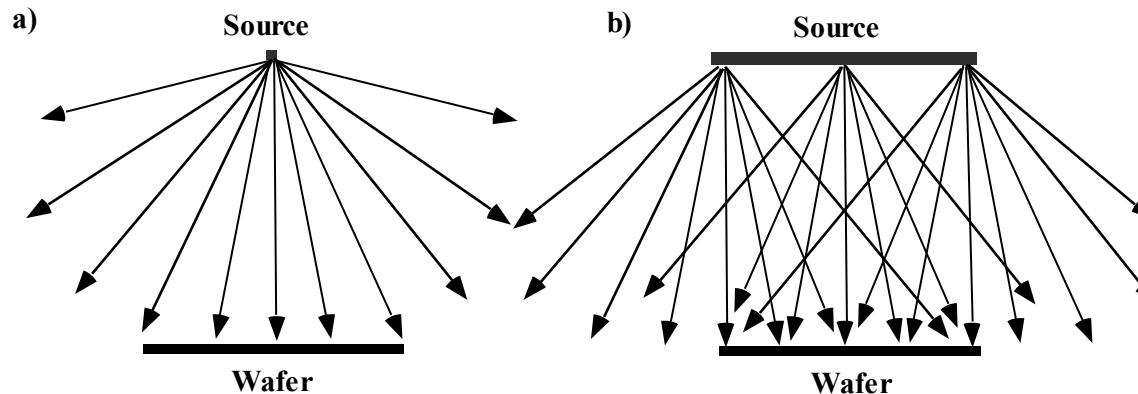


Figure 9-26

- (a) Small source, wide emitted angle distribution, but a narrow arrival angle distribution.
(b) Wider arrival angle distribution.

Isotropic flux arrival

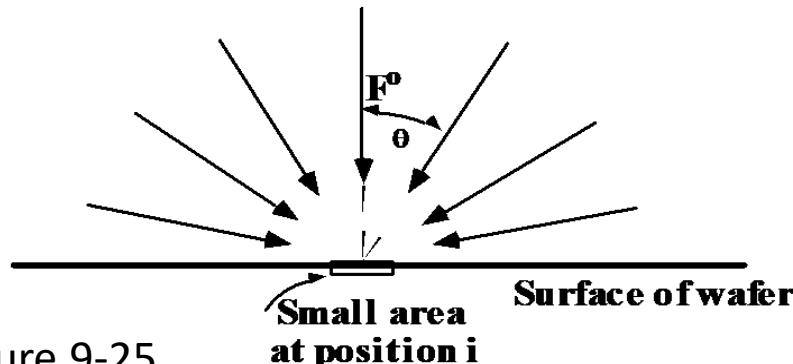
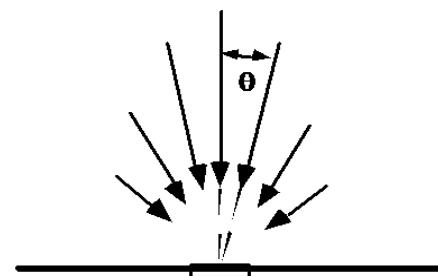


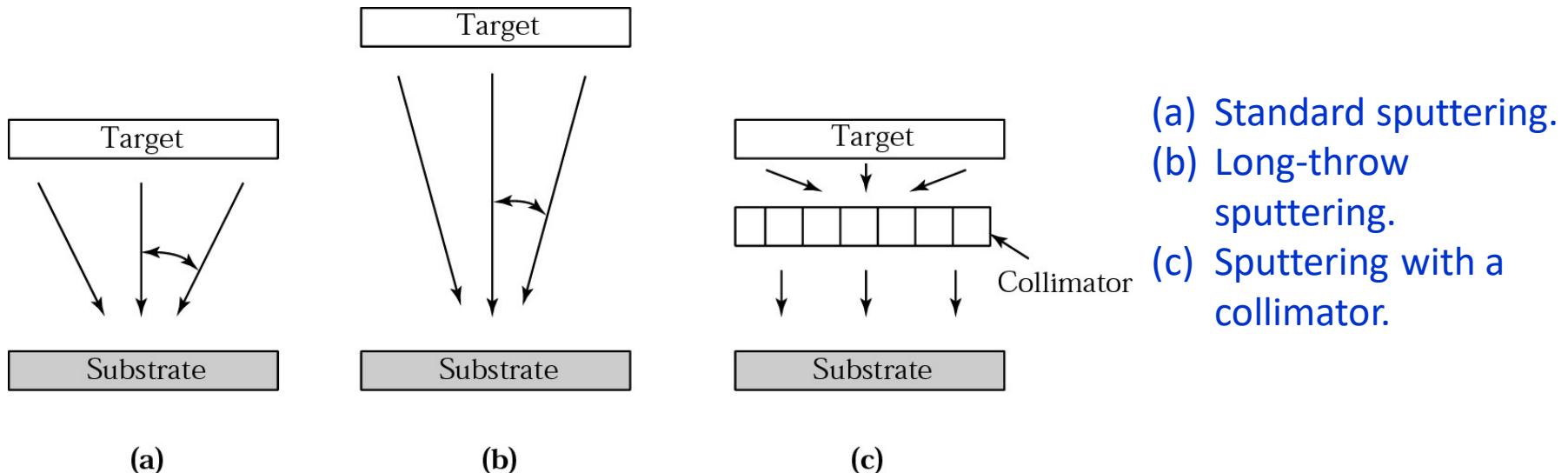
Figure 9-25

Anisotropic flux arrival

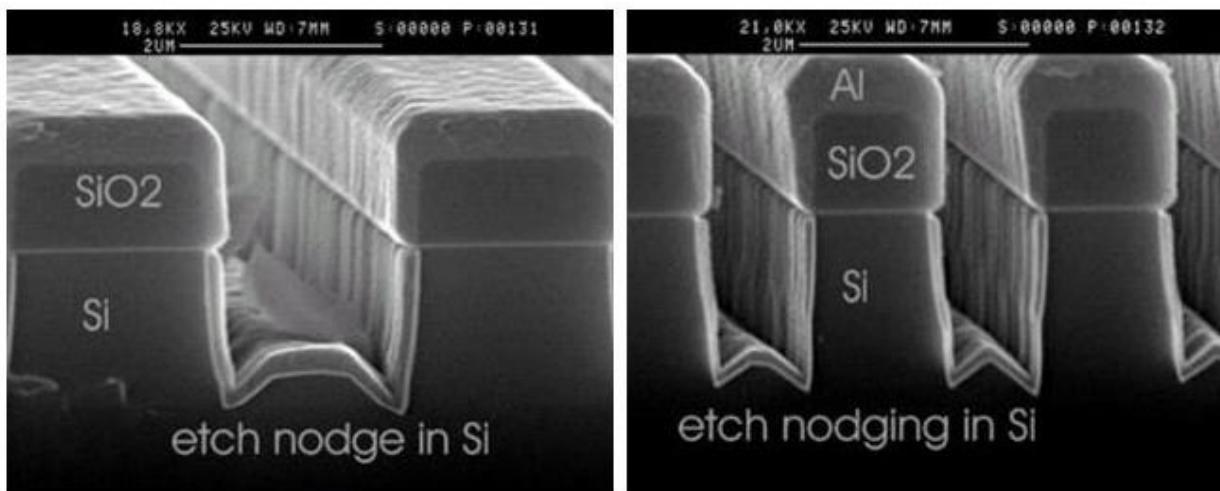


Arrival angle distribution is generally described by $\cos^n\theta$ distribution.
Size of source, system geometry and collisions in gas phase are important in arrival angle distribution.

Arrival angle can be tailored to some degree



However, when the mean free path of the *target* atom (determined by gas pressure, order 10cm for 1mTorr pressure/1cm for 10mTorr) is much shorter than target-substrate separation, many collisions will occur, which broaden the arrival angle distribution.



More deposition
on top surface.

Adatom migration along surface also important

- Atoms ejected from cathode escape with energies of 10 to 50 eV, which is 10-100 times the energy of evaporated atoms.
- This additional energy (together with bombardment by other ions) provides sputtered atoms with additional surface mobility for improved step coverage relative to evaporation. (This additional energy also makes the deposited film “denser” - better film quality than evaporated film).

Besides tilting and rotating substrate, step coverage can be further improved by:

- Substrate heating: improve step coverage due to surface diffusion, but may produce unacceptably large grains.
- Apply bias to wafers: increase bombardment by energetic ions, but it will also sputter the deposited material off the film and thus reduce deposition rate.

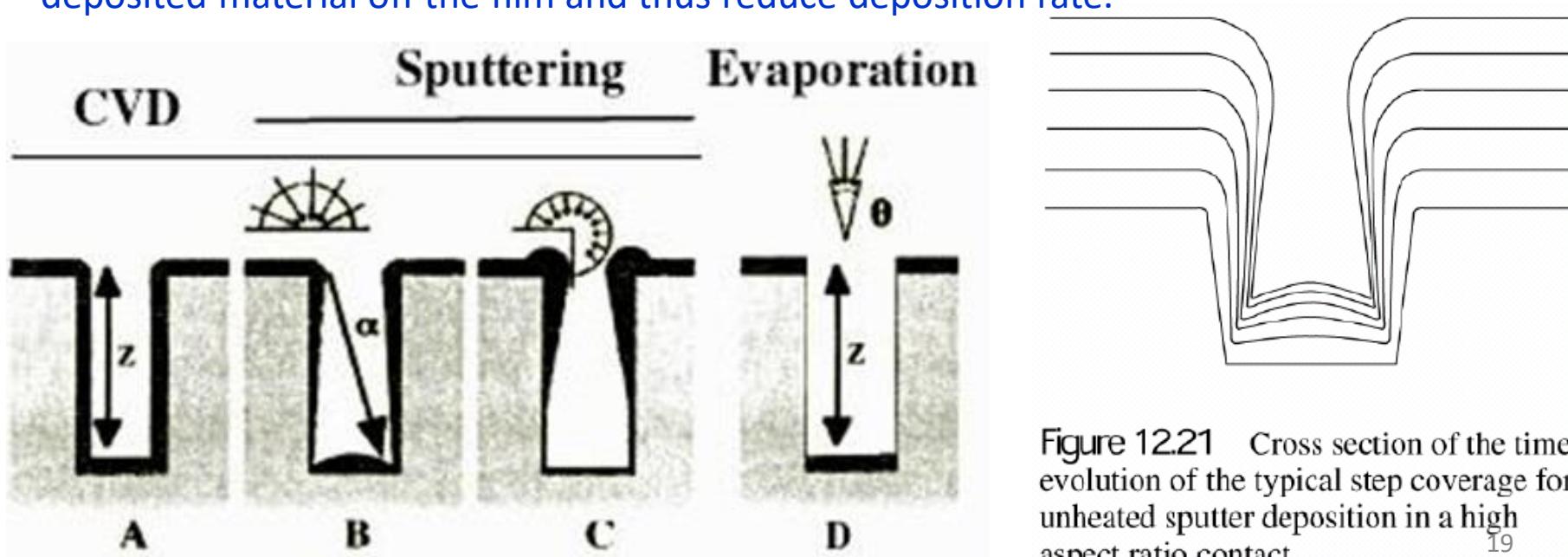


Figure 12.21 Cross section of the time evolution of the typical step coverage for unheated sputter deposition in a high aspect ratio contact. 19

Film morphology: the zone model

Zone model: film morphology as a function of substrate temperature and incident ion energy.

Once reach wafer surface, adatoms (newly added atom) diffuse along surface until they form nuclei.

Nuclei capture more adatoms, forming islands.

If surface mobility is high, islands may merge, forming a smooth continuous film.

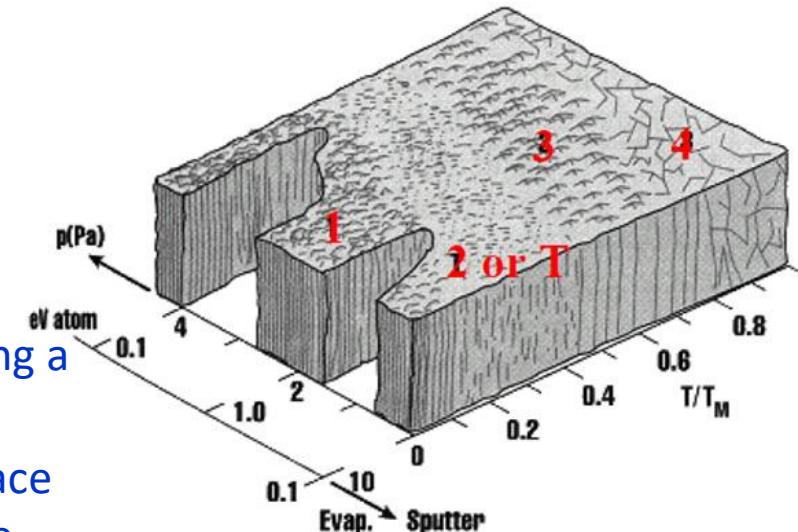
Zone 1: porous and/or amorphous due to poor surface mobility, which is in-turn caused by low temperature and/or low ion energy (due to low RF power/DC bias or higher pressures - less acceleration between collisions).

Metal films in this region can readily oxidize when exposed to air and so may have high resistivity.

Zone 2 (“T-zone”): most desirable, small grain polycrystalline, dense, smooth (high reflectance) due to higher surface mobility (higher temperature and/or ion energy).

Zone 3: further increases in surface mobility result in large columnar grains that have rough surfaces. These rough surfaces lead to poor coverage in later steps.

Zone 4: still further increases in surface mobility result in large non-columnar grains. These grains can pose problems for lithography due to light scatter off of large grains, and tend to be more rigid leading to more failures in electrical lines.



Zone model of film deposition.
 T_m : melting temperature.

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF bias, magnetron, collimated, and ion beam deposition.

Reactive: a mixture of inert + reactive gases used for sputtering

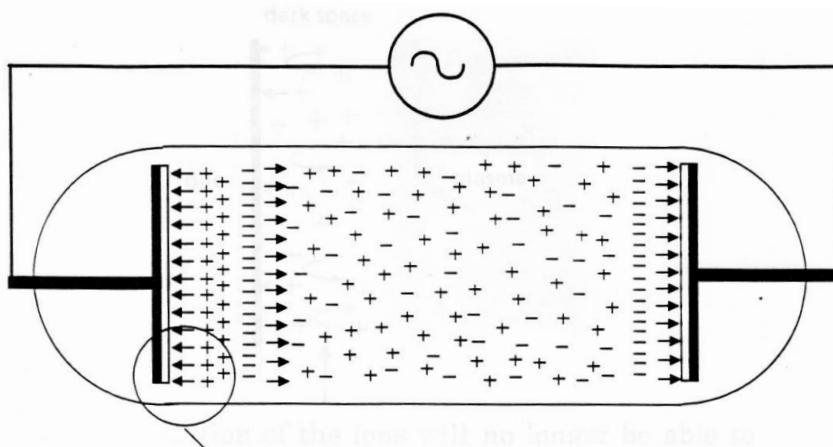
Oxides – Al_2O_3 , SiO_2 , Ta_2O_5 (O_2 mixed with Ar)

Nitrides – TaN , TiN , Si_3N_4 (N_2 , NH_3 , mixed with Ar)

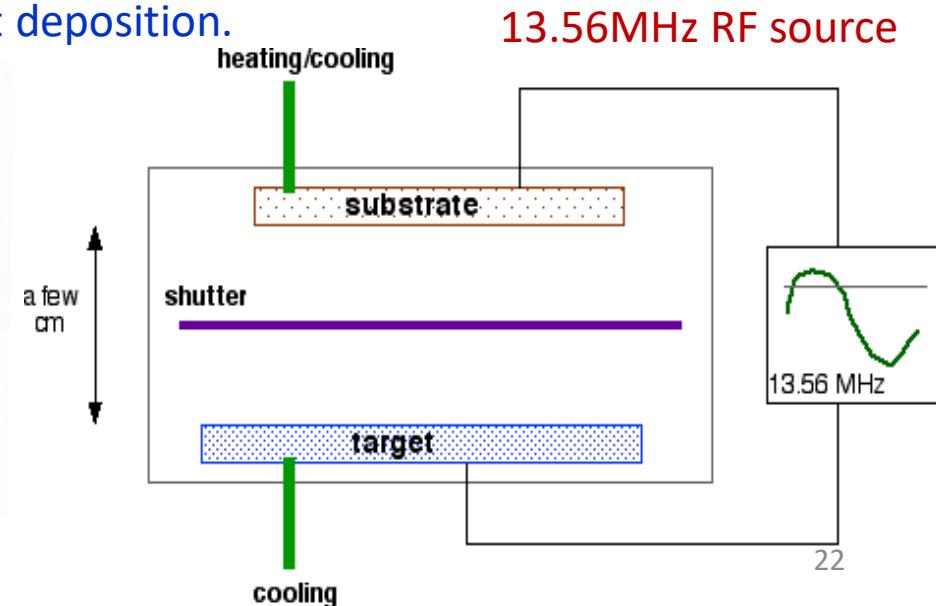
Carbides – TiC , WC , SiC (CH_4 , C_2H_4 , C_3H_8 , mixed with Ar)

RF (radio frequency) sputter deposition

- Good for insulating materials because current cannot flow across an insulating target in DC sputtering systems.
- Impedance for target $Z=1/(i\omega C)$, C is target capacitance, ω is angular frequency. Z is too high for low frequency; and if so, most voltage drop is across the insulating target.
- When frequencies less than $\sim 50\text{kHz}$, both electrons and ions can follow the switching of the anode and cathode, and once electrons / ions are all lost to the electrodes, the plasma will be off and deposition will stop.
- When frequencies well above $\sim 50\text{kHz}$, ions (heavy) can no longer follow the switching, and the target becomes effectively conductive (Z is small enough).
- As now electrons gain energy directly from RF power, and oscillating electrons are more efficient to ionize the gas, RF sputter is capable of running in lower pressure (1-15 mTorr), so fewer gas collisions and more line of sight deposition.



Switch polarities before the target surface saturates with ions.



RF plasma

- For symmetric target-substrate configuration, sputtering of both surfaces will occur, though in the opposite half cycles.
- When the electrode areas are not equal, the field must be higher at the smaller electrode (higher current density), to maintain overall current continuity.
- It was found that voltage drop across the dark sheath of the two electrodes satisfy the relation: (A is the area of the electrode)

$$\frac{V_1}{V_2} = \left(\frac{A_2}{A_1} \right)^m \quad (m = 1-2 \text{ experimentally})$$

- Thus by making the target electrode much smaller, sputtering occurs "only" on the target.
- Wafer electrode can also be connected to chamber walls, further increasing V_2/V_1 .

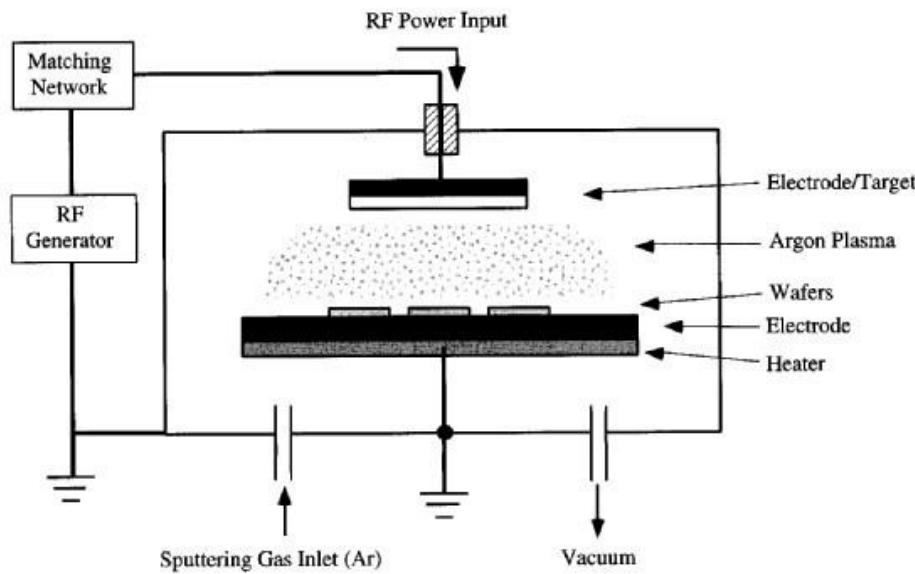


Figure 9-28 Schematic diagram of RF-powered sputter deposition system.

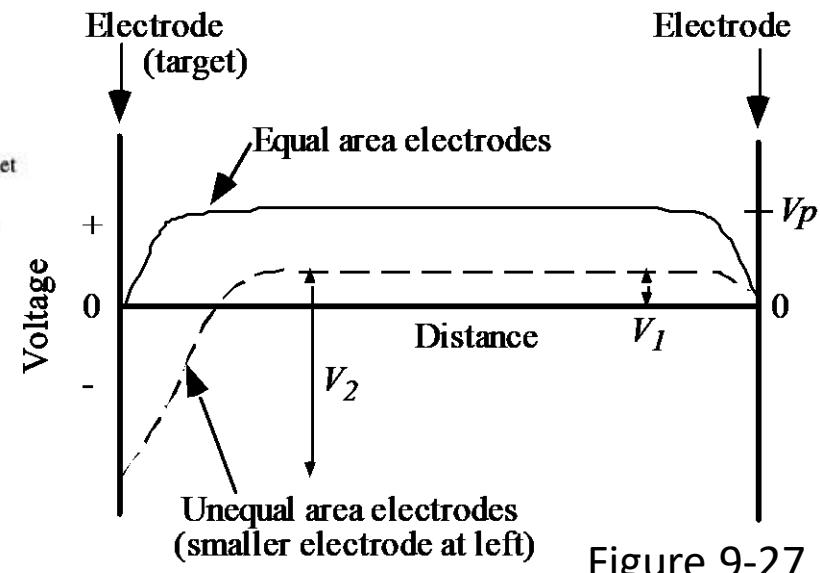
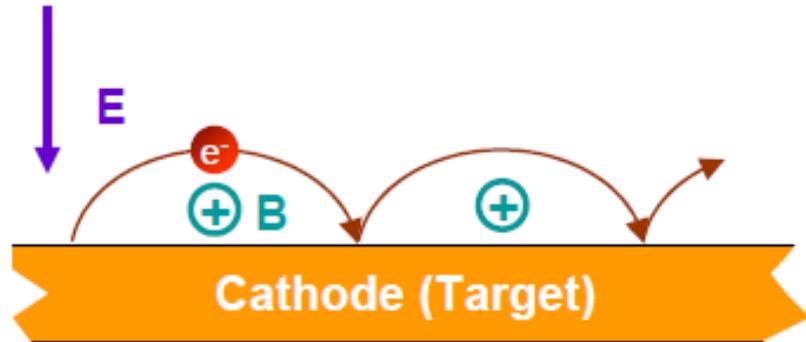


Figure 9-27
23

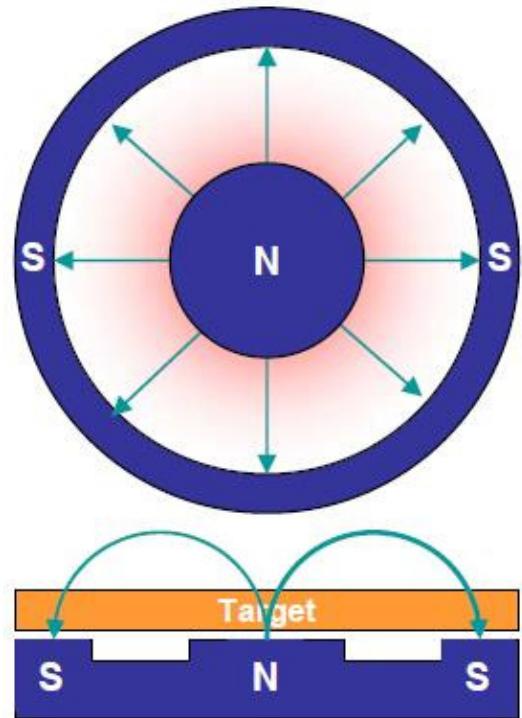
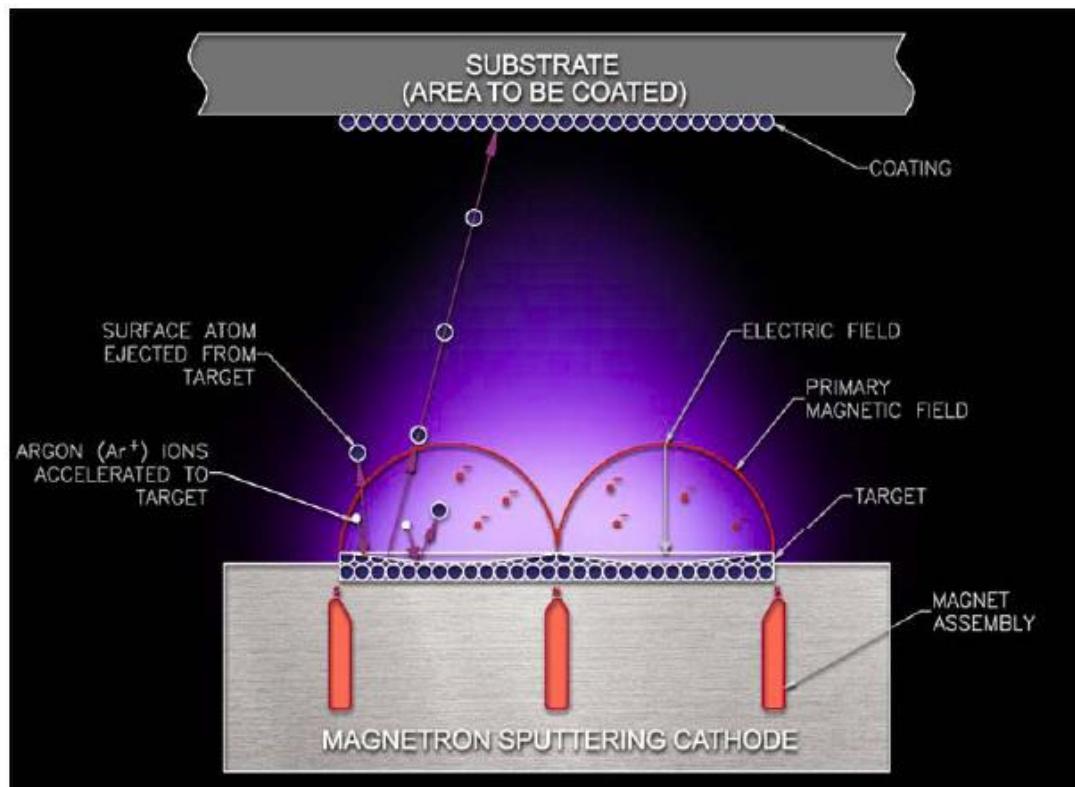
Magnetron sputtering

- To increase deposition rates, magnets are used to increase the percentage of electrons that take part in ionization events, increasing the ionization efficiency.
- A magnetic field is applied at right angles to the electric field by placing large magnets behind the target.
- This traps electrons near the target surface and causes them to move in a spiral motion until they collide with an Ar atom.
- The ionization and sputtering efficiencies are increased significantly - **deposition rates increase by 10-100×**.
- Unintentional wafer heating is reduced since the dense plasma is confined near the target and ion loss to the wafers is less.
- A lower Ar pressure (to 0.5mTorr, can still sustain plasma) can be utilized since ionization efficiency is larger which can improve film quality as less argon will be incorporated into the film.
- Magnetron sputtering can be done in DC or RF, though more often it is done in DC.

Magnetron sputtering



Orbital motion of electrons increases probability that they will collide with neutral species and create ions.



Magnetron sputtering for high density of plasma near target.

Impact of magnetic field on ions

Hopping radius r :

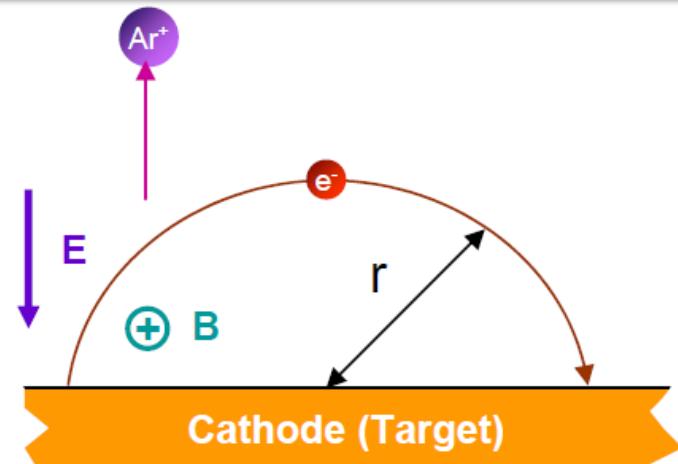
$$r \sim \frac{1}{B} \sqrt{\frac{2m}{e} V_d}$$

V_d : voltage drop across dark space/sheath ($\sim 100V$)

B: magnetic field ($\sim 100G$)

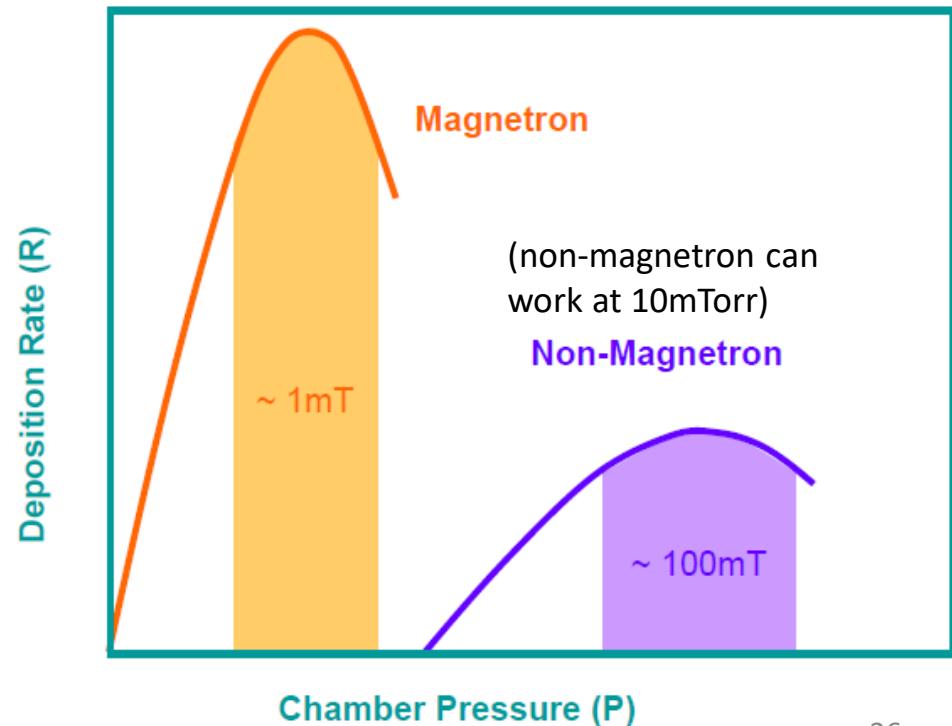
For electron: $r \sim 0.3\text{cm}$

For Ar^+ ion: $r \sim 81\text{cm}$



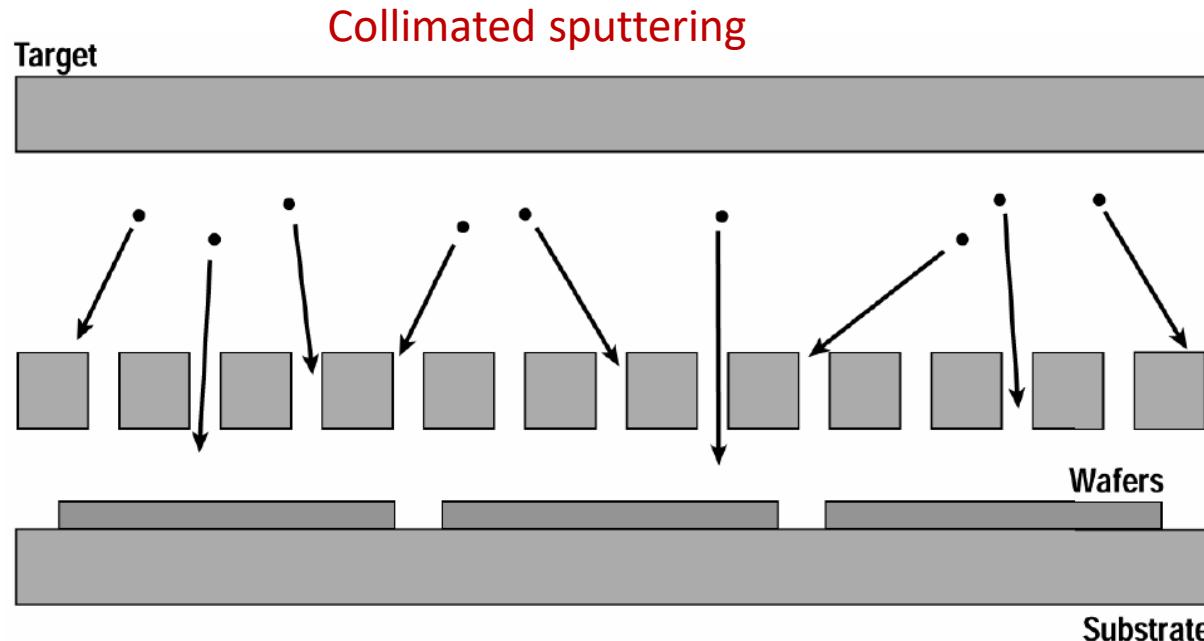
As a result:

- Current density (proportional to ionization rate) increases by many times.
- Required discharge pressure drops.
- Deposition rate increases greatly.



Collimated sputtering

The goal is to fill high aspect ratio holes by more directional sputtering with narrow arrival angle distribution.



- Insert a plate with high-aspect-ratio holes.
- Sputter at low pressure, mean path is long enough that few collisions occur between collimator and wafer.
- Species with velocities nearly perpendicular to wafer surface pass through the holes.
- Reduce deposition rate considerably (most sputtered atoms cannot reach the substrate).

Ionized sputter deposition

DC target bias

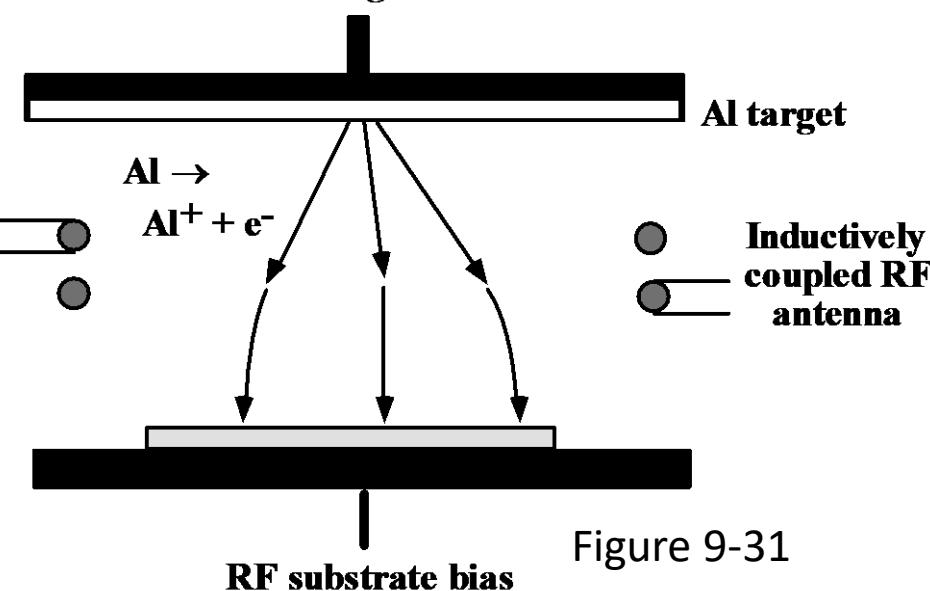


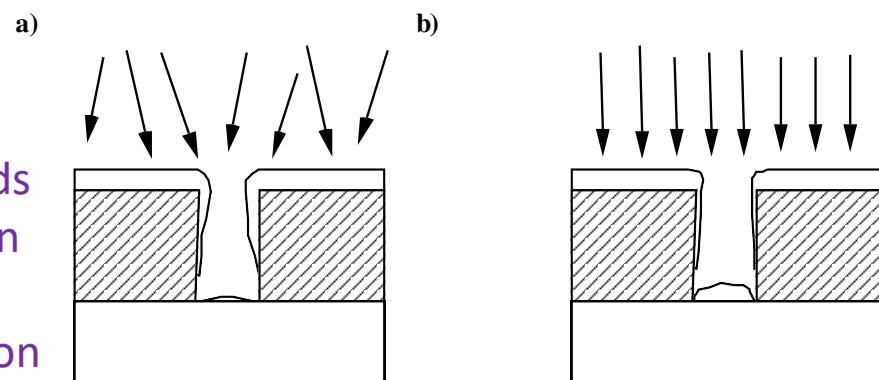
Figure 9-31

Figure 9-30 (at the right)

- Regular sputter deposition.
- Collimated sputter deposition, by using long throw configuration, a collimator, or ionized sputter deposition.

Overall, too narrow arrival angle distribution leads to non-conformal deposition (too little deposition onto hole sidewall), whereas too broad leads to fast closing up of the opening thus little deposition onto hole bottom and sidewall. Both are not good.

- The depositing atoms themselves are ionized.
- An RF coil around the plasma induces collisions in the plasma, creating the ions (50-85% ionized).
- Most sputtered atoms can reach the substrate, thus it is a better solution than a collimator.
- This, again, provides a narrow distribution of arrival angles, which may be useful when filling or coating the bottom of deep contact hole.



Comparison of evaporation and sputtering

Evaporation	Sputtering
Low energy atoms (~ 0.1 eV)	High energy atoms / ions (1 – 10 eV) <ul style="list-style-type: none">• denser film• smaller grain size• better adhesion
High Vacuum <ul style="list-style-type: none">• directional, good for lift-off• lower impurity	Low Vacuum <ul style="list-style-type: none">• poor directionality, better step coverage• gas atom implanted in the film
Point Source <ul style="list-style-type: none">• poor uniformity	Parallel Plate Source <ul style="list-style-type: none">• better uniformity
Component Evaporate at Different Rate <ul style="list-style-type: none">• poor stoichiometry	All Component Sputtered with Similar Rate <ul style="list-style-type: none">• maintain stoichiometry

Comparison of typical thin film deposition technology

Process	Material	Uniformity	Impurity	Grain Size	Film Density	Deposition Rate	Substrate Temperature	Directional	Cost
Thermal Evaporation	Metal or low melting-point materials	Poor	High	10 ~ 100 nm	Poor	1 ~ 20 A/s	50 ~ 100 °C	Yes	Very low
E-beam Evaporation	Both metal and dielectrics	Poor	Low	10 ~ 100 nm	Poor	10 ~ 100 A/s	50 ~ 100 °C	Yes	High
Sputtering	Both metal and dielectrics	Very good	Low	~ 10 nm	Good	Metal: ~ 100 A/s Dielectric: ~ 1-10 A/s	~ 200 °C	Some degree	High
PECVD	Mainly Dielectrics	Good	Very low	10 ~ 100 nm	Good	10 - 100 A/s	200 ~ 300 °C	Some degree	Very High
LPCVD	Mainly Dielectrics	Very Good	Very low	1 ~ 10 nm	Excellent	10 - 100 A/s	600 ~ 1200 °C	Isotropic	Very High

Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

Common deposition methods for thin films in IC fabrication

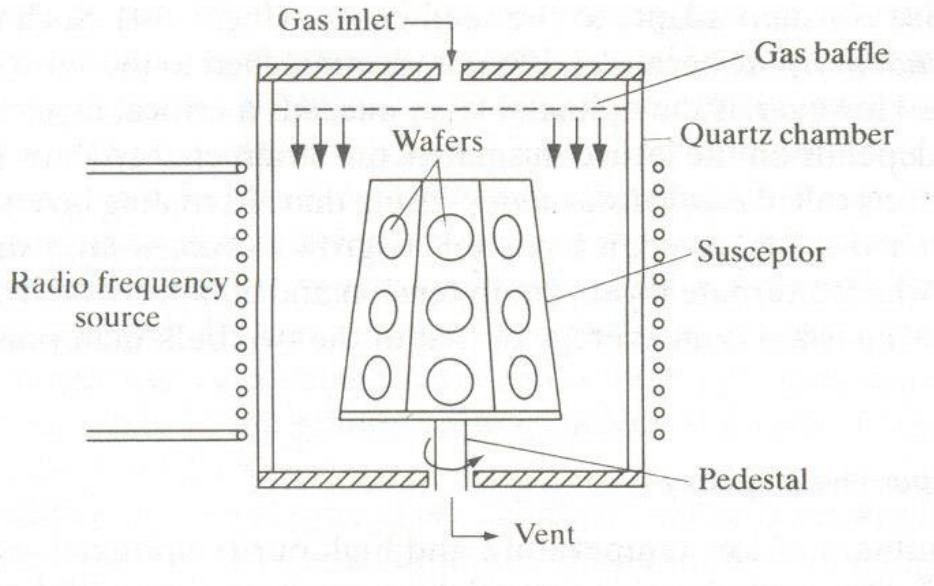
Table 9–1 Common deposition methods for thin films in integrated circuit fabrication

Thin Film	Equipment	Typical Reactions	Comments
Epitaxial silicon	APCVD, LPCVD	$\text{SiH}_4 \rightarrow \text{Si} + 2\text{H}_2$	1000–1250°C.
		$\text{SiCl}_4 + 2\text{H}_2 \rightarrow \text{Si} + 4\text{HCl}$ Also $\text{SiHCl}_3, \text{SiH}_2\text{Cl}_2$	Reduce pressure for lower-temperature deposition.
Polysilicon	LPCVD	Same as epitaxial Si	575–650°C. Grain structure depends on deposition conditions and doping.
Si_3N_4	LPCVD, PECVD	$3\text{SiH}_4 + \text{NH}_3 \rightarrow \text{Si}_3\text{N}_4 + 12\text{H}_2$	650–800°C for oxidation mask. 200–400°C (PECVD) for passivation.
SiO_2	LPCVD, PECVD, HDPCVD, APCVD	$\text{SiH}_4 + \text{O}_2 \rightarrow \text{SiO}_2 + 2\text{H}_2$ $\text{Si}(\text{OC}_2\text{H}_5)_4 (+ \text{O}_3) \rightarrow \text{SiO}_2 + \text{byproducts}$	200–800°C. 200–500°C (LTO)—may require high T anneal. 25–400°C (TEOS-ozone, PECVD, HDPCVD).
Al	Magnetron sputter deposition		25–300°C (standard deposition). 440–550°C (hot Al for in situ reflow). CVD difficult for alloys (Al-Cu-Si).
Ti and Ti-W	Magnetron sputter deposition (standard, ionized, or collimated)		CVD difficult. Nitrogen can be added to Ti-W to stuff grain boundaries.
W	LPCVD	$2\text{WF}_6 + 3\text{SiH}_4 \rightarrow 2\text{W} + 3\text{SiF}_4 + 6\text{H}_2$ $\text{WF}_6 + 3\text{H}_2 \rightarrow \text{W} + 6\text{HF}$	250–500°C. Blanket deposition with two-step process using both reactions is common.
TiSi_2	Sputter and surface reaction Cospattering or CVD	$\text{Ti}(\text{sputtered}) + \text{Si}(\text{exposed}) \rightarrow \text{TiSi}_2$	Sputter/reaction gives self-aligned silicide.
TiN	Reactive sputter deposition	$\text{Ti} + \text{N}_2 (\text{in plasma}) \rightarrow \text{TiN}$	Organometallic source possible for MOCVD deposition.
	CVD	$6\text{TiCl}_4 + 8\text{NH}_3 \rightarrow 6\text{TiN} + 24\text{HCl} + \text{N}_2$	TiN can also be formed in TiSi_2 process.
Cu	Electroplating, electroless, sputtering, CVD	$\text{Cu}^{2+} + 2e^- \rightarrow \text{Cu}$	Electroplating is most common method today.

Epitaxial silicon deposition

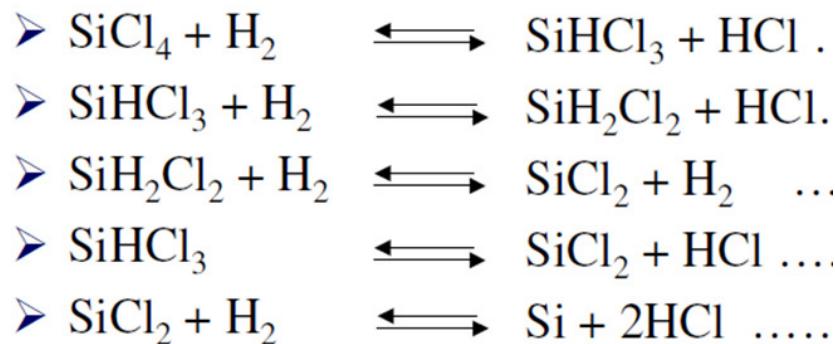
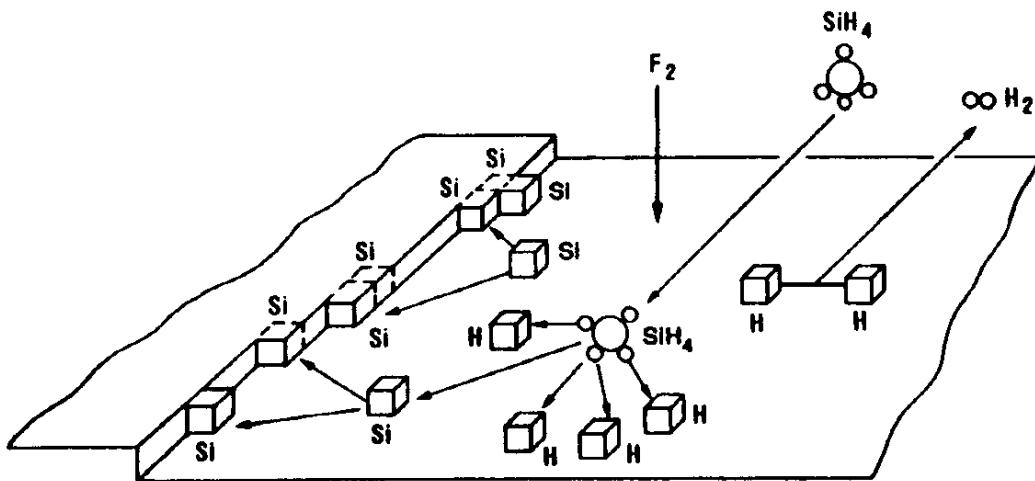
Advantages of epitaxial wafers over bulk wafers

- Offers means of controlling the doping profile (e.g. lightly doped on heavily doped possible)
- Epitaxial layers are generally oxygen and carbon free



Gases used in silicon epitaxy

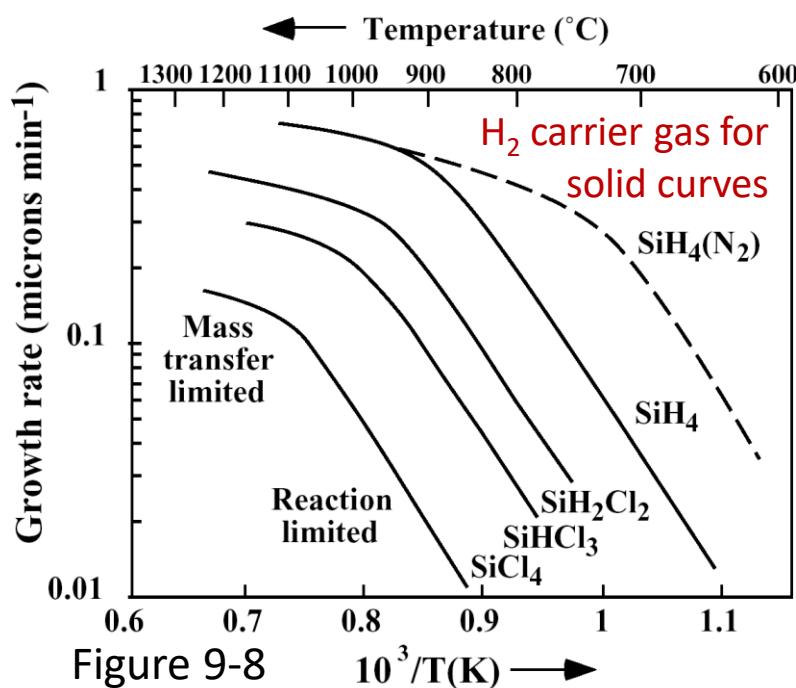
- a) Silane (SiH₄) Pyrolysis : SiH₄ (H₂)
- b) Dichlorosilane (DCS) SiH₂Cl₂
- c) Trichlorosilane (TCS) SiHCl₃
- d) Silicon tetrachloride SiCl₄
- e) Disilane Si₂H₆
- f) Dopant gases
 - Diborane (B₂H₆)
 - Phosphine (PH₃)
 - Arsine (AsH₃)



Si APCVD epitaxy growth process

- Hydrogen gas purges of air from the reactor.
 - Reactor is heated to a temperature.
 - After thermal equilibrium, an HCl etch takes place at 1150°C and 1200°C for 3 minutes.
 - Temperature is reduced to growth temperature.
 - Silicon source and dopant flows are turned on.
 - After growth, temperature is reduced by shutting off power.
 - Hydrogen flow replaced by nitrogen flow.
 - Depending on wafer diameter and reactor type, 10 to 50 wafers per batch can be formed.
 - Process cycle times are about one hour.
-
- Epitaxy film need high temperature ($>1000^{\circ}\text{C}$) because at high temperature the (amorphous) native oxide will become unstable and desorb from the surface, exposing the single crystalline silicon lattice for epitaxy.

Polycrystalline silicon deposition



- Application: gate of MOSFET.
- Usually deposited in a LPCVD chamber at 25-150Pa, 600-650°C, 10-20nm/min.
- SiH_4 is preferred because of its lower deposition temperature.

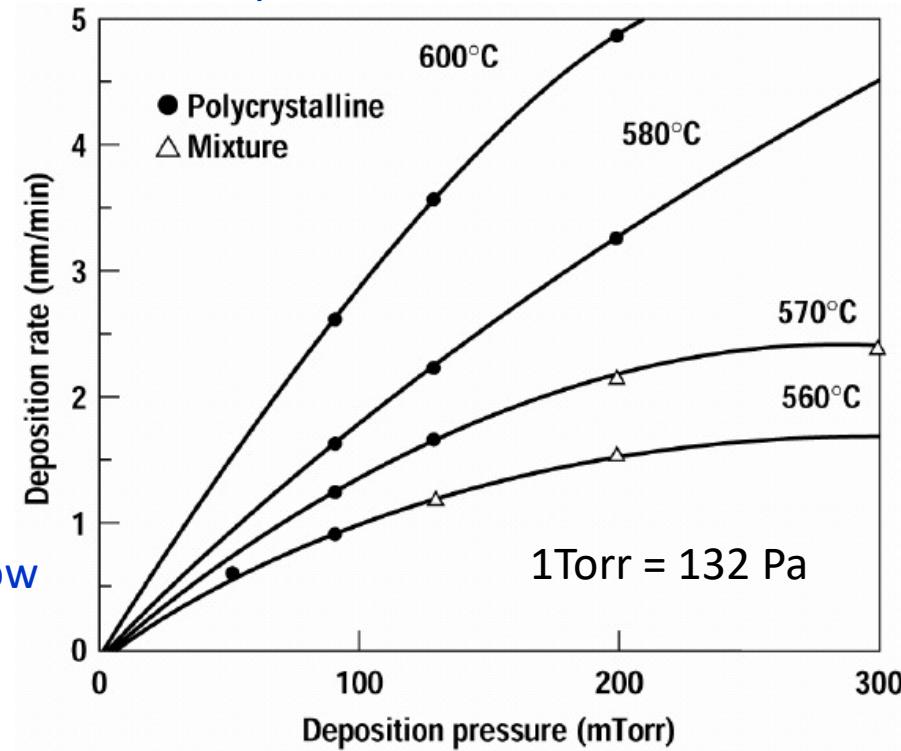


Figure 13.15 Deposition rate of polysilicon as functions of temperature and silane flow rate (*after Voutsas and Hatalis, reprinted by permission of the publisher, The Electrochemical Society*).

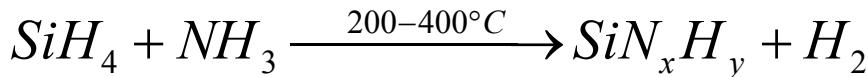
- Usually amorphous when deposition at <575°C; but may be polycrystalline if deposition rate is low enough.
- Columnar grain structure/texture, and the grain will grow when annealed.
- When annealed, amorphous Si will become polycrystalline Si with even large grain size than poly-Si under same annealing.

Silicon nitride deposition

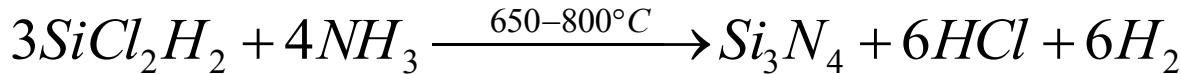
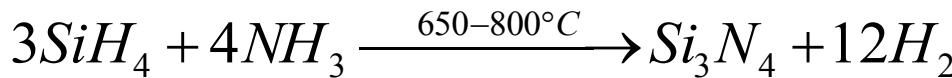
- Application:

- Masks to prevent oxidation for LOCOS process
- Final passivation barrier for moisture and sodium contamination
- Etch stop for Cu damascene process

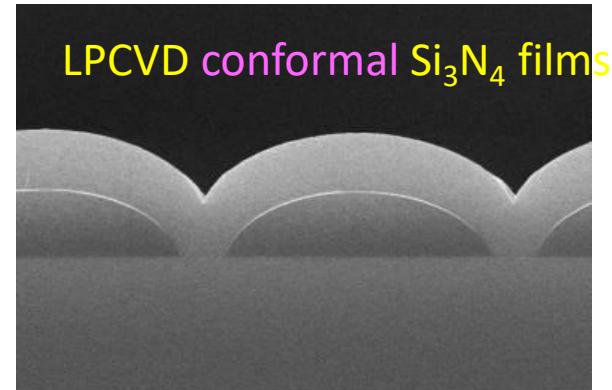
- PECVD



- LPCVD



Deposition Type	LPCVD	PECVD
Typical Temp.	700 to 800°C	< 250 to 350°C
Composition	Si ₃ N ₄ (H)	SiN _x H _y
Si/N Ratio	0.75	0.8 to 1.2
% H	4 to 8	20 to 25
Refractive Index	2.01	1.8 to 2.5
Density (g/cm ³)	2.9 to 3.1	2.4 to 2.8
Resistivity (Ω·cm)	10 ¹⁶	10 ⁶ to 10 ¹⁵
Dielectric Strength (10 ⁶ V/cm or 10 ² V/μm)	10	5
Energy gap (eV)	5	4 to 5
Stress (MPa)	1,000 tens (can be = zero for Si rich films)	200 comp to 500 tens



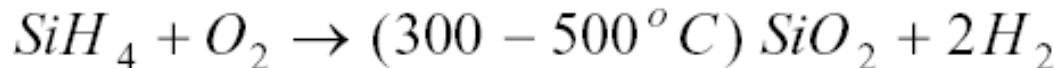
LPCVD film quality is much better than PECVD in almost every aspect.

Silicon dioxide deposition

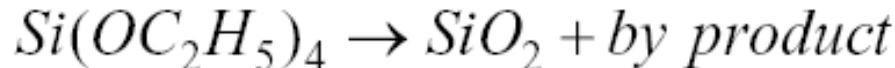
Sputtered oxide has poorer step coverage than CVD.

APCVD has been used for many years, but today LPCVD and PECVD are more popular.

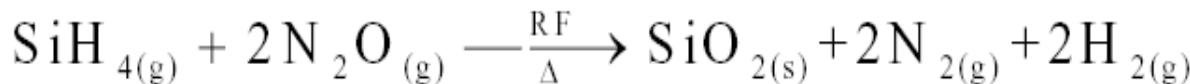
- **Silane based LPCVD**



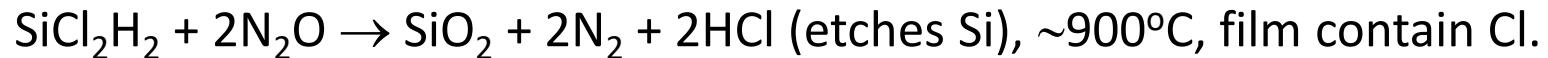
- TEOS (tetra-ethoxy-silane). LPCVD 650-800°C, PECVD 350°C.
Lower sticking coefficient, thus more conformal film.



- **Silane based PECVD**



- **Others**



TEOS + Ozone (O_3). Ozone is more reactive and lowers deposition temperature to $\sim 400^\circ C$.

Comparison of varied silicon dioxide

Property	PECVD SiH_4+O_2	LPCVD SiH_4+O_2	LPCVD TEOS	LPCVD $\text{SiCl}_2\text{H}_2+\text{N}_2\text{O}$	Thermal oxidation
Deposition temp	200°C	450°C	700°C	900°C	1000°C
Composition	$\text{SiO}_2(\text{H})$	$\text{SiO}_2(\text{H})$	$\text{SiO}_2(\text{C}...)$	$\text{SiO}_2(\text{Cl})$	SiO_2
Thermal stability	Loses H	Densifies	Stable	Loses Cl	stable
Density (g/cm ³)	2.3	2.1	2.2	2.2	2.2
Stress (MPa)	3C-3T	3T	1C	3C	3C
Dielectric Strength (10^6 V/cm)	5	8	10	10	11
Etch Rate (Å/min) (100 H_2O :1 HF)	400	60	30	30	25
Step coverage	Non-conformal	Non-conformal	Conformal	Conformal	Conformal

Lower HF etch rate means better film quality (denser film).

For stress, C=compressive, T=tensile

Deposition of metals

Al	Magnetron sputter deposition		25-300°C (standard deposition) 440-550°C (hot Al for in-situ reflow) CVD difficult for alloys (Al-Cu-Si)
Ti and Ti-W	Magnetron sputter deposition (standard, ionized or collimated)		CVD difficult Nitrogen can be added to Ti-W to stuff grain boundaries.
W	LPCVD	$2\text{WF}_6 + 3\text{SiH}_4 \rightarrow 2\text{W} + 3\text{SiF}_4 + 6\text{H}_2$ $\text{WF}_6 + 3\text{H}_2 \rightarrow \text{W} + 6\text{HF}$	250-500°C Blanket deposition with two step process using both reactions is common.
TiSi ₂	Sputter and surface reaction Co-sputtering or CVD	Ti(sputtered) + Si(exposed) → TiSi ₂	Sputter/reaction give self-aligned silicide Two step anneal process required (600/800°C)
TiN	Reactive sputter deposition	$\text{Ti} + \text{N}_2(\text{in plasma}) \rightarrow \text{TiN}$	Organometallic source possible for MOCVD deposition

MOCVD: metal-organic-CVD

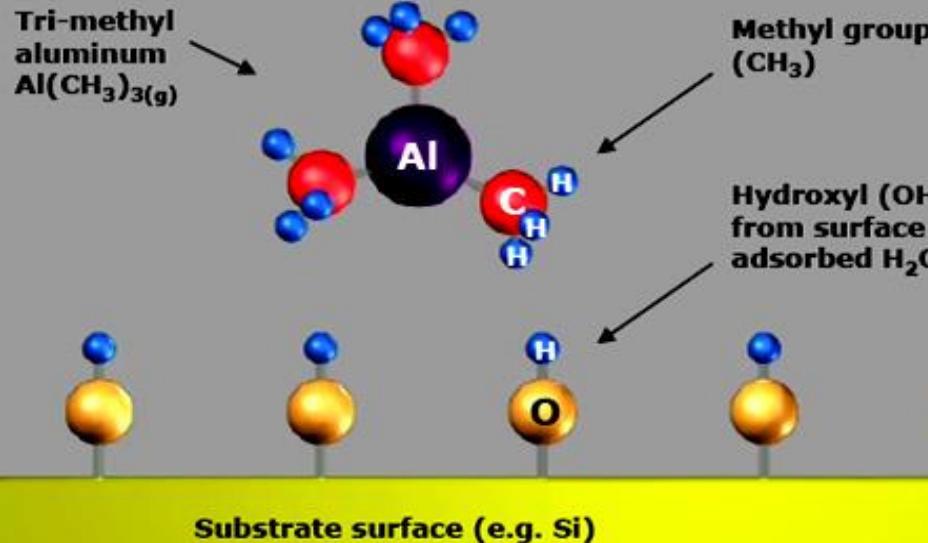
Chapter 9 Thin film deposition

1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

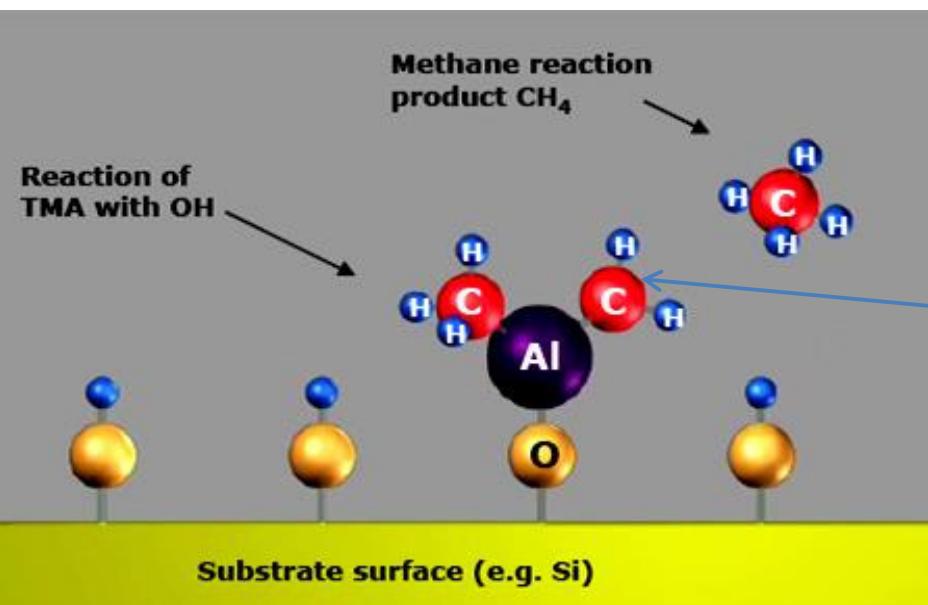
Atomic layer deposition (ALD, break CVD into two steps)

- Similar in chemistry to CVD, except that the ALD reaction breaks the CVD reaction into two half-reactions, keeping the precursor materials separate during the reaction.
- The precursor gas is introduced into the process chamber and produces a monolayer of gas on the wafer surface. A second precursor gas is then introduced into the chamber reacting with the first precursor to produce a *monolayer* of film on the wafer surface.
- Film growth is self-limited (monolayer adsorption/reaction each half-cycle), hence atomic layer thickness control of film growth can be obtained.
- That is, one layer per cycle; thus the resulting film thickness may be precisely controlled by the number of deposition cycles.
- Introduced in 1974 by Dr. Tuomo Suntola and co-workers in Finland to improve the quality of ZnS films used in electroluminescent displays.
- Recently, it turned out that ALD also produces outstanding dielectric layers and attracts semiconductor industries for making High-K dielectric materials.

Example: ALD cycle for Al_2O_3 deposition



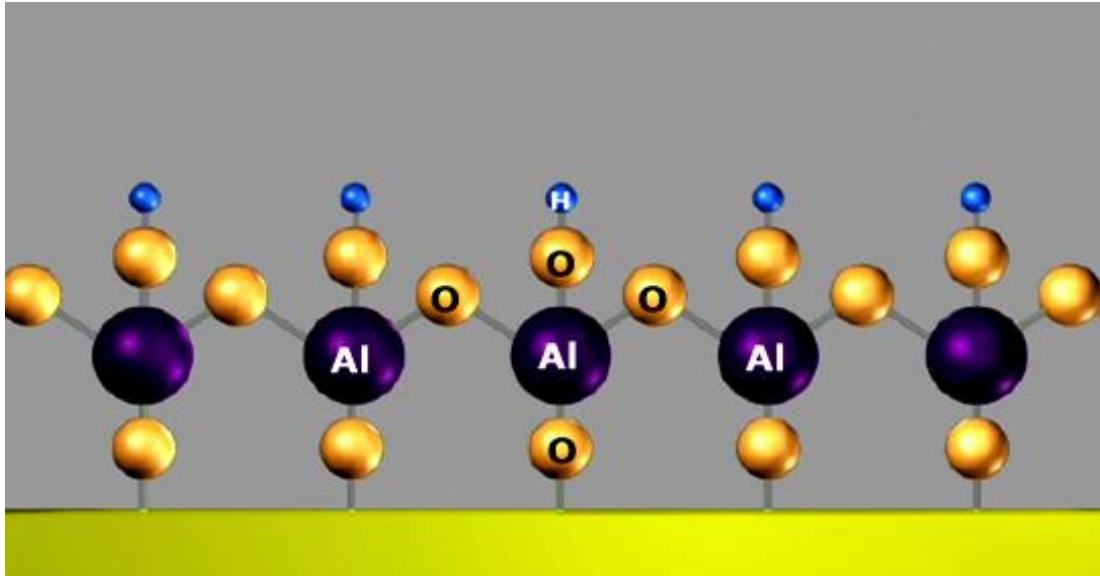
1. Introduce TMA
(tri-methyl aluminum)
In air, H_2O vapor absorb on Si to form $\text{Si}-\text{O}-\text{H}$.



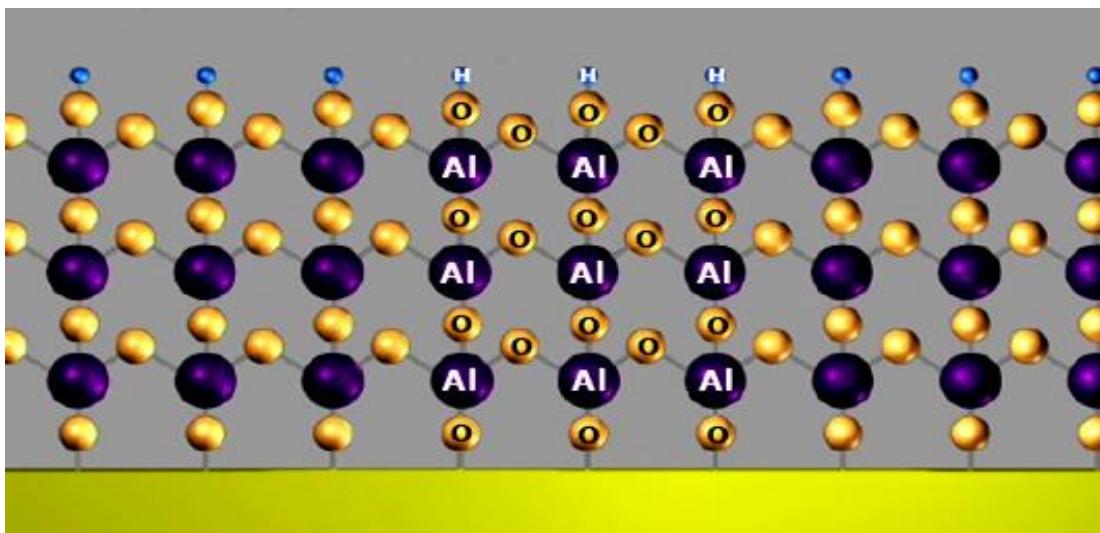
2. TMA reacts with hydroxyl groups to produce methane.

Should be one more "H" here

ALD cycle for Al₂O₃ deposition



3. Introduce H₂O. Reaction product methane is pumped away, leaving an OH⁻ passivation layer on surface.

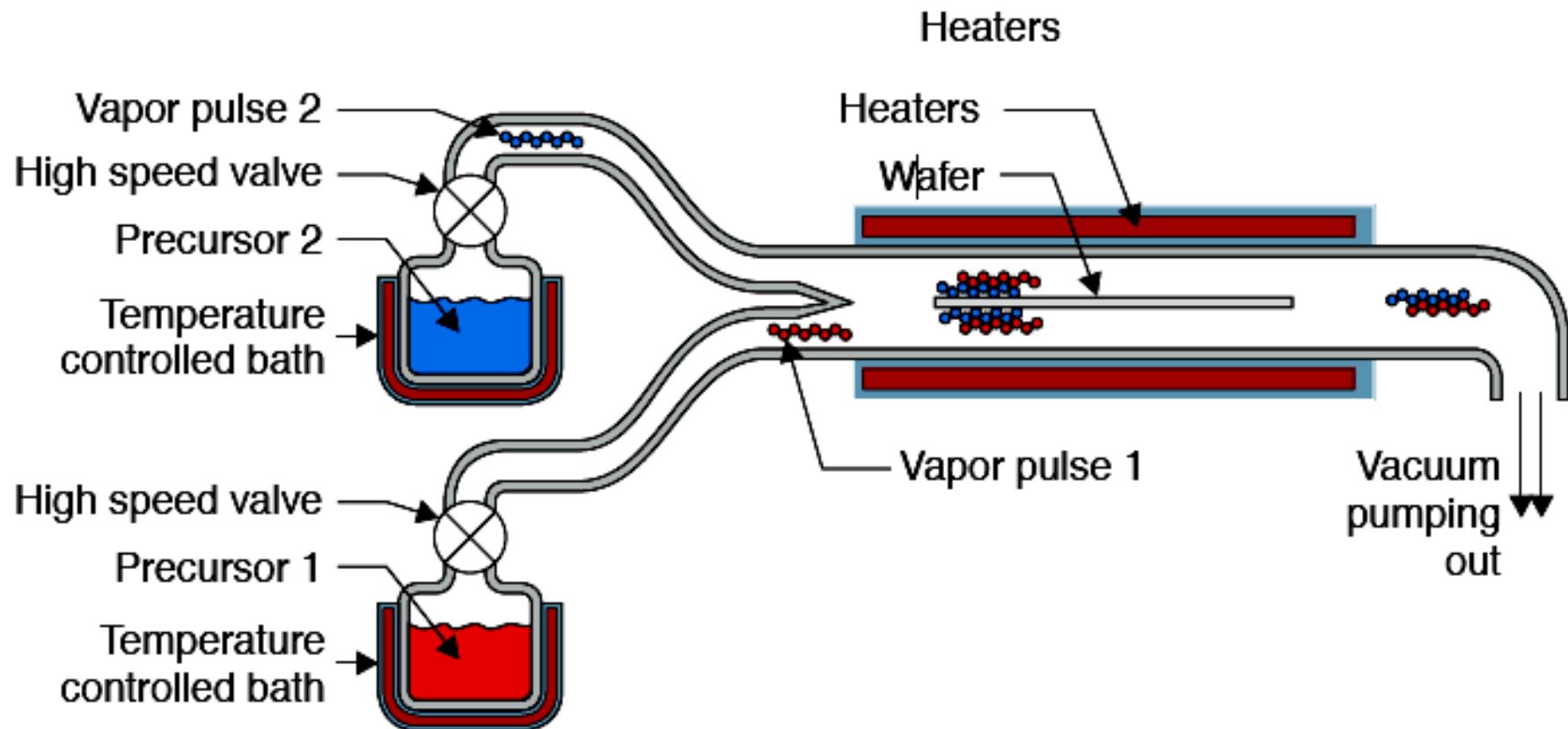


4. After three cycles.
One TMA and one H₂O vapor pulse form one cycle. Here ~1Å/cycle, each cycle including gas injection and pumping takes few seconds.

Two steps each cycle



Closed system chambers (most common) for ALD



Advantages and disadvantages

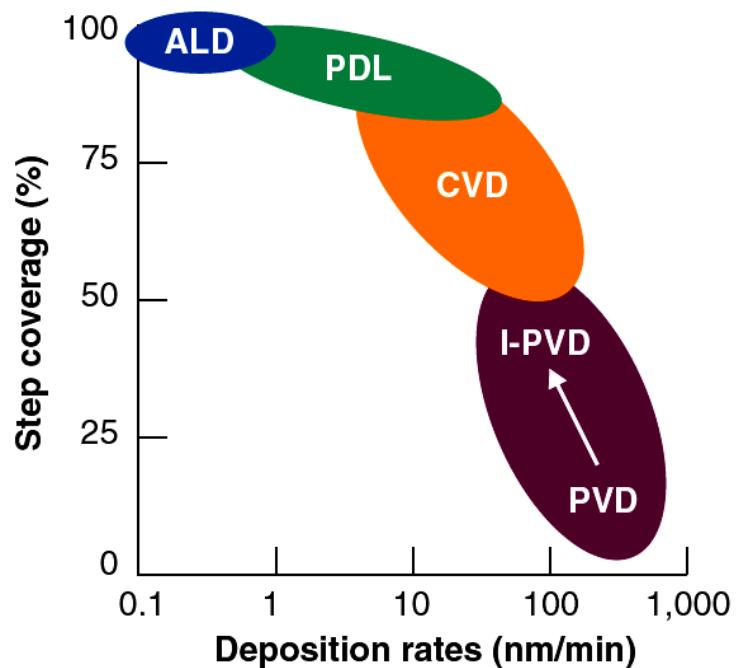
Advantages

- Stoichiometric films with large area uniformity and 3D conformality.
- Precise thickness control.
- Low temperature deposition possible.
- Gentle deposition process for sensitive substrates.

Disadvantages

- Deposition rate slower than CVD.
- Number of different materials that can be deposited is fair compared to MBE.

ALD: slowest, best step coverage
PDL: is PLD?



Chapter 9 Thin film deposition

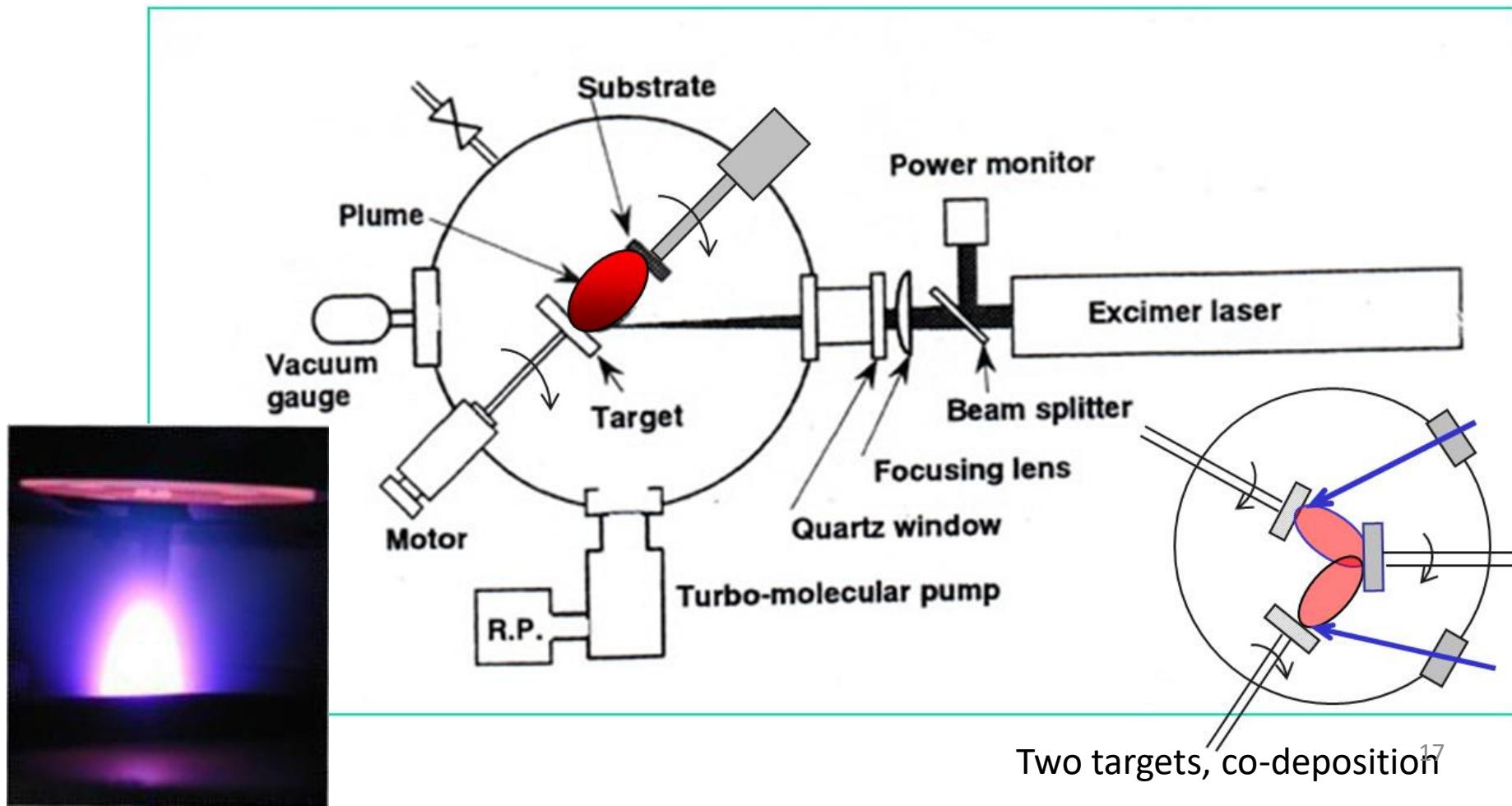
1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

Pulsed laser deposition (PLD)

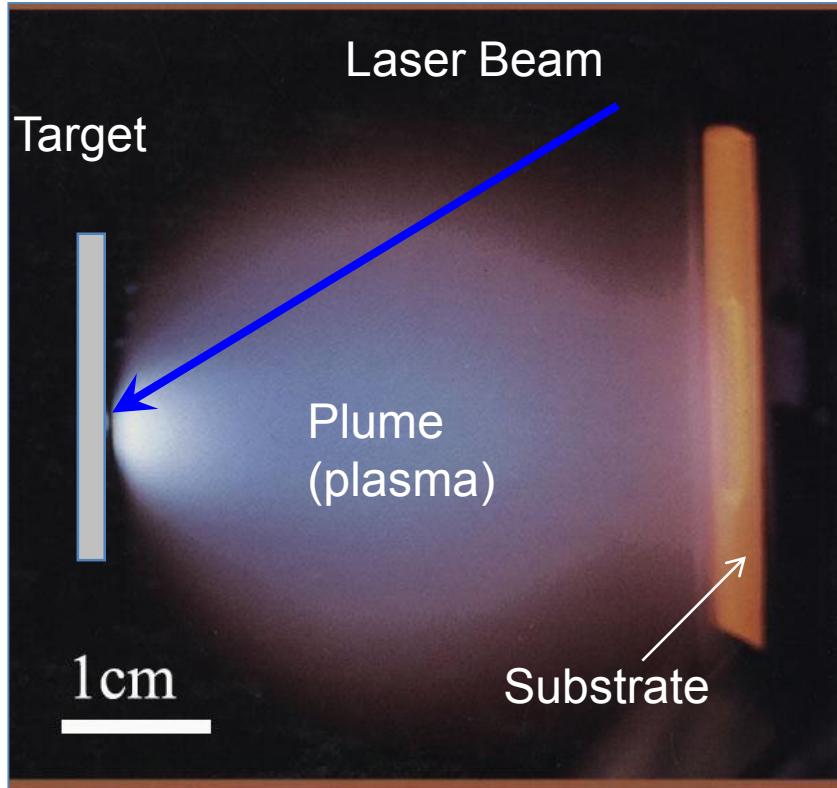
PLD Characteristics:

- Reproduce the composition of the target
- Fast response, well controlled deposition rate
- Flexible, easy to implement.

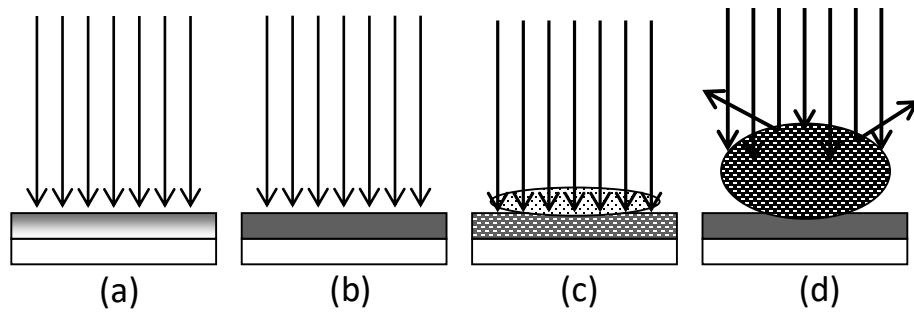
- Growth in any environment.
- Atoms arrive in bunches.
- Uneven coverage.
- High defect or particulate concentration.
- Not well suited for large-scale film growth.



Plume generated by laser ablation with different tiny or micro-particles

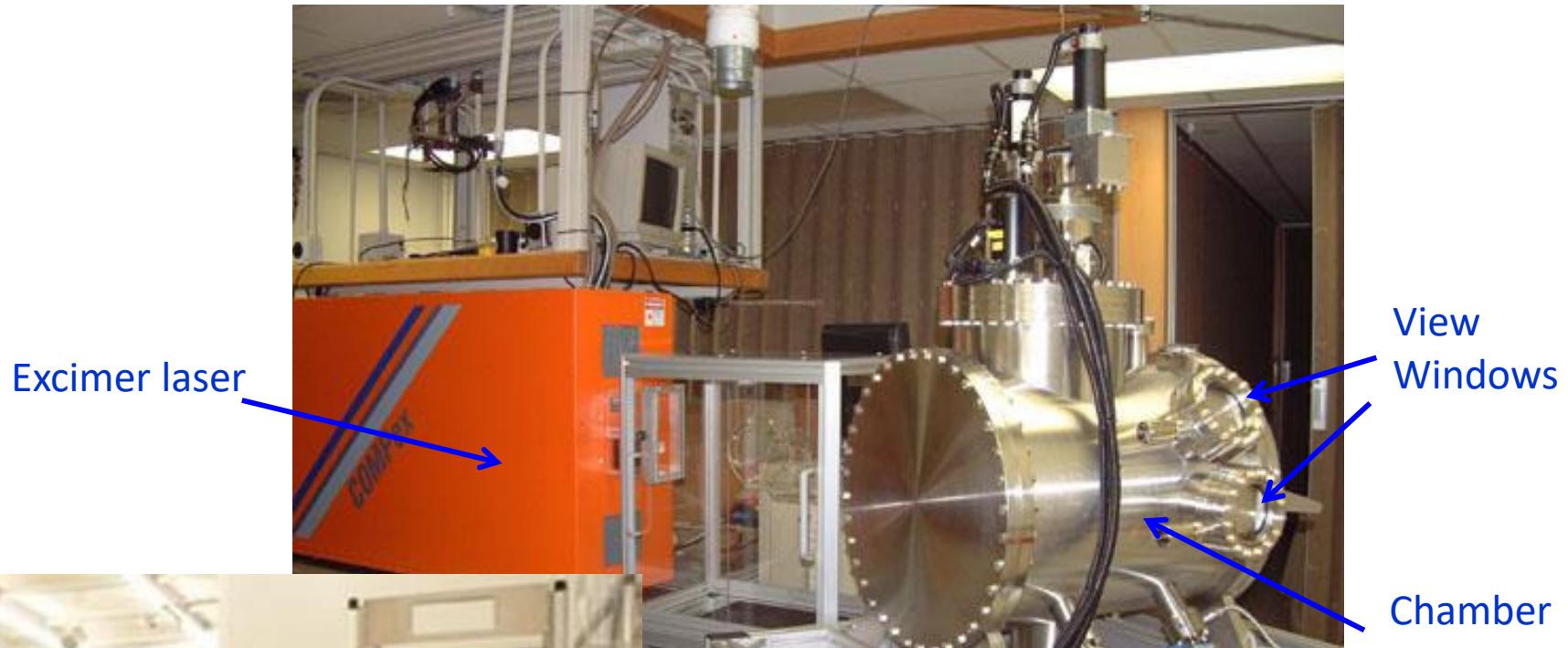


Pulse of fs to ns with peak power high enough (hundreds of MW/cm^2) to melt \rightarrow boil \rightarrow vaporize \rightarrow ablate the target surface material, to atoms, ions, electrons, and clusters.



Laser-material interaction. (a) Absorption and heating; (b) Melting and flowing; (c) Vaporization; (d) Plasma formation in front of the target. Under certain conditions the plasma can detach from the target and propagate toward the laser beam.

Pulsed laser deposition (PLD) system



Another system

Ceramic films deposited by PLD

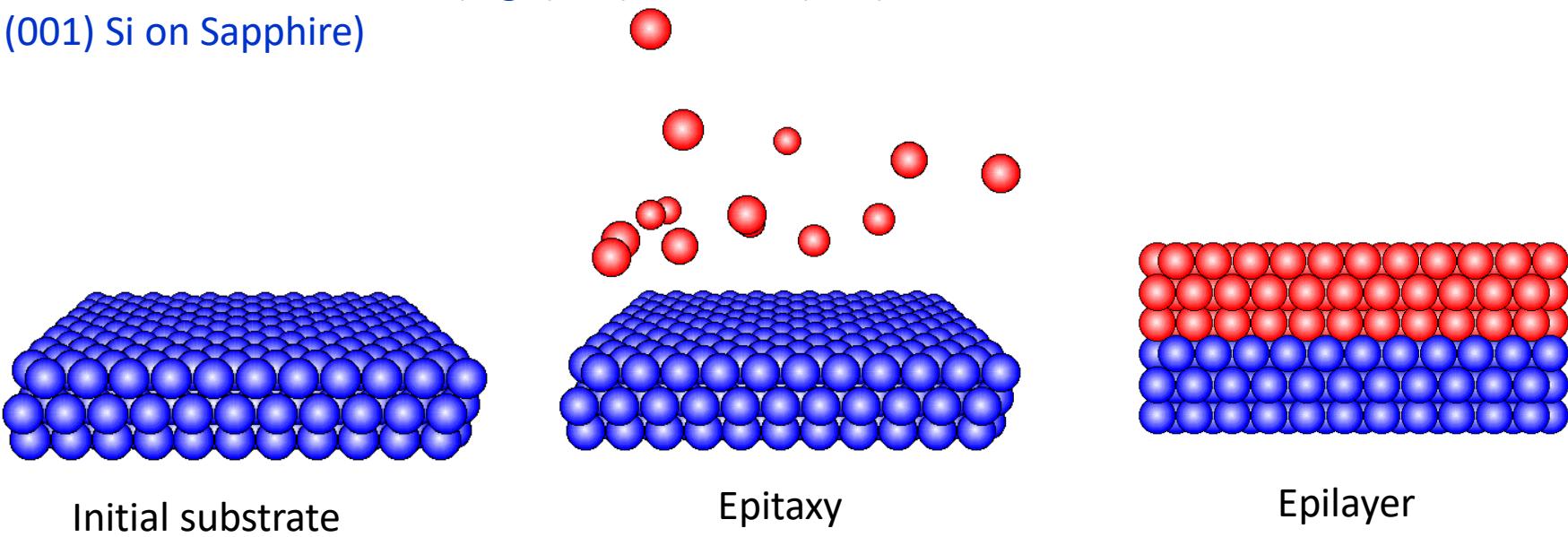
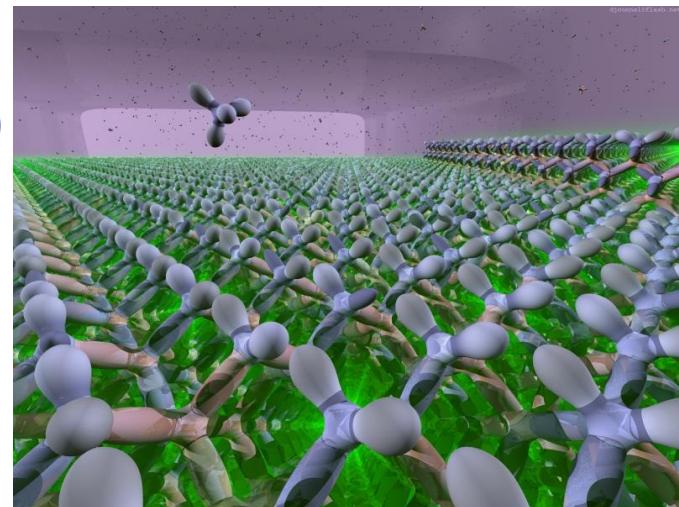
Property	Applications	Materials
High-temperature superconductivity	Microwave filters and delay lines, digital electronics, sensors	$\text{YBa}_2\text{Cu}_3\text{O}_7$, $\text{Tl}_2\text{Ca}_2\text{Sr}_2\text{Cu}_3\text{O}_x$, $\text{Nd}_{1.85}\text{Ce}_{0.15}\text{CuO}_4$
Ferroelectricity	DRAM capacitors, nonvolatile RAMS, optoelectronics, microwave devices	$\text{Pb}(\text{Zr})\text{TiO}_3$, $(\text{Sr}, \text{Ba})\text{TiO}_3$, $(\text{Sr}, \text{Ba})\text{Nb}_2\text{O}_6$, LiNbO_3
Ferrimagnetism	Circulators, phase shifters, magnetic recording, antennas	$\text{BaFe}_{12}\text{O}_{19}$, $\text{Y}_3\text{Fe}_5\text{O}_{12}$, $(\text{Mn}, \text{Zn})\text{Fe}_2\text{O}_4$, Li_2FeO_4
Electrochromic effects	Optical modulators, sunroofs, sensor protection	WO_3 , MoO_3 , V_2O_5
Electro-optical effects	Transparent conductors, solar energy, photovoltaics	F-doped ZnO_2 , $\text{In}_2\text{O}_3/\text{SnO}_2$, $(\text{La}, \text{Sr})\text{CoO}_3$
Piezoelectricity	Microelectrical-mechanical (MEM) devices	$\text{Pb}(\text{Zr})\text{TiO}_3$
Giant magnetoresistance	Magnetic recording head field sensors	$(\text{La}, \text{Ca})\text{MnO}_3$
Thermal and corrosive stability	Oxidation and thermal protection coatings for turbine blades	Y-ZrO_2 , MgAl_2O_4
Friction and wear	Hard, low-friction, wear-resistant coatings	MoS_2 , BN, SiC, diamond-like carbon
Biocompatibility	Prostheses, hip/knee implants	Hydroxylapatite, Al_2O_3

Chapter 9 Thin film deposition

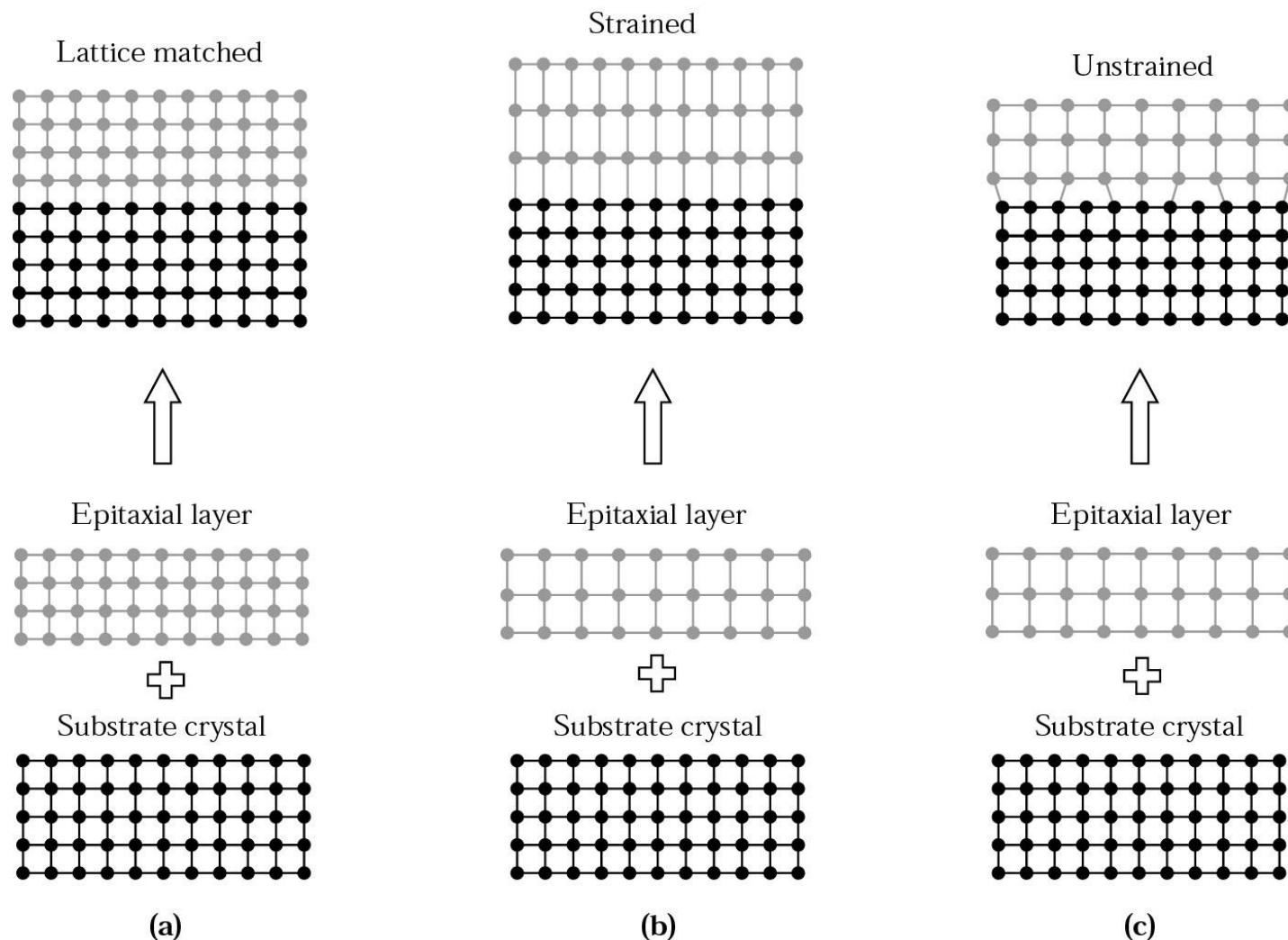
1. Introduction to thin film deposition.
2. Introduction to chemical vapor deposition (CVD).
3. Atmospheric Pressure Chemical Vapor Deposition (APCVD).
4. Other types of CVD (LPCVD, PECVD, HDPCVD...).
5. Introduction to evaporation.
6. Evaporation tools and issues, shadow evaporation.
7. Introduction to sputtering and DC plasma.
8. Sputtering yield, step coverage, film morphology.
9. Sputter deposition: reactive, RF, bias, magnetron, collimated, and ion beam.
10. Deposition methods for thin films in IC fabrication.
11. Atomic layer deposition (ALD).
12. Pulsed laser deposition (PLD).
13. Epitaxy (CVD or vapor phase epitaxy , molecular beam epitaxy).

Introduction

- Epitaxy refers to the method of depositing a monocrystalline film on a monocrystalline substrate.
- The word “epitaxy” comes from the Greek for ‘above’ (epi) and ‘in an ordered manner’ (taxis): to arrange upon.
- Autoepitaxy (or Homoepitaxy): extension of the substrate lattice by the overgrowth of a layer of identical material (e.g. Si on Si or GaAs on GaAs) with no problem of compatibility or lattice mismatch.
- Heteroepitaxy: any two materials of different crystalline structure and orientation (e.g. (001) GaAs on (001) Si or (001) Si on Sapphire)



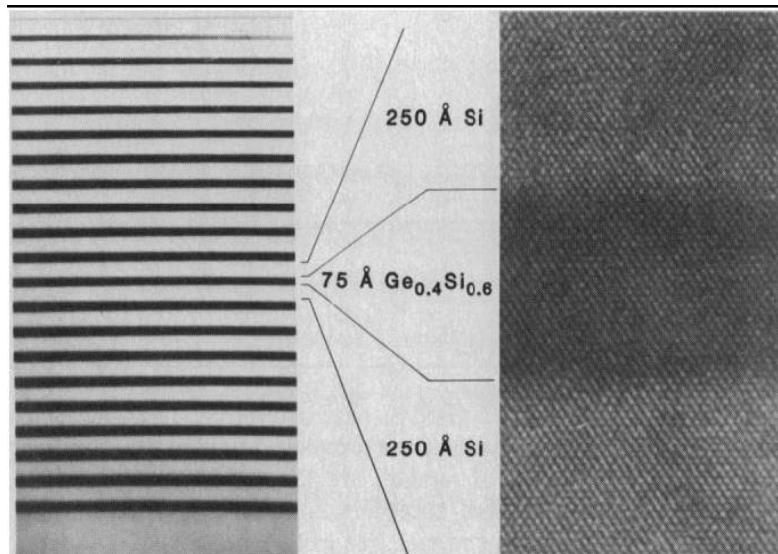
Strained and unstrained



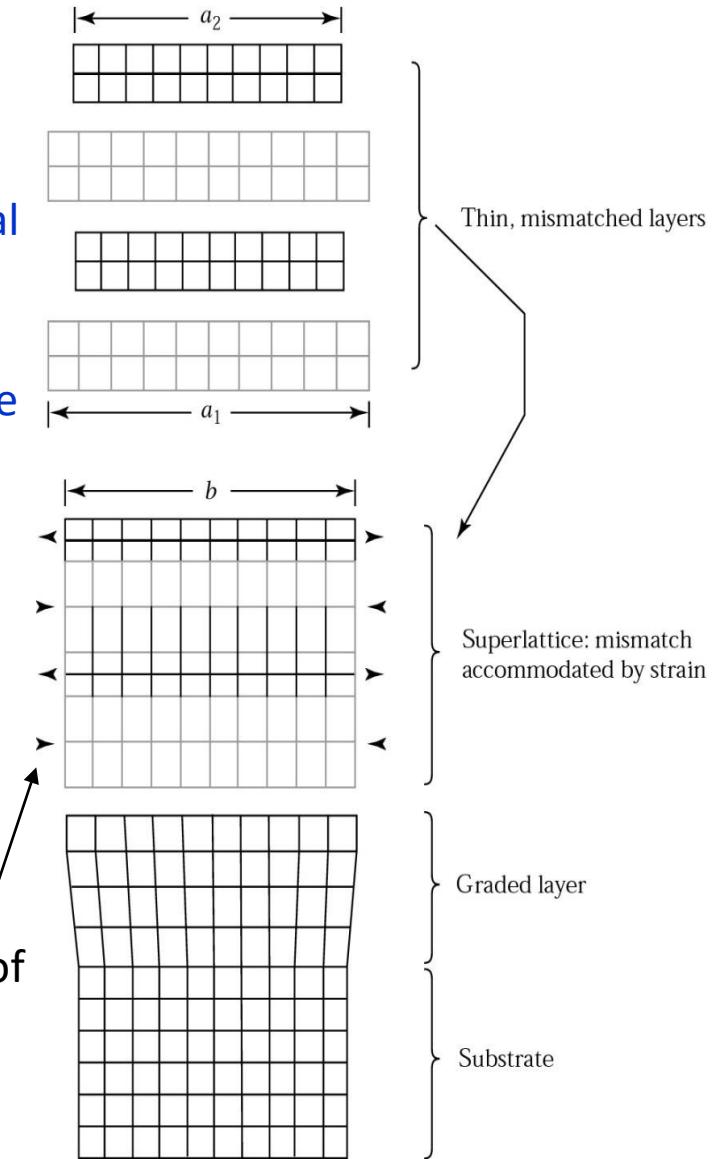
Schematic illustration of (a) lattice-matched, (b) strained, and (c) relaxed hetero-epitaxial structures. Homoepitaxy is structurally identical to the lattice-matched heteroepitaxy.

Strained-layer super-lattices

- Strained-layer super-lattices : in some devices we need repeated, regular alternation between two mono-crystalline.
- With these structures we can alter the basic physical properties of the material
- Because the layers are sufficiently thin, the lattice mismatch is accommodated by uniform strains in the layers.
- The new lattice will have an equilibrium lattice constant b such that $a_1 > b > a_2$ (for $a_1 > a_2$)



direction of
the strain



Growth methods

- Vapor-Phase Epitaxy (VPE, a form of CVD, e.g. MOCVD): transport of the epi-layer constituents (Si, Ga, As, dopants,...) in the form of one or more volatile compounds to the substrate where they react to form the epi-layer.
- Molecular Beam Epitaxy (MBE): physical transport of material to a heated substrate through vacuum evaporation (like PVD, no chemical reaction).
- Liquid-Phase Epitaxy (LPE): the growth of epitaxial layer on crystalline substrate by direct precipitation from the liquid phase.

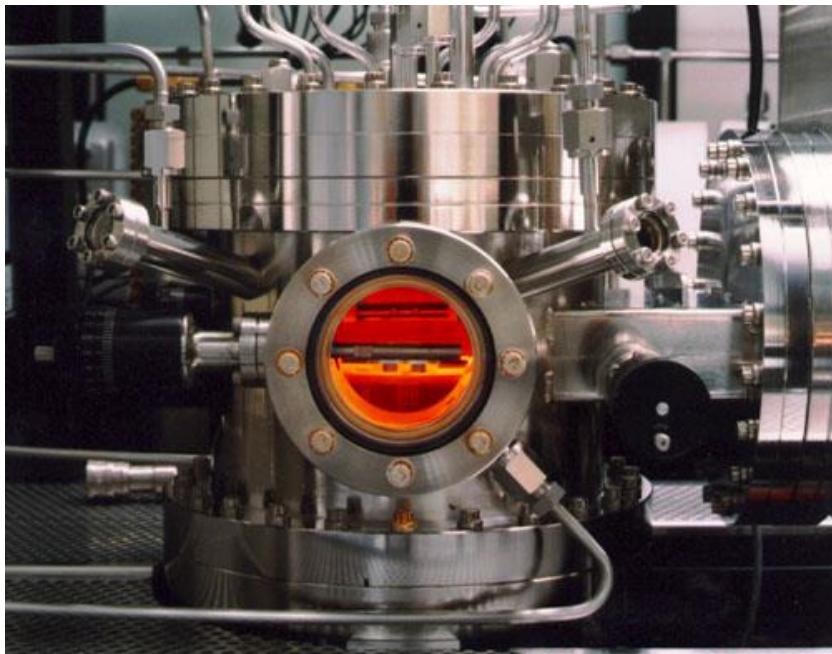
Liquid-phase epitaxy (LPE)

- LPE involves the growth of epitaxial layers on crystalline substrate by direct precipitation from the liquid phase.
- In LPE, the substrate is held in contact with the supersaturated solution (As saturated solution of Ga).
- Cooling the arsenic saturated solution of gallium causes the arsenic to precipitate in the form of GaAs.
- Typical deposition rates for monocrystalline films range from 0.1 to 1 $\mu\text{m}/\text{min}$.
- LPE in most cases is a very economic deposition techniques, especially when up-scaled to mass-production.

The organo-metallic process

MOCVD: metal-organic CVD = OMVPE: organo-metallic vapor phase epitaxy

- Many materials that we wish to deposit have very low vapor pressures and thus are difficult to transport via gases.
- One solution is to chemically attach the metal (Ga, Al, Cu, etc...) to an organic compound that has a very high vapor pressure.
- The organic-metal bond is very weak and can be broken via thermal means on wafer, depositing the metal with the high vapor pressure organic being pumped away.



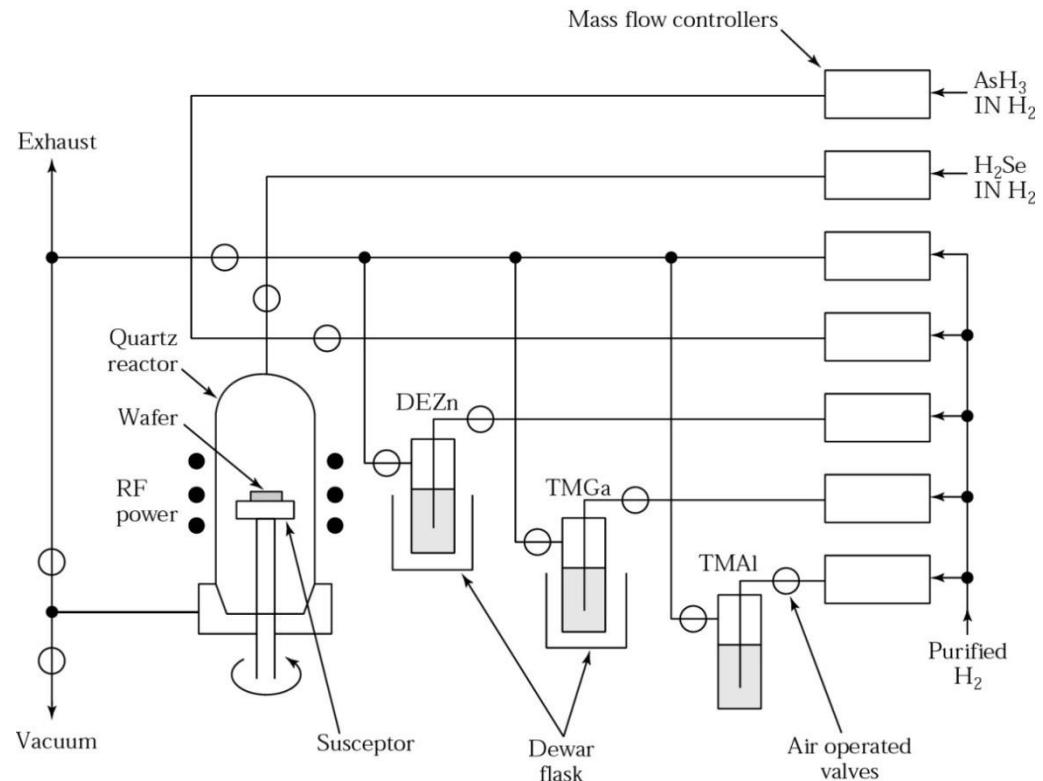
Some MOCVD precursor gases:
Tri-methyl-aluminum, liquid
Tri-methyl-gallium, liquid
Arsine AsH₃, gas
Di-methyl selenide, liquid
Di-methyl zinc, liquid

http://en.wikipedia.org/wiki/Metalorganic_vapour_phase_epitaxy

The organo-metallic process

Advantages: highly flexible, can deposit semiconductors, metals, dielectrics

Disadvantages: highly toxic, very expensive source material, and environmental disposal costs are high.



Material deposited:

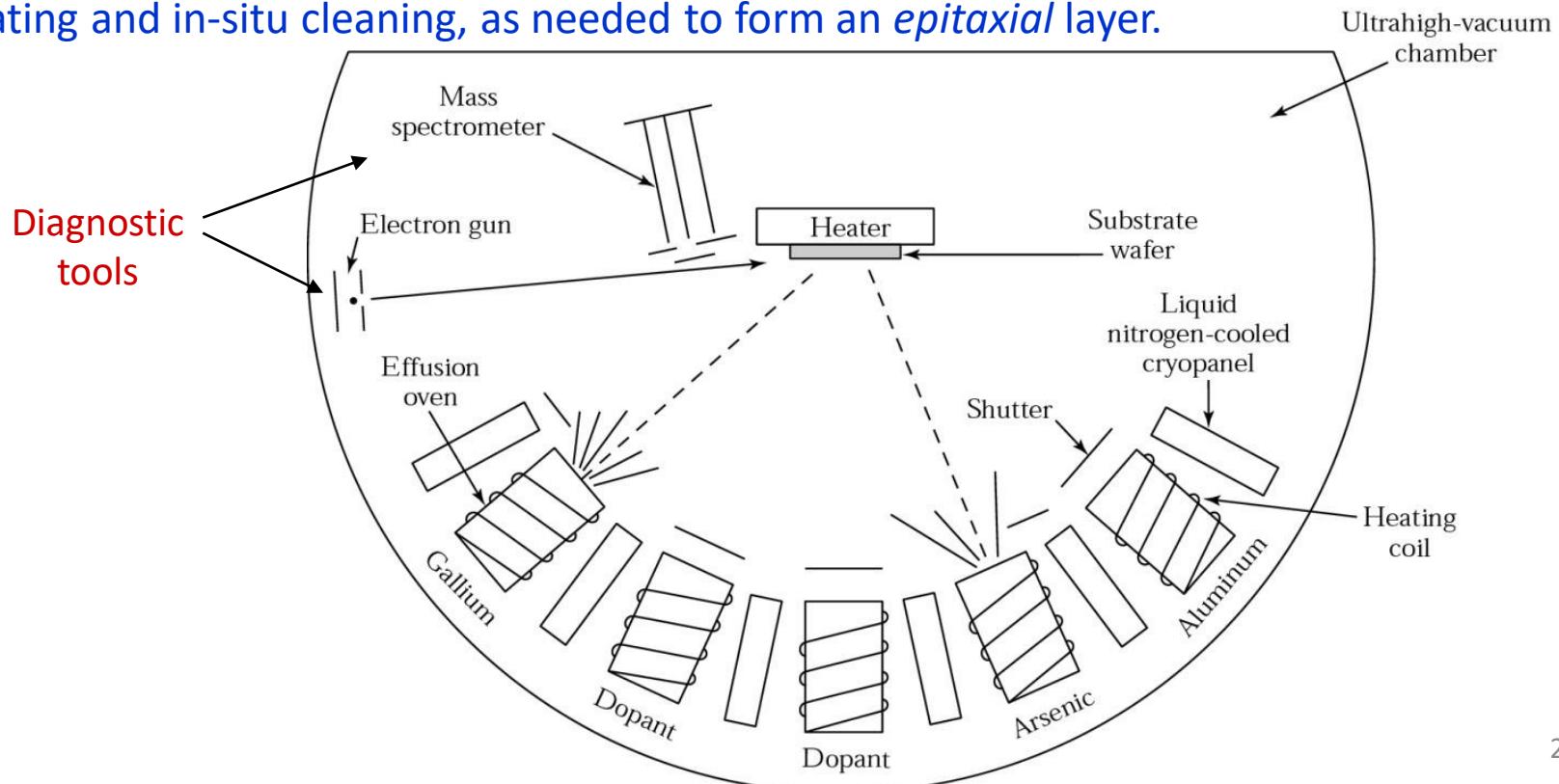
III-V semiconductors - AlGaAs, AlGalnP, AlGaN, AlGaP, GaAsP, GaAs, GaN, GaP, InAlAs, InAlP, InSb , InGaN, GaInAlAs, GaInAlN, GaInAsN, GaInAsP, GaInAs, GaInP, InN, InP.

II-VI semiconductors - Zinc selenide (ZnSe), HgCdTe, ZnO, Zinc sulfide (ZnS)

IV semiconductors - Si, Ge, strained silicon

Molecular-beam epitaxy (MBE)

- MBE is an epitaxial process involving the reaction of one or more thermal beams of atoms or molecules with a crystalline surface under UHV (ultra-high vacuum) conditions.
- Precise control in both chemical composition and doping profiles.
- It has a very low growth rate (e.g. for GaAs, a value of $1\mu\text{m}/\text{hr}$ is typical.)
- Single-crystal multilayer structures with dimensions on the order of atomic layers can be made.
- It is a kind of evaporation (PVD) process, but under ultra-high vacuum and with substrate heating and in-situ cleaning, as needed to form an *epitaxial* layer.



Effusion cell (Knudsen cell)

- The most common type of MBE source is the effusion cell. Sources of this type are sometimes called Knudsen, or K-cells.
- The crucible and source material are heated by radiation from a resistively heated filament. A thermocouple is used to allow closed-loop feedback control.
- A typical Knudsen cell contains a crucible (made of pyrolytic boron nitride (PBN), quartz, tungsten or graphite), heating filaments (often made of metal tantalum), water cooling system, heat shields and shutter.
- Knudsen cell is also used for thermal evaporation, but without precise control of temperature, and the cell is much more (by an order) simpler/cheaper.

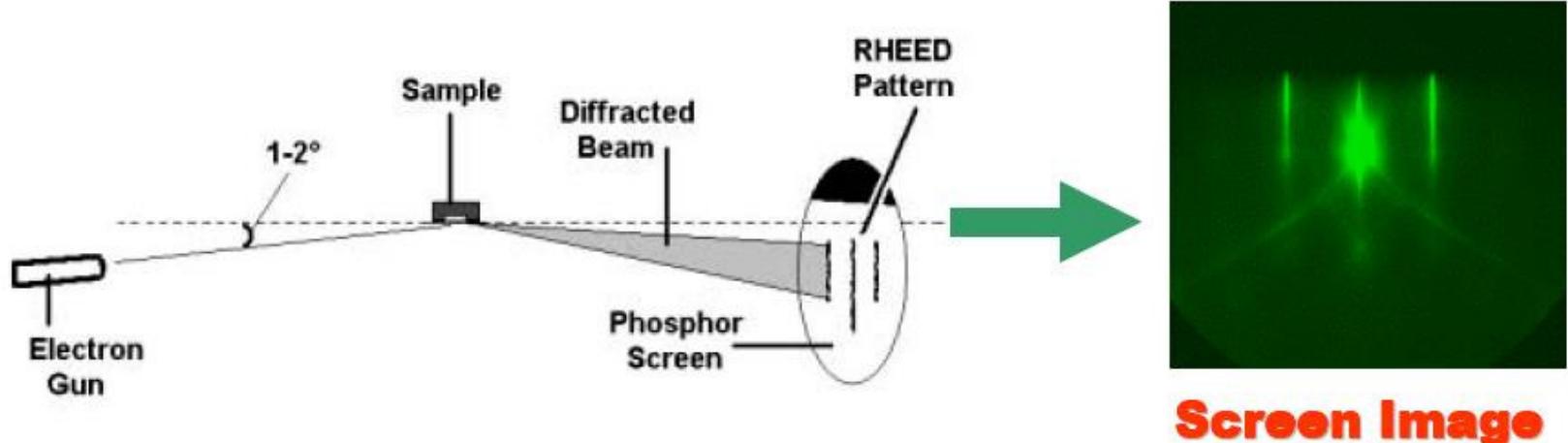


1. PBN crucible
2. Resistive heater filament
3. Metal foil radiation shields
4. Thermocouple
5. Mounting flange



RHEED

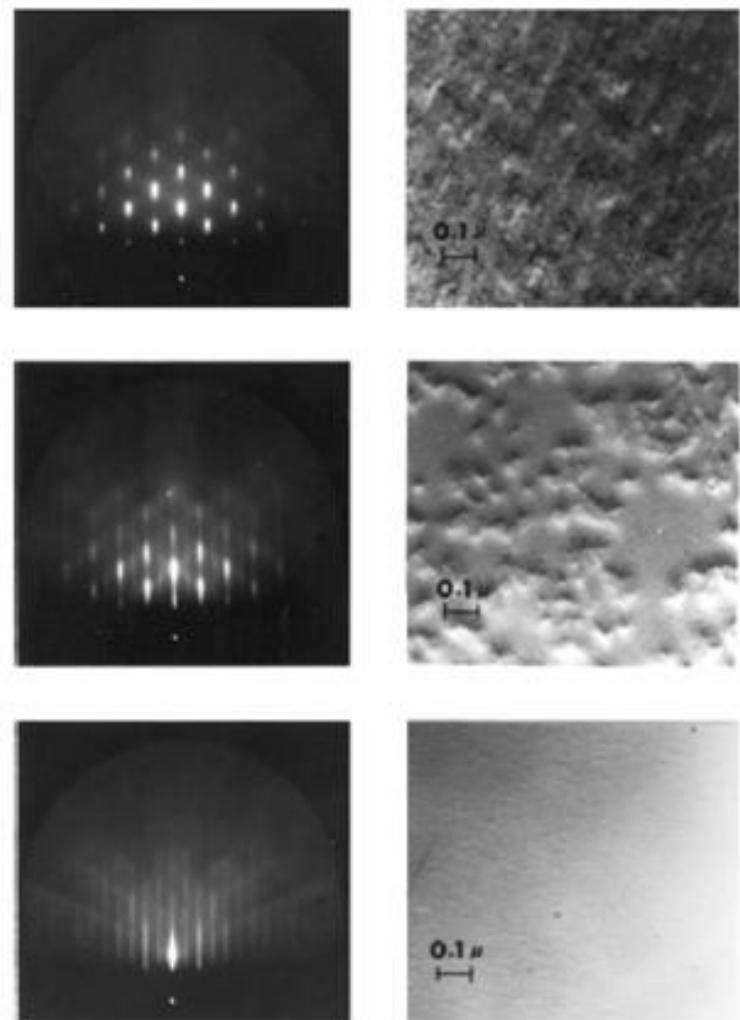
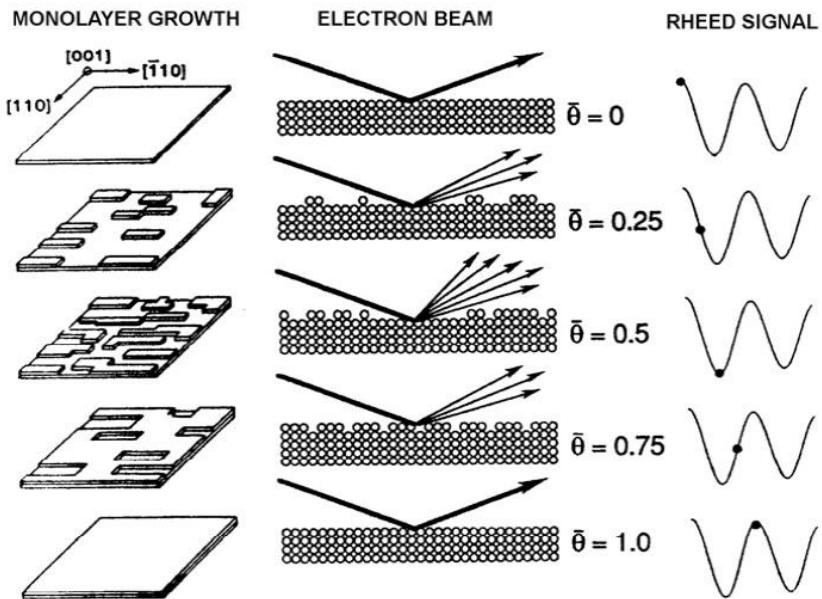
- Reflection high-energy electron diffraction (RHEED) is a technique used to characterize the surface of crystalline materials.
- A RHEED system consists of an electron source (gun), and a photoluminescent detector screen.
- The electron gun generates a beam of electrons which strike the sample at a very small angle relative to the sample surface.
- Incident electrons diffract from atoms at the surface of the sample, interfere constructively at specific angles and form regular patterns.



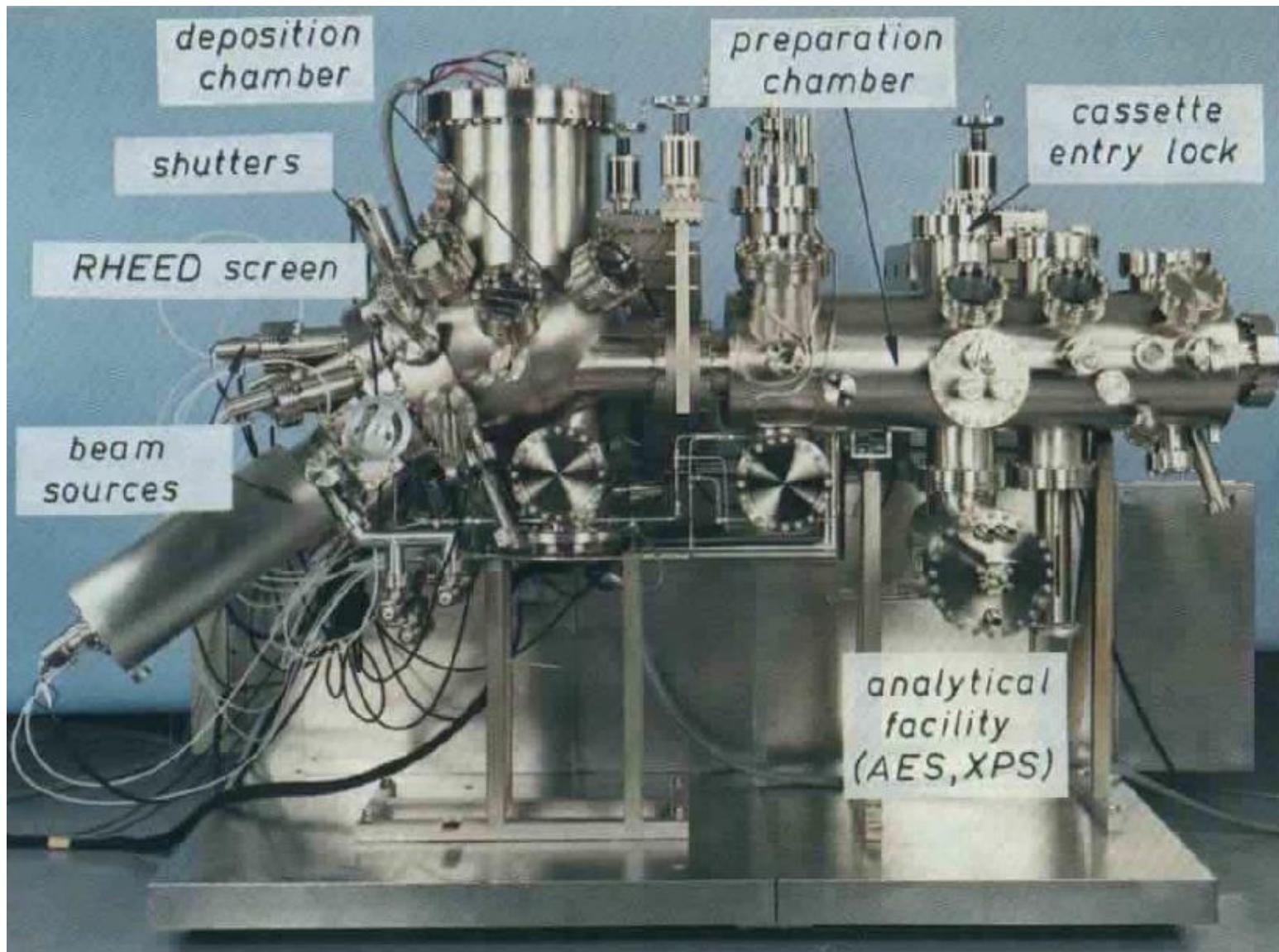
Screen Image

RHEED

The 2D surface giving rise to the commonly observed RHEED streaks.



MBE system



Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

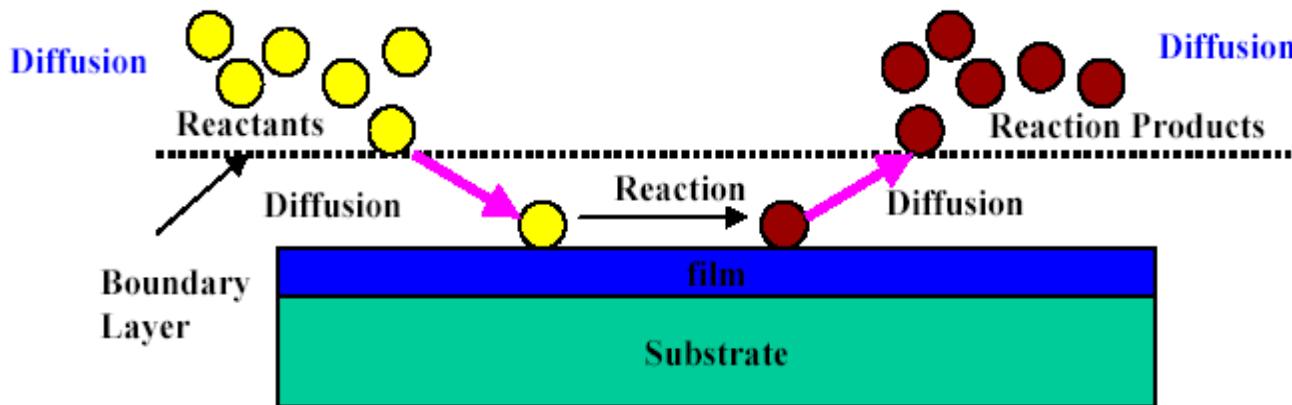
Material removal: etching processes

Etching is done either in “dry” or “wet” methods:

- Wet etching uses liquid etchants with wafers immersed in etchant solution.
- Wet etch is cheap and simple, but hard to control (not reproducible), not popular for micro-nanofabrication.
- Dry etch uses gas phase etchants in plasma, both chemical and physical (sputtering process).
- Dry plasma etch works for many dielectric materials and some metals (Al, Ti, Cr, Ta, W...).
- For other metals, ion milling (Ar^+) can be used, but with low etching selectivity. (as a result, for metals that cannot be dry-etched, it is better to pattern them using liftoff)

Etching is consisted of 3 processes:

- Mass transport of reactants (through a boundary layer) to the surface to be etched.
- Reaction between reactants and the film to be etched at the surface.
- Mass transport of reaction products from the surface through the surface boundary layer.



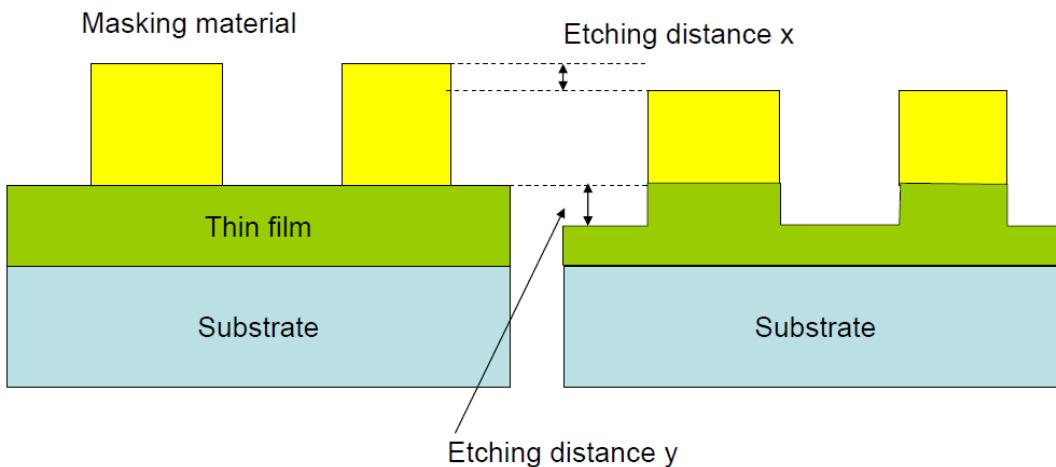
Figures of merit: etch rate, etch rate uniformity, selectivity, and anisotropy.

Figures of merit: selectivity

Etching selectivity:

- The ratio of etching rate between different materials, usually the higher the better.
- Generally, chemical etching has high selectivity, physical etching (sputtering, milling) has low selectivity.
- For fabrication, the selectivity is usually between film material and mask material, and is defined by S_{fm} , higher is better. (f: film; m: mask)

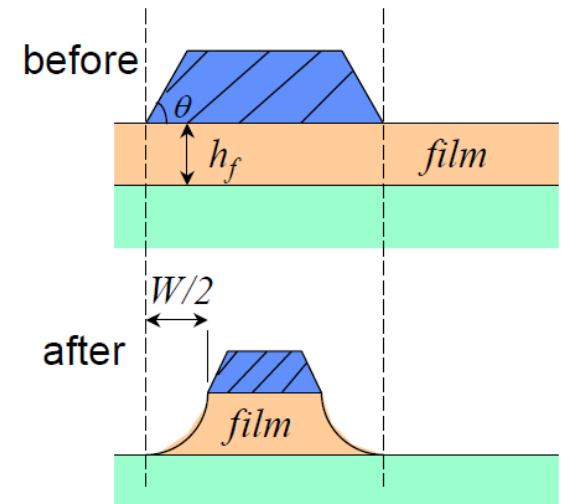
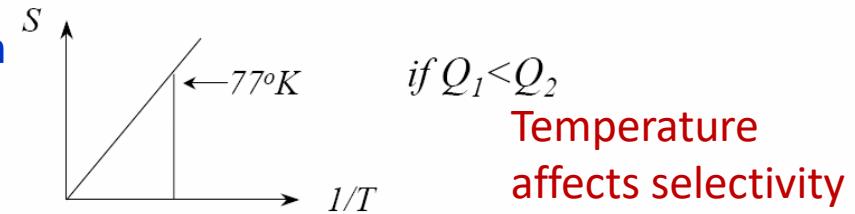
$$S_{fm} \equiv \frac{V_{f\perp}}{V_{m\perp}} \text{ (vertical components only)}$$



$$R_1 = A_1 e^{-Q_1/kT}$$

$$R_2 = A_2 e^{-Q_2/kT}$$

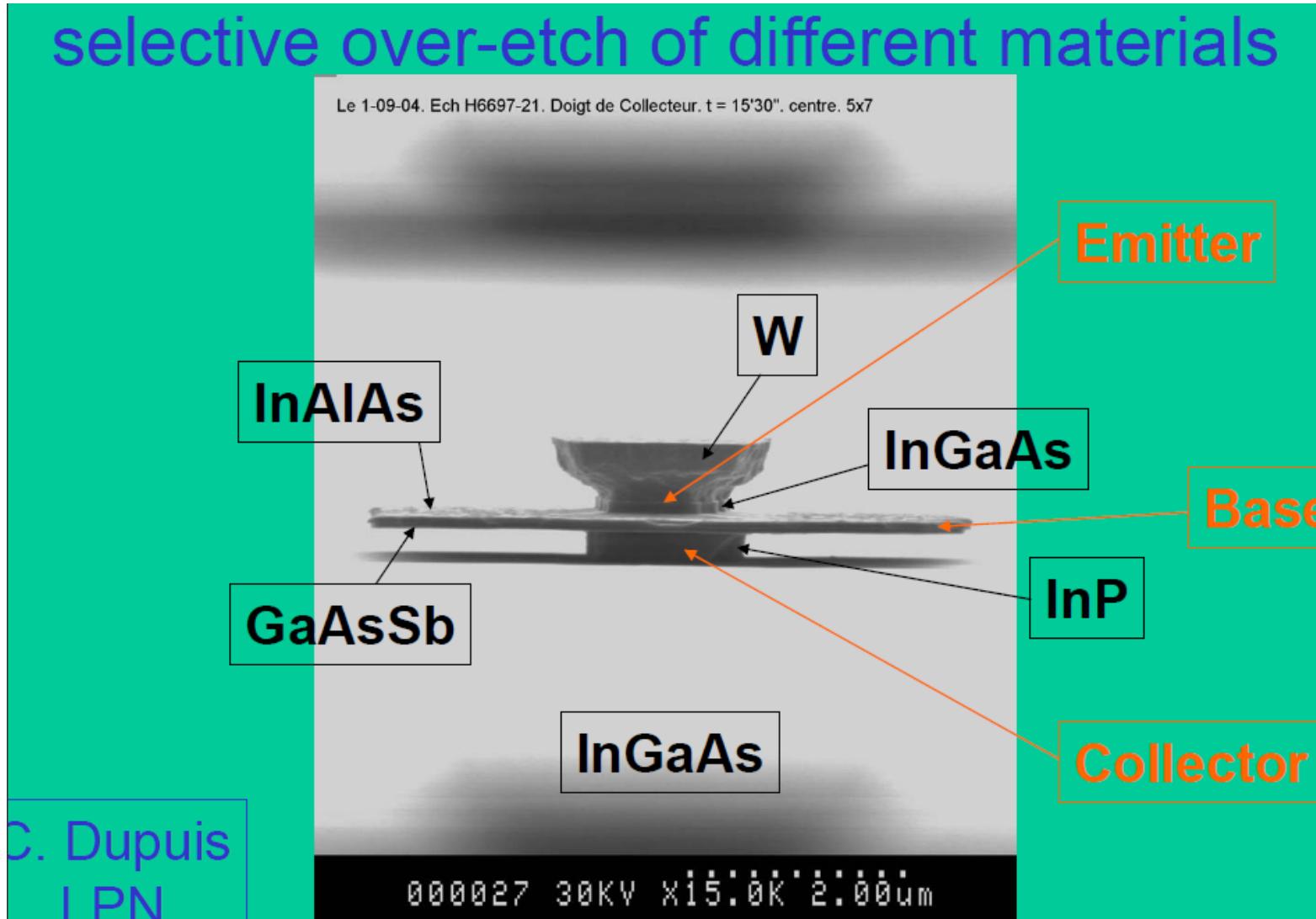
$$\therefore S = \frac{R_1}{R_2} = \frac{A_1}{A_2} e^{-(Q_1-Q_2)/kT}$$



Etching with mask erosion 3

Selective over-etch of different materials

The film is etched through to the bottom, plus over-etch to etch laterally for under-cut profile.



Figures of merit: anisotropy

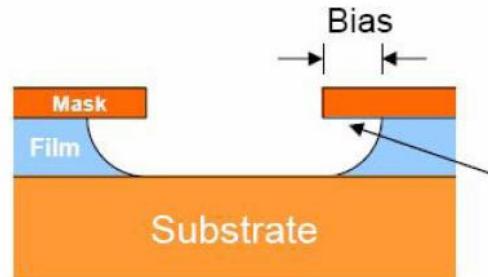
Isotropic: etch rate is the same along all directions. Similar to “conformal” used for deposition

Anisotropic: etch rate depends on direction, usually vertical vs. horizontal.

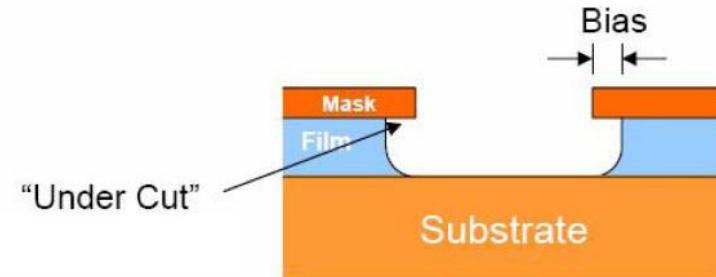
$$R_I = \frac{\text{Horizontal Etch Rate}(R_h)}{\text{Vertical Etch Rate}(R_v)}$$

For isotropic, $R_I=1$.

For complete
anisotropic, $R_I=0$.



($R_I = 1$, pattern dimension
is poorly defined)



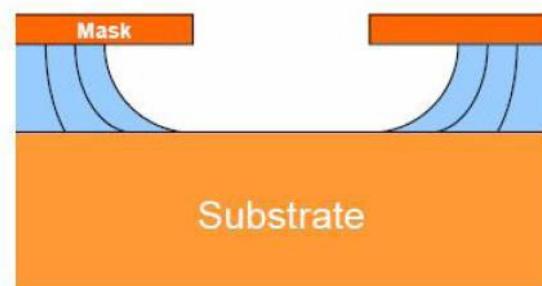
($R_I = 0.5$, pattern dimension
is better defined)



Over-Etch

- results in more vertical profile
but larger bias

CD: critical dimension



Worse in thick film

- Poor CD control in
thick film using wet etch

Figures of merit: anisotropy

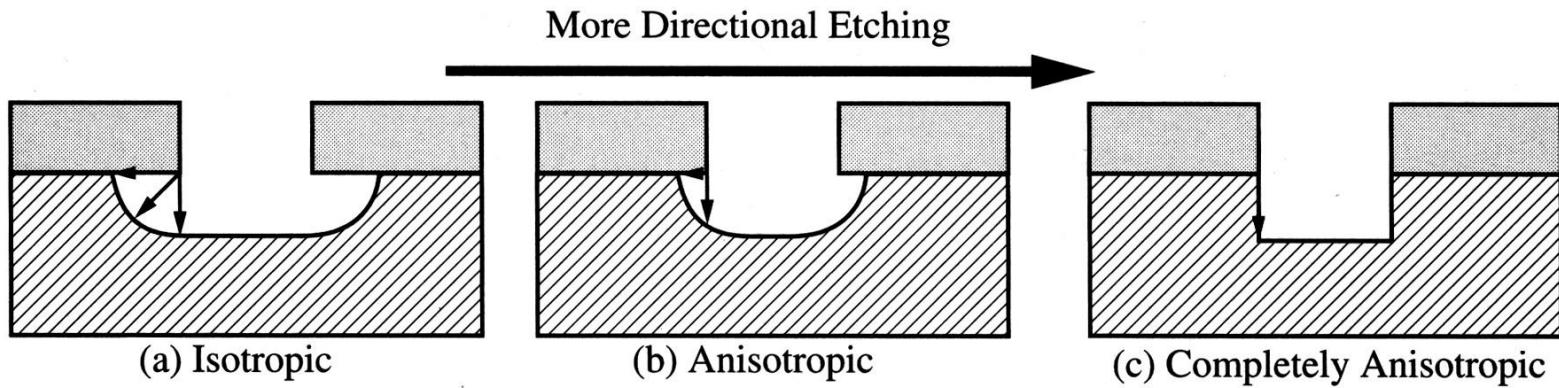


Figure 10-3 Etch profiles for different degrees of anisotropic, or directional, etching: (a) purely isotropic etching; (b) anisotropic etching; (c) completely anisotropic etching.

Generally speaking, chemical process (wet etch, plasma etch) leads to isotropic etch; whereas physical process (directional energetic bombardment) leads to anisotropic etch.

Isotropic:

- Best to use with large features when sidewall slope does not matter.
- Large critical dimension (CD, i.e. feature size) loss, generally not for nano-fabrication.
- Quick, easy, and cheap.

Anisotropic:

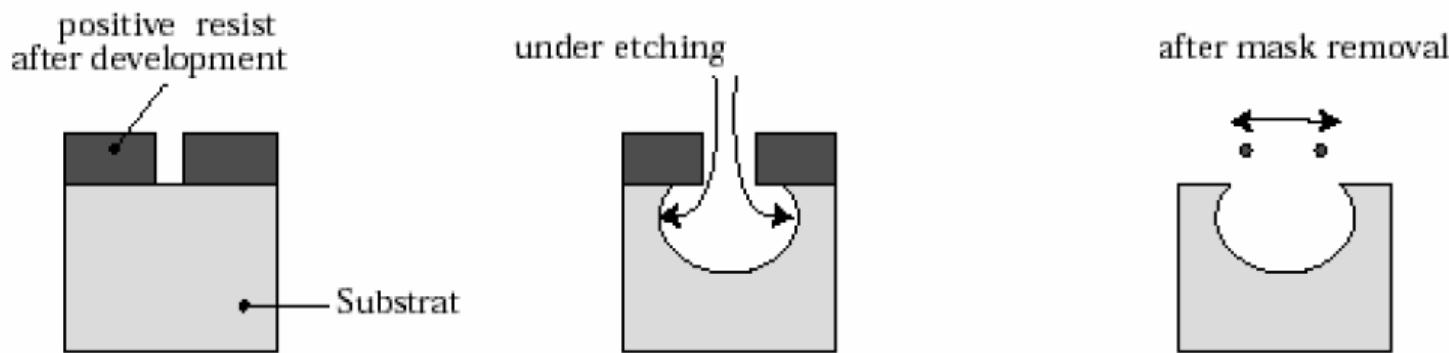
- Best for making small features with vertical sidewalls, preferred pattern transfer method for nano-fabrication and some micro-fabrication.
- Typically more costly.

Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Wet etching

- Wet etching was used exclusively till 1970's when feature size >3um.
- For today's IC industry, wet etching is used for noncritical feature sizes.
- **Advantages:** high selectivity, relatively inexpensive equipment, batch system with high throughput, etch rate can be very fast (many $\mu\text{m}/\text{min}$).
- **Disadvantages:** generally isotropic profile, high chemical usage, poor process control (not so reproducible), excessive particulate contamination.

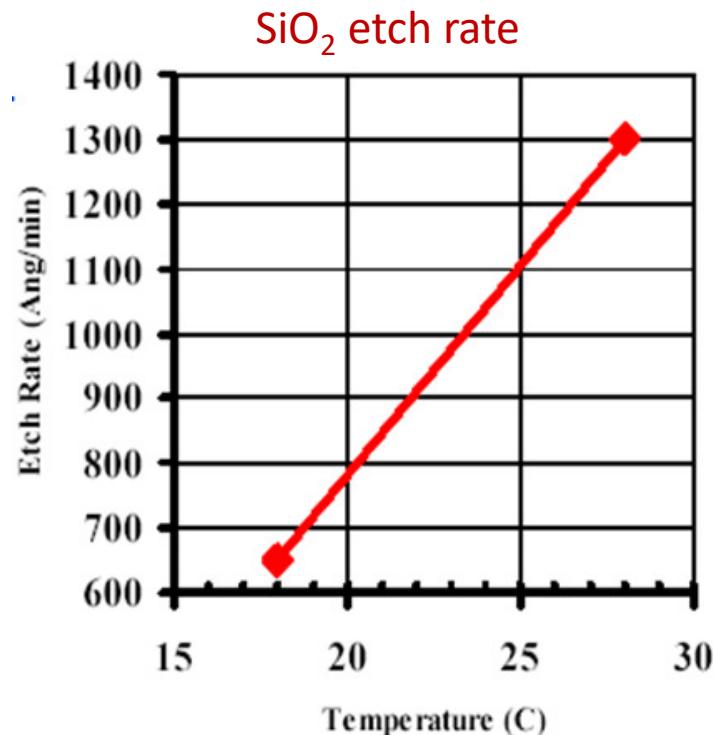


- The etch rate can be controlled by any of the three serial processes: reactants transport to the surface (depends on chemical concentration and stirring...), reaction rate (depends on temperature), reaction products transport from the surface (depends on stirring...).
- Preference is to have reaction rate controlled process because
 - Etch rate can be increased by temperature
 - Good control over reaction rate – temperature of a liquid is easy to control
- Mass transport control will result in non-uniform etch rate: edge etches faster.
- Etchant is often stirred to minimize boundary layer and make etching more uniform.

Isotropic wet etching (silicon dioxide)



- Etch is isotropic and easily controlled by dilution of HF in H₂O.
- Thermally grown oxide etches at
 - 120nm/min in 6H₂O:1HF
 - ~1 μm/min in 49 wt% HF (i.e. undiluted, as purchased HF).
- Faster etch rate for doped or deposited oxide.
- High etch selectivity (SiO₂/Si) > 100
- Buffered HF (BHF) or buffered oxide etchant (BOE) provides consistent etch rate
 - In regular HF etches, HF is consumed and the etch rate drops.
 - HF buffered with NH₄F to maintain HF concentration, typically 6NH₄F : 1HF
$$\text{NH}_4\text{F} \rightarrow \text{NH}_3 \uparrow + \text{HF}$$



HF is very dangerous! Because:

- It is not a so strong acid (you don't feel the pain for diluted HF).
- Deceptive: it looks just like water.
- It penetrates skin and attacks slowly the flesh and the bone.
- It might be too late when you begin to feel the pain.

Isotropic etch (silicon)

- Silicon is etched by nitric acid and hydrofluoric acid mixtures.
- HNO_3 partially decomposes to NO_2 , which oxidizes the surface of Si.

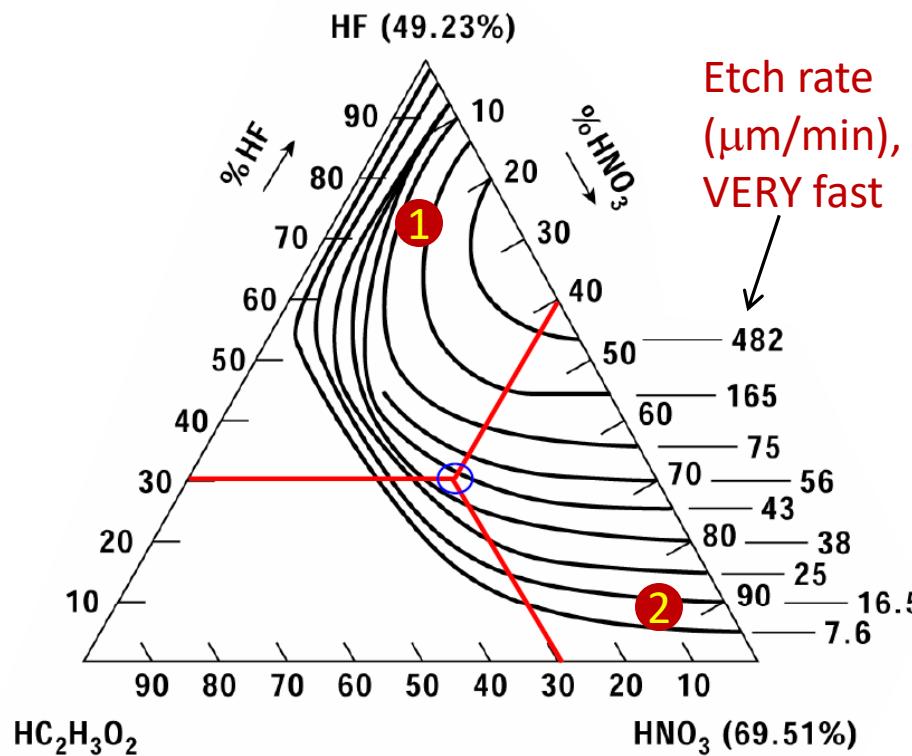


- The HF then dissolves the SiO_2 . The overall reaction is:



- Excess nitric acid results in a lot of silicon dioxide formation and etch rate becomes limited by HF removal of oxide (polishing).
- CH_3COOH (acetic acid) or H_2O can be added as diluent, but etch differently.
- Acetic acid is preferred because it prevents HNO_3 dissociation.

Si iso-etch curves



As-purchased HF is 49.23%, and HNO₃ 69.51%)

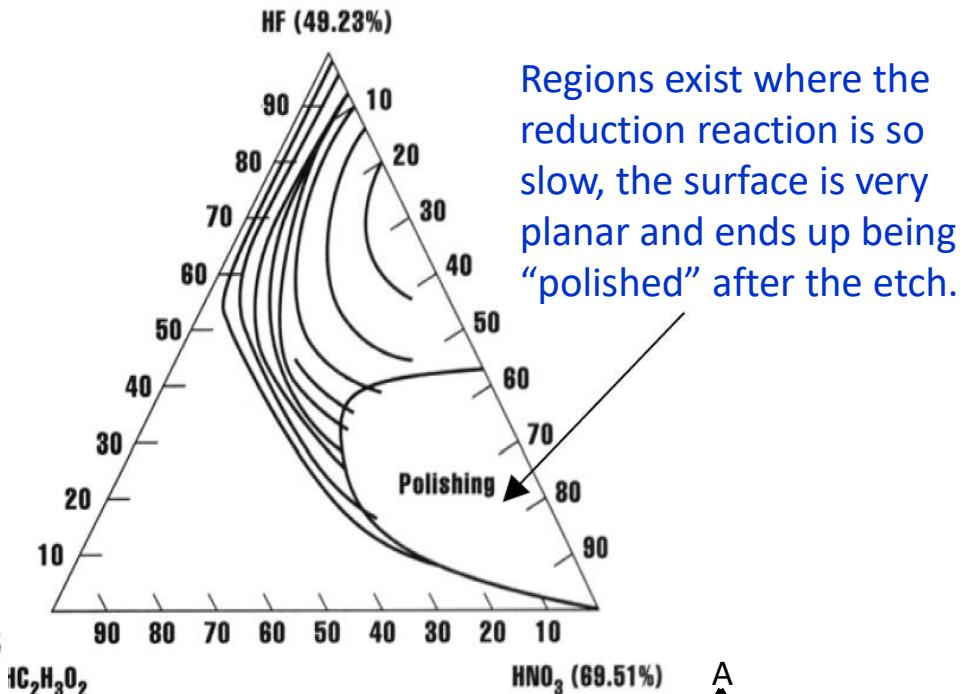
Region 1:

High HF concentrations, reaction limited by HNO₃, follow constant HNO₃% lines.

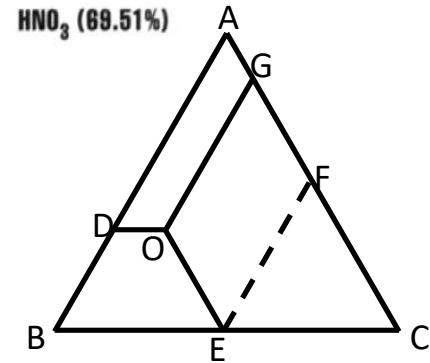
Region 2:

High HNO₃ concentrations, reaction limited by HF, follow constant HF % lines.

Rate limited by reduction, etched wafer surface have more oxide.



Need to show the total concentration is 100%
That is, AG+CE+BD=constant=AC=100%
This is the case, since CE=CF, BD=OE=GF

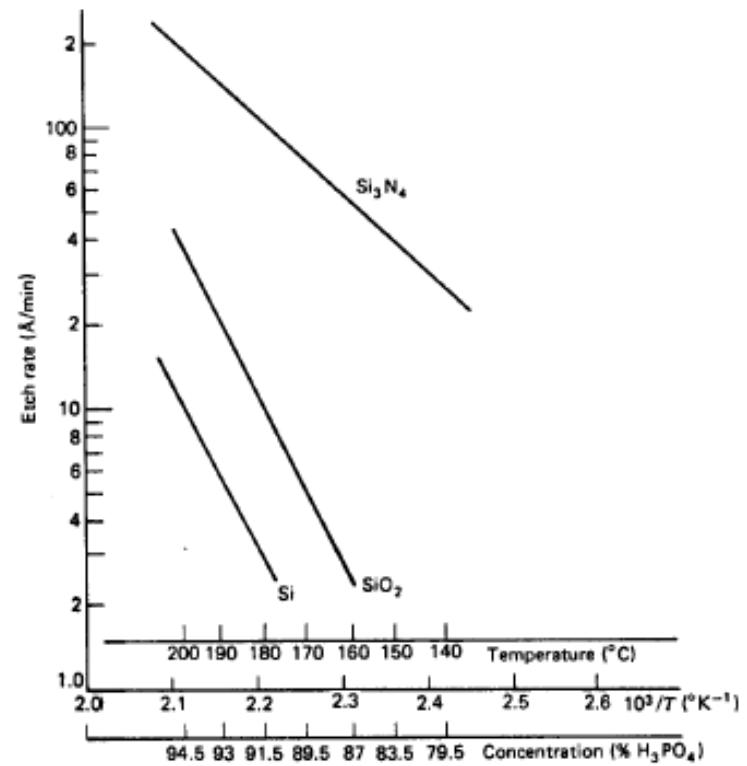


Isotropic etch (silicon nitride)

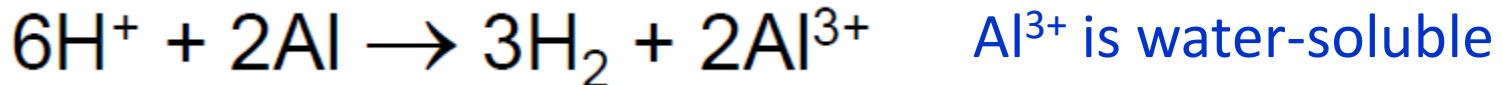
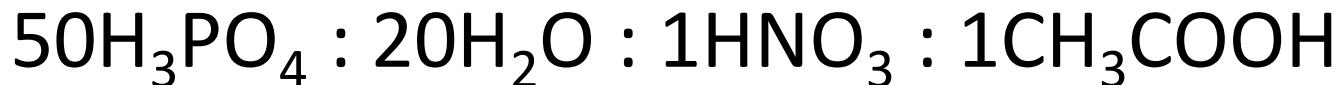


- Silicon Nitride is etched very slowly by HF solutions at room temperature, for example 20:1 BOE @20 °C
 - Etch rate of SiO_2 - 300 Å/min
 - Etch rate of Si_3N_4 - 5-15 Å/min
 - Very good selectivity of oxide to nitride
- Silicon nitride etches in 49% HF at room temperature at about 500 Å/min
- Phosphoric acid at 150 °C [140-200 °C] etches Si_3N_4 at fairly fast rate
 - Etch rate of Si_3N_4 - 100 Å/min
 - Etch SiO_2 - 10 Å/min
 - Selectivity of Si_3N_4 over SiO_2 : S = 10
 - Selectivity of Si_3N_4 over Si: S=30

Phosphoric Acid Etch Rate



Isotropic etching (aluminum)



- Aluminum etches in water, phosphoric, nitric and acetic acid mixtures.
- Converts Al to Al_2O_3 with nitric acid (evolves H_2).
- Dissolve Al_2O_3 in phosphoric acid.
- Gas evolution leading to bubbles.
- Local etch rate goes down where bubble is formed, leading to non-uniformity.
- Al can also be etched in (diluted) acid or base, such as HCl , HNO_3 , H_2SO_4 , NaOH or KOH , but less controllable (etch the native oxide slowly and un-controllably, then once oxide all etched away, etch Al metal very fast).
- The etch seems more repeatable in diluted HF (1:100 diluted).

Etchant

H_3PO_4 (19), Hac(1), HNO_3 (1), H_2O (2)
 HF , BOE ($\text{HF} + \text{NH}_4\text{F}$)
 H_2SO_4 (3), H_2O_2 (1) pirahna
 I_2 (1), KI (2), H_2O (10)
 NH_4OH (5), H_2O_2 (1)
 HNO_3 (64), NH_4F (3), H_2O (33)
 HCl (3), HNO_3 (1) (aqua regia)

Etches

Al, SiN, M
 SiO_2 , M
Organics, M
Au, M
Polymers, Al
Si, M
Au, other M

Doesn't etch

SiO_2 , Si, PR
Si, SiN, Au
Si, SiO_2 , SiN
Si, SiO_2 , SiN, M, PR
Si, SiO_2 , SiN, M
SiN, PR
Cr, Si, SiN, SiO_2

Summary

M: metal;

PR: photoresist;

Hac: acetic acid

Those are just starting point, can use different ratios.

E.g. the ratio for the Al etchant is different from previous slide.

In addition, most metals can be etched by (diluted) acid, except Ti and Cr that form a dense stable oxide on top.

Ti can be etched by HF; Cr by ceric ammonium nitrate plus acid.

Table 10-1 Common wet chemical etchants for various thin films used in IC fabrication

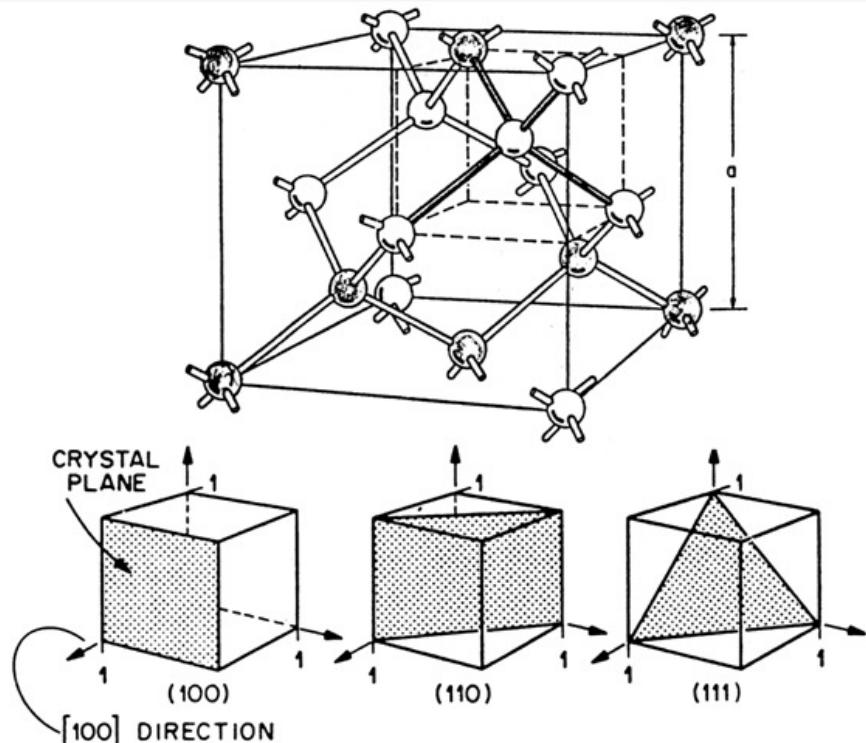
Material	Etchant	Comments
SiO_2	HF (49% in water) “straight HF” $\text{NH}_4\text{F:HF}$ (6:1) “Buffered HF” or “BOE”	Selective over Si (i.e., will etch Si very slowly in comparison). Etch rate depends on film density, doping. About $\frac{1}{20}$ th the etch rate of straight HF. Etch rate depends on film density, doping. Will not lift up photoresist like straight HF.
Si_3N_4	HF (49%) $\text{H}_3\text{PO}_4:\text{H}_2\text{O}$ (boiling @ 130–150°C)	Etch rate depends strongly on film density, O, H in film. Selective over SiO_2 . Requires oxide mask.
Al	$\text{H}_3\text{PO}_4:\text{H}_2\text{O}:\text{HNO}_3:\text{CH}_3\text{COOH}$ (16:2:1:1)	Selective over Si, SiO_2 , and photoresist.
Polysilicon	$\text{HNO}_3:\text{H}_2\text{O:HF}$ (+ CH_3COOH) (50:20:1)	Etch rate depends on etchant composition.
Single crystal Si	$\text{HNO}_3:\text{H}_2\text{O:HF}$ (+ CH_3COOH) (50:20:1) $\text{KOH:H}_2\text{O:IPA}$ (23 wt. % KOH, 13 wt. % IPA)	Etch rate depends on etchant composition. Crystallographically selective; relative etch rates: (100): 100 (111): 1
Ti	$\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (1:1:5)	Selective over TiSi_2 .
TiN	$\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (1:1:5)	Selective over TiSi_2 .
TiSi_2	$\text{NH}_4\text{F:HF}$ (6:1)	
Photoresist	$\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ (125°C) Organic strippers	For wafers without metal. For wafers with metal.

Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

An-isotropic wet etching of Si: overview

- Orientation selective etch of silicon occur in hydroxide solutions partly because of the closer packing of some orientations relative to other orientations
 - Density of planes: $\langle 111 \rangle > \langle 110 \rangle, \langle 100 \rangle$
 - Etch rate: $R(111) \ll R(110), R(100)$
- $\langle 100 \rangle$ direction etches faster than $\langle 111 \rangle$ direction, with etch rate
 - $R(100) = \text{few } 100 \times R(111)$
 - It is reaction rate limited
- Used very widely in MEMS (micro electro mechanical systems), since it is inexpensive, fast etching and easy to control.



{100} and {110} have 2 bonds below surface & 2 dangling bonds that can react.

{111} plane has three of its bonds below surface & only one dangling bond to react
→ much slower etch rate.

An-isotropic wet etching of Si

KOH etch example:

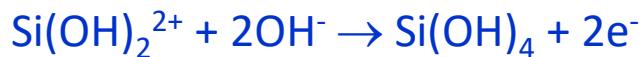
250 g KOH: 200 g 2-propanol, 800 g H₂O at 80°C

1000 nm/min of [100]

Etch stops at p++ layers

Selectivity: {111}:{110}:{100} ~ 1:600:400

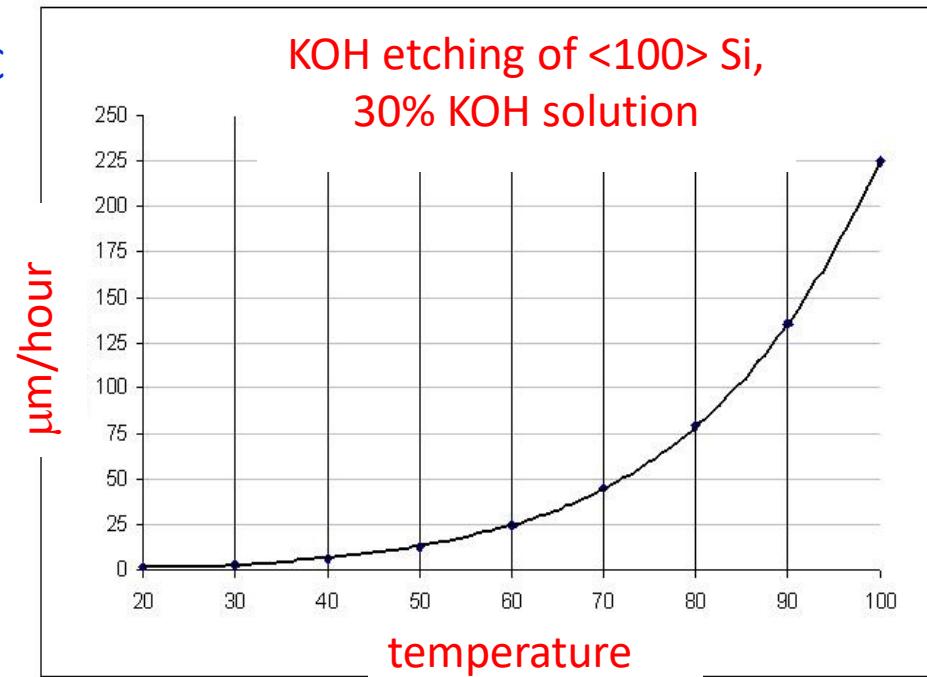
Seidel's etching model:



This is a model, real reaction is complicated.

Si(OH)₄ is soluble.

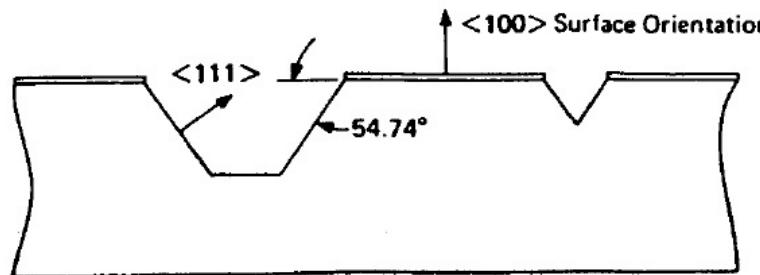
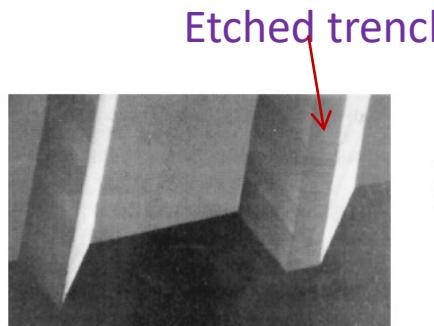
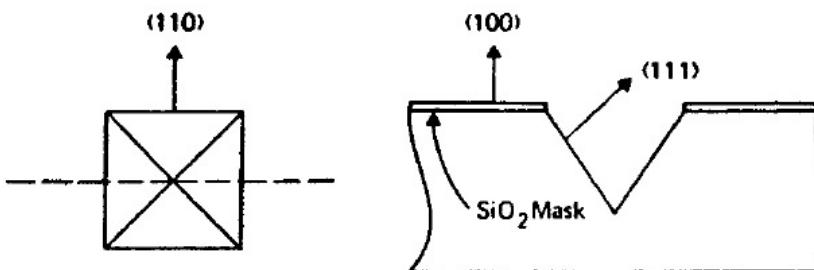
H₂ is generated and form bubbles.



Materials	Etchants	Etch Rates	
Silicon in <100> Silicon in <100>	KOH EDP	0.25 – 1.4 μm/min 0.75 μm/min	Rate drops Harder to etch
Silicon dioxide Silicon dioxide	KOH EDP	40 – 80 nm/hr 12 nm/hr	
Silicon nitride Silicon nitride	KOH EDP	5 nm/hr 6 nm/hr	
		EDP: see later slides	

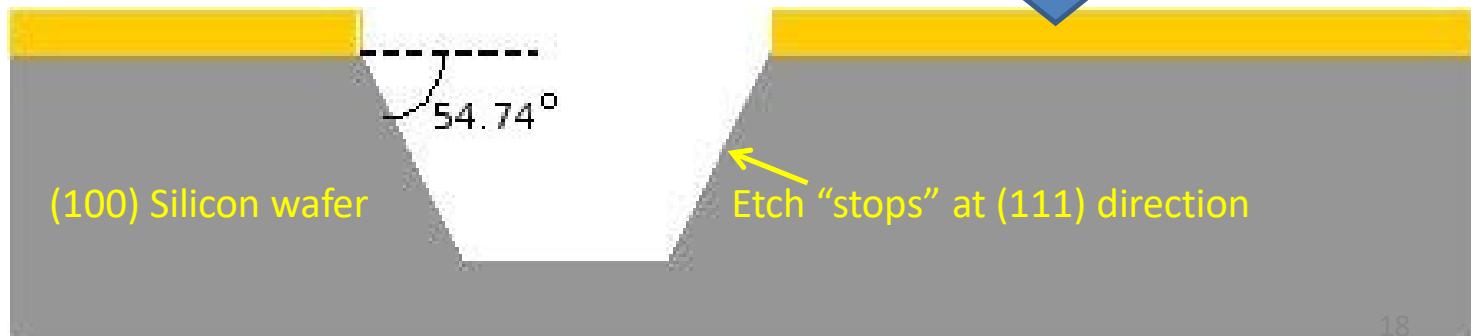
Examples: for (100) wafer

Effect of slow {111} etching with KOH: etching virtually stops at {111} plane.

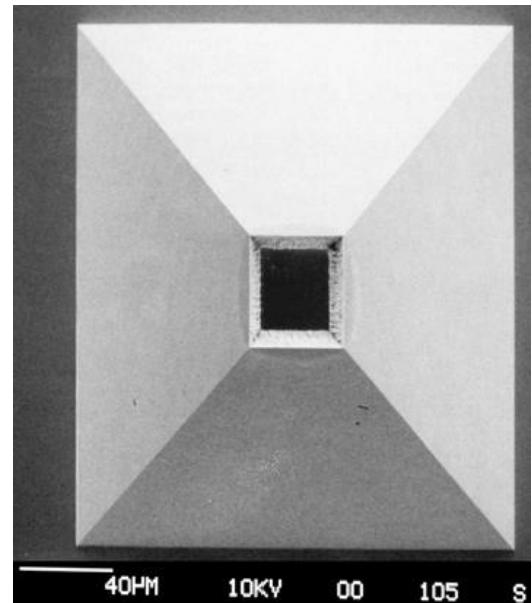
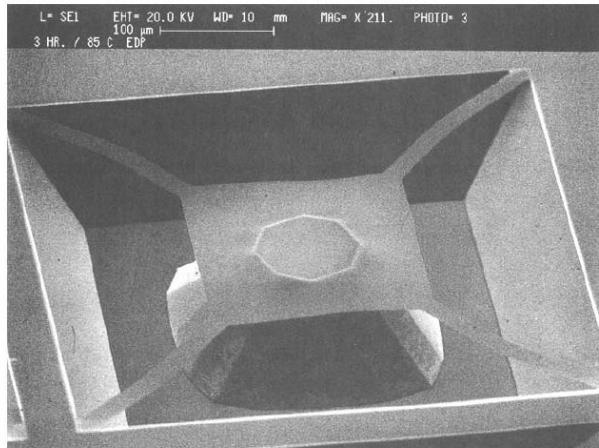
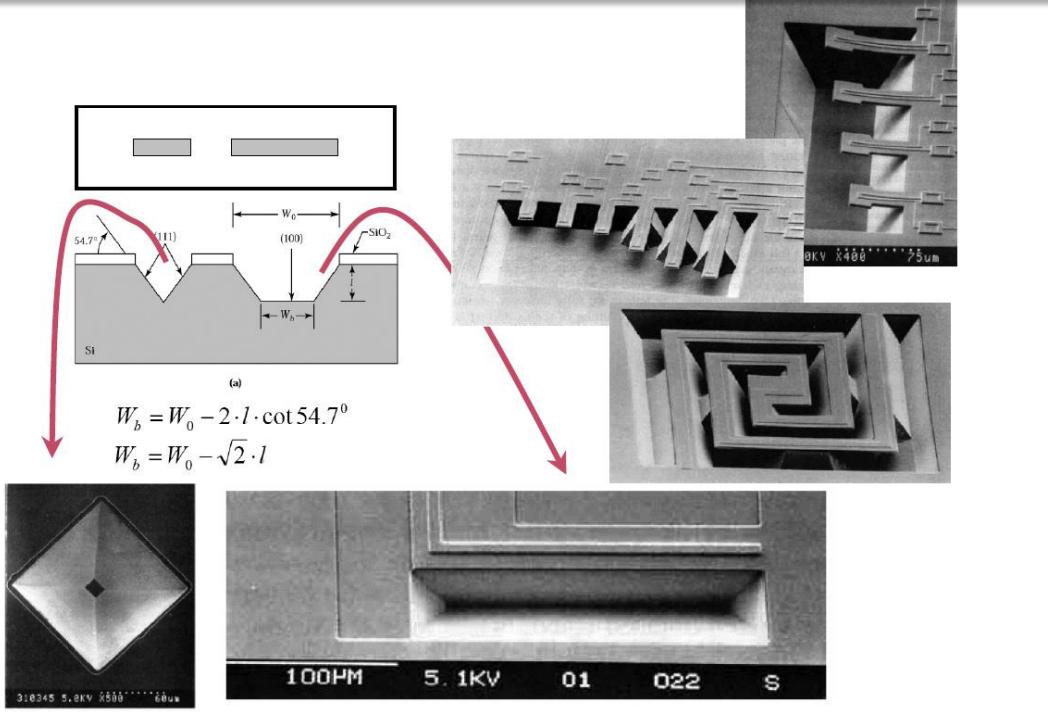


Etching selectivity to thermal oxide ~1000,
to *LPCVD* nitride ~infinity (>10⁴!).

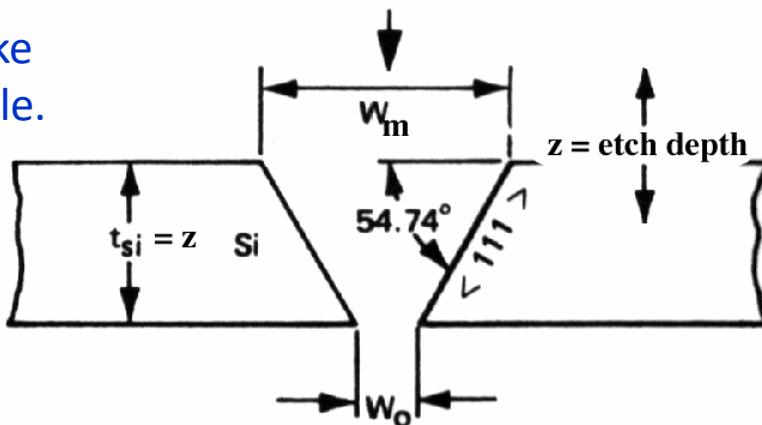
But KOH attacks
PECVD oxide
and nitride.



Examples: for (100) wafer

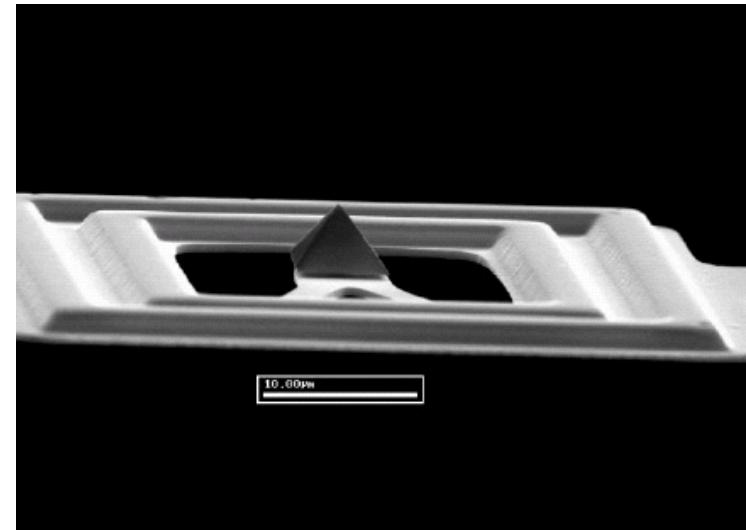
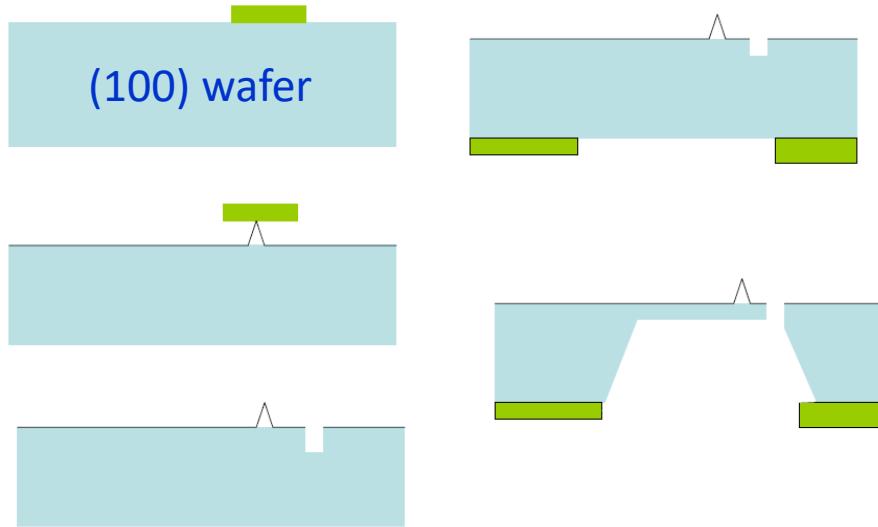


Have been used to make bubble-jet printer nozzle.

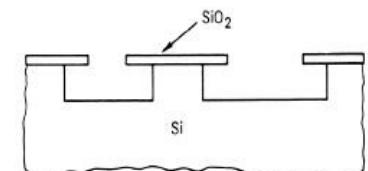
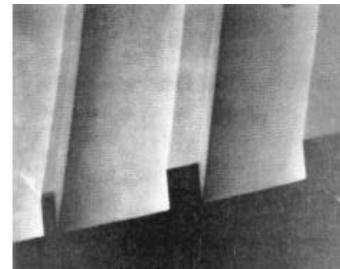
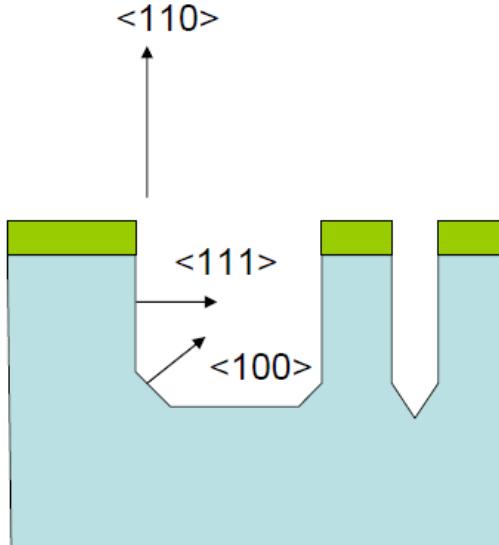


ANISOTROPIC ETCHING OF<100> SILICON

AFM (atomic force microscope) tips



For (110) (not (100))
wafer, vertical (not
tapered) trench
possible.



What is the
direction along the
grating lines?

- Bottom of pits are
- flat ($\{110\}$ plane) if KOH is used
 $\{100\}$ etches slower than $\{110\}$
 - V-shaped ($\{100\}$ planes) if EDP is used
 $\{110\}$ etches slower than $\{100\}$

Other anisotropic silicon etchants

Tetramethyl Ammonium Hydroxide (TMAH)

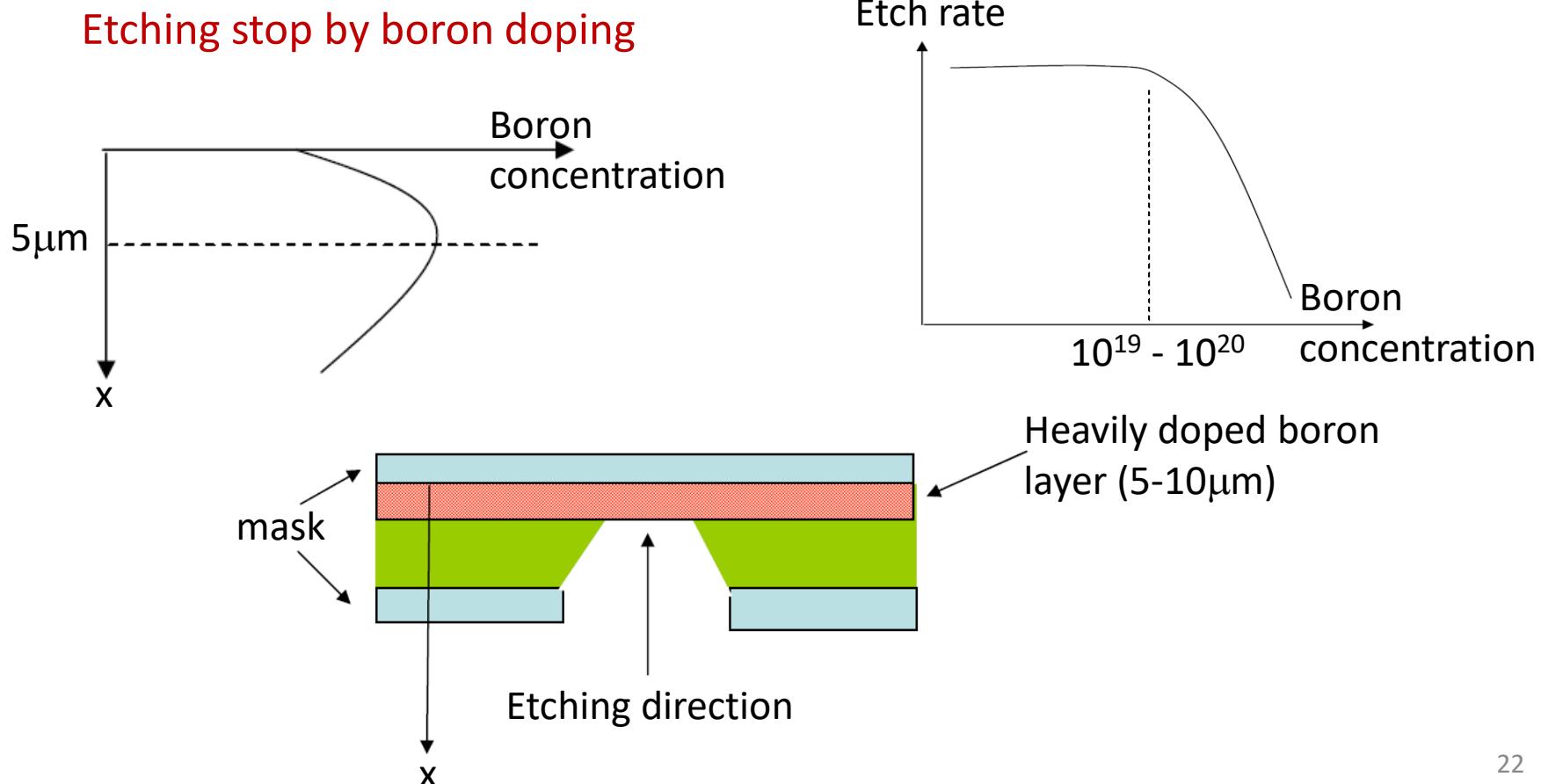
- Used widely as positive photoresist developer (since it contains no metal like K or Na, which are harmful for device.)
- Typical etching at 80-90°C.
- Etching rate ~0.5-1.5 µm/min (10-40%w.t)
- Selectivity $<100>$: $<111>$ ~ 10 - 35, much lower than KOH.
- Result in rough surface(H_2 bubble), KOH etch is smoother.
- Like KOH, attacks aluminum
- Like KOH, can use boron-stop-etching technique (etching rate decreases 40 times for $10^{20}/cm^3$ boron doping).
- Excellent selectivity of $<100>Si$: oxide/nitride (1: 5000-50000)

Ethylene Diamine Pyrochatechol (EDP)

- Typical etching temperature 115°C.
- Etching rate 1µm/min.
- Selectivity of $<100>Si$: oxide/nitride ~ 3000-7000.
- Doesn't attack metal (Au, Cr, Cu, Ta) but attacks Al.
- Selectivity $<100>$: $<111>$ ~35; (100) etches faster than (110), ((110) etches faster for KOH).
- Excellent for boron stop technique, etching rate drops 50 times for $7 \times 10^{19}/cm^3$ boron doping.
- Carcinogenic.

Etch stop

In wet etching process, etching depth is hard to control, so need etch stop layer. Besides oxide and nitride, etching may be stopped by controlled by doping: doped Si dissolved slower than pure Si.



Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Why dry etching?

Dry etching advantages

- Eliminates handling of dangerous acids and solvents
- Uses small amounts of chemicals
- Isotropic or anisotropic/vertical etch profiles
- Faithful pattern transfer into underlying layers (little feature size loss)
- High resolution and cleanliness
- Better process control

Dry etching disadvantages:

- Some gases are quite toxic and corrosive.
- Re-deposition of non-volatile compound on wafers.
- Expensive equipment (\$200-500K for R&D, few million for industrial tools).

Types of dry etching:

- Non-plasma based - uses spontaneous reaction of appropriate reactive gas mixture.
- Plasma based - uses radio frequency (RF) power to drive chemical reaction.

Non-plasma based dry etching

This is very rare. For example,



Although there is a large gain in free energy, the large activation energy does not allow low temperature processes - reaction is only effective above $\sim 800^\circ\text{C}$.

One exception is room temperature XeF_2 etching of Si. (same for BrF_3 & ClF_3)

Xenon di-fluoride (XeF_2) etching of Si:



- XeF_2 is a white powder, with vapor pressure ~ 3.8 Torr at 25°C .
- Isotropic etching, non-polish etching (rough)
- High selectivity for Al, SiO_2 , Si_3N_4 , photoresist, and PSG (phospho-silicate glass).
- Typical etch rate $\sim 1\mu\text{m/min}$
- Heat is generated during exothermic reaction
- XeF_2 reacts with water (or vapor) to form HF

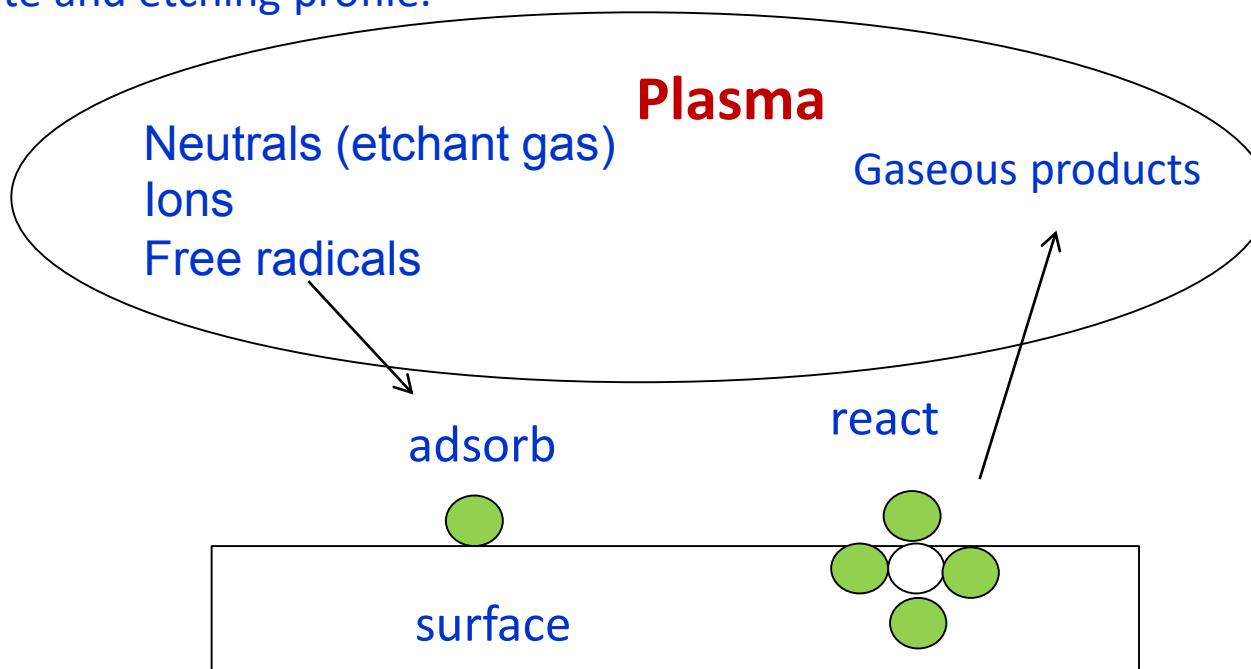
Gas phase etching, no stiction between freed structure and substrate (no liquid involved like KOH etch, so no need of drying that collapses pattern due to capillary force).

Popular for MEMS application.

MEMS: micro electro mechanical systems

Plasma-based etching

- Directional etching due to presence of ionic species in plasma and (self-) biased electric field. (The self-bias electric field is not applied *externally*, but is created *spontaneously* in RF plasma)
- Two components exist in plasma
 - Ionic species result in directional etching.
 - Chemical reactive species result in high etch selectivity.
- Control of the ratio of ionic/reactive components in plasma can modulate the dry etching rate and etching profile.

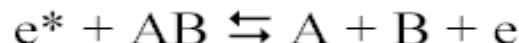


RF plasma chemistry

RF plasma is more widely used for dry etching than DC plasma.

For a plasma with inlet flow of molecule AB, Plasma processes are

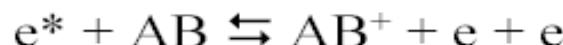
Dissociation



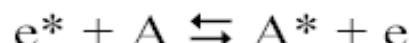
Atomic Ionization



Molecular Ionization



Atomic Excitation



Molecular Excitation

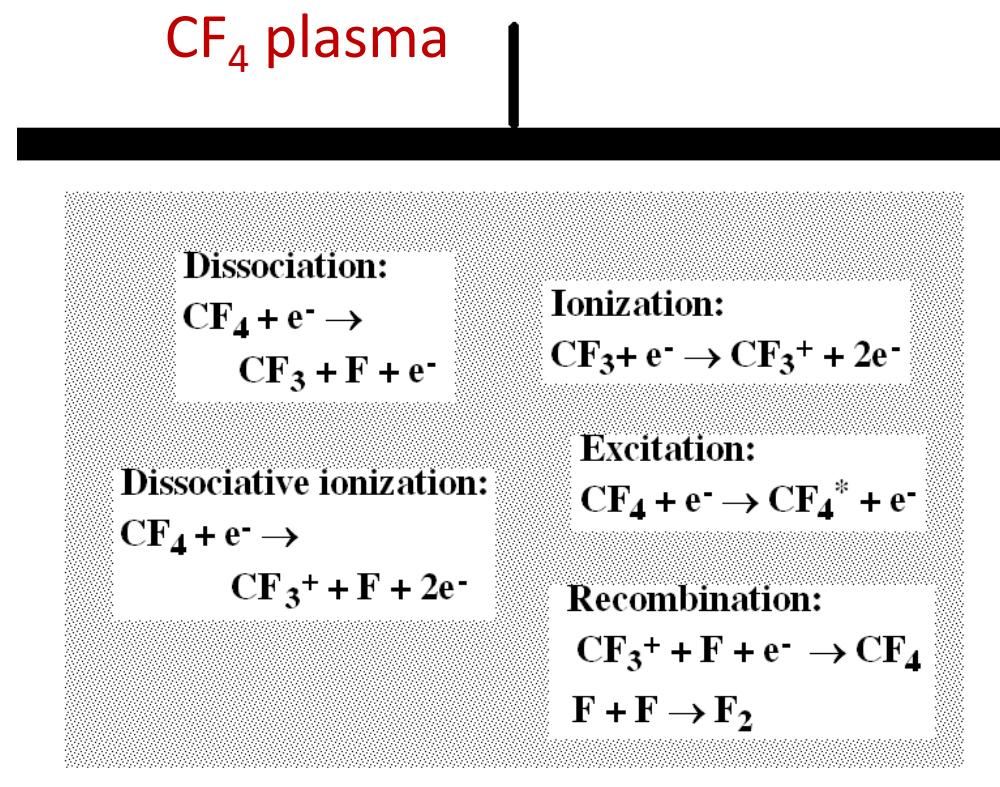
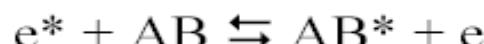


Figure 10-9

Loss mechanisms

- As seen in previous slide, in a plasma, unstable particles are continuously generated.
- The concentrations of ions, radicals, active atoms, & electrons increase until their loss rate is equal to the generation rate, forming a steady-state plasma.
- Recombination of ions and electrons: they attract each other and are annihilated.
- Drift, diffusion to walls: electrons are lost at conductive surfaces, chamber walls or electrodes. Ions are lost (converted to neutral particles) by contact with conductive surfaces, especially positive electrode.
- Recombination of radicals: e.g. $2O \rightarrow O_2$.
- Chemical reaction (what we want): e.g. $4F + Si \rightarrow SiF_4$ (fluorine radical combines with silicon wafer to produce silicon tetra-fluoride gas. This is a typical dry etching process.)

In equilibrium, degree of ionization typically 10^{-3} - 10^{-6} , very low, meaning majority gas not ionized.

(plasma density = number of ions/cm³ ~ typically 10^9 - $10^{13}/cm^3$.)

In a plasma TV, the recombination of ion-electron or radical, or de-excitation of atom or molecule, generates the colorful light we see through fluorescent process.



Chapter 10 Etching

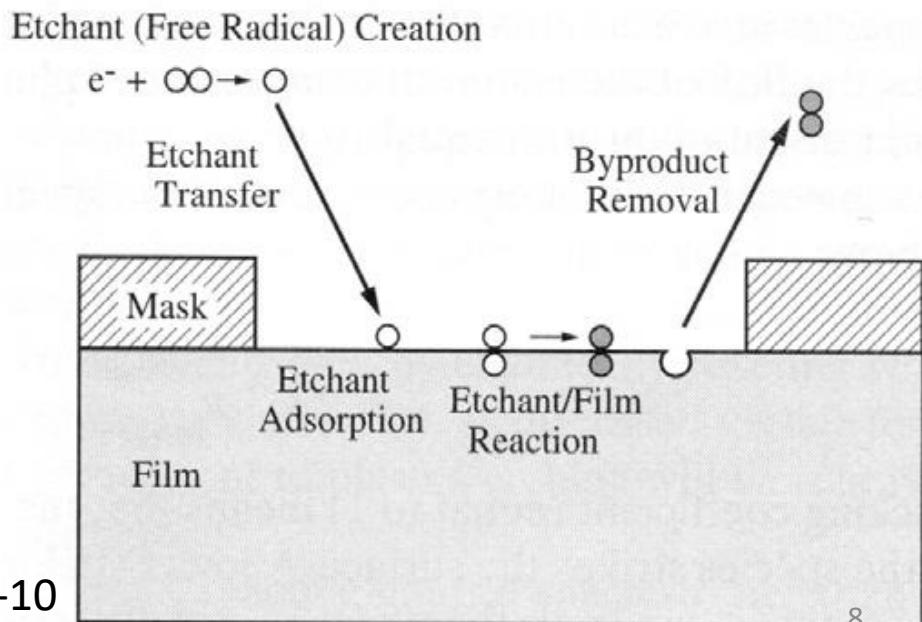
1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Plasma etching mechanism overview

- In a plasma, reactive *neutral* chemical species (free radicals, e.g. F atoms or molecular species CF_3) are mainly responsible for the chemical reaction due to their much greater numbers compared to ions (e.g. CF_3^+ is also reactive, but with low concentration in a plasma).
- Those free radicals and molecules also serve as primary deposition species in PECVD.
- Those free radicals are more abundant than ions because: 1) they are generated at lower threshold energy (e.g. < 8eV; in comparison, Ar is ionized at 15.7eV); and 2) they (*uncharged* radicals) have longer lifetime in the plasma.
- The neutral radicals arrive at cathode surface by *diffusion* (thus non-directional).

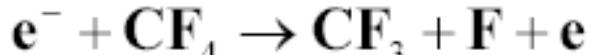
- Charged ions are *accelerated* to the cathode due to self-bias.
- (Unless with very high energy of >100eV as in ion beam/sputter etching), ion itself doesn't contribute significantly to the chemical reaction mostly due to its very low concentration, but ion bombardment can greatly enhance the chemical reaction in ion-enhanced etching.

Figure 10-10

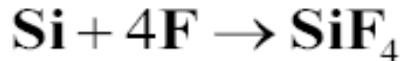


Chemical etch: highly selective, but isotropic

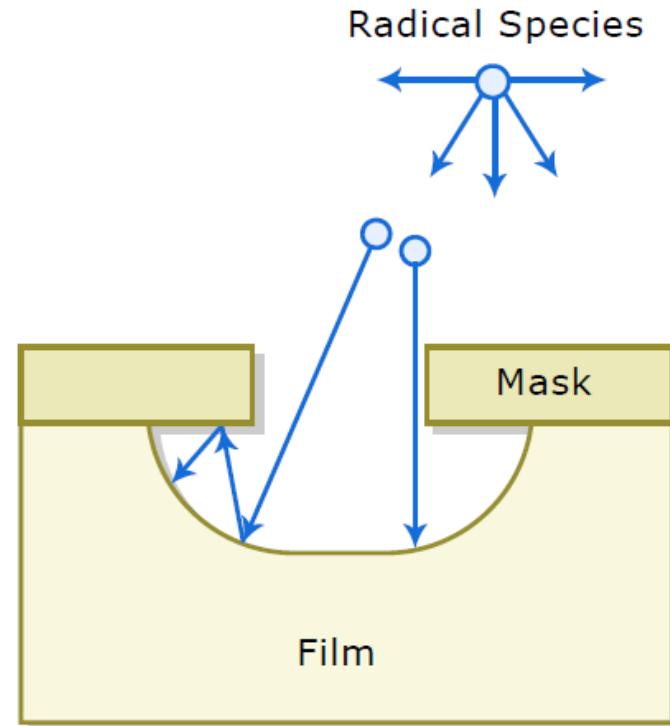
- Due to their incomplete bonding (incomplete outer shells), free radicals (neutral, e.g. CF_3 and F from CF_4 plasma) are highly reactive chemical species.



- Free radicals react with film to be etched and form volatile by-products.



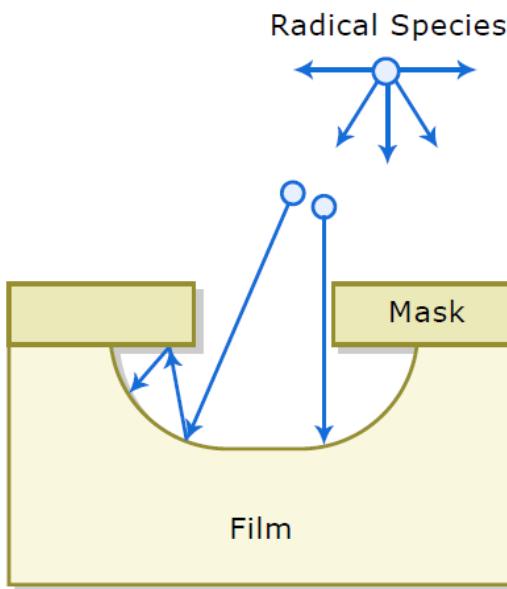
- Pure chemical etch is isotropic or nearly isotropic, and the etching profile depends on arrival angle and sticking coefficients of free radicals.
- Free radicals (un-charged) in plasma systems have isotropic arrival angles.
- The sticking coefficient S is very low, typically only $S \sim 0.01$ (i.e. most free radicals adsorb then just bounce back without reaction).
- This leads to isotropic character of etch, as free radicals can etch area beneath the mask due to bouncing, as seen in the figure. The resulted profile has large undercut.



Adsorption rate onto surface

$$\phi = \alpha \frac{1}{4} n < v > = \frac{\alpha p}{\sqrt{2\pi mkT}}$$

Sticking coefficient S



Most adsorbed species just left the adsorption site without doing anything, so $S \ll 1$.

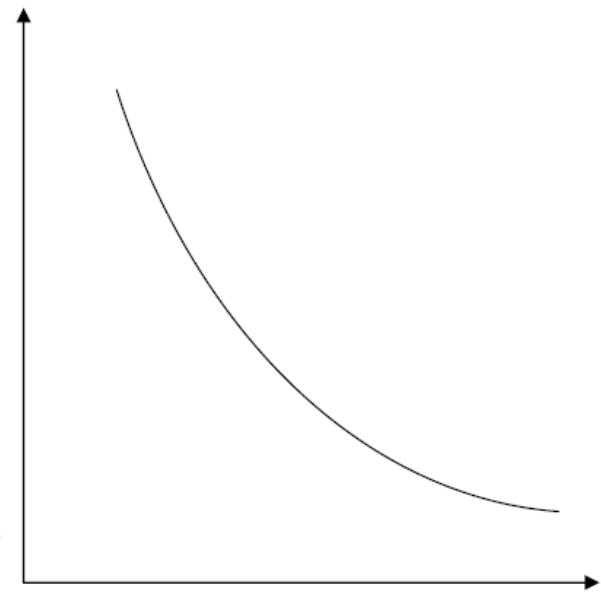
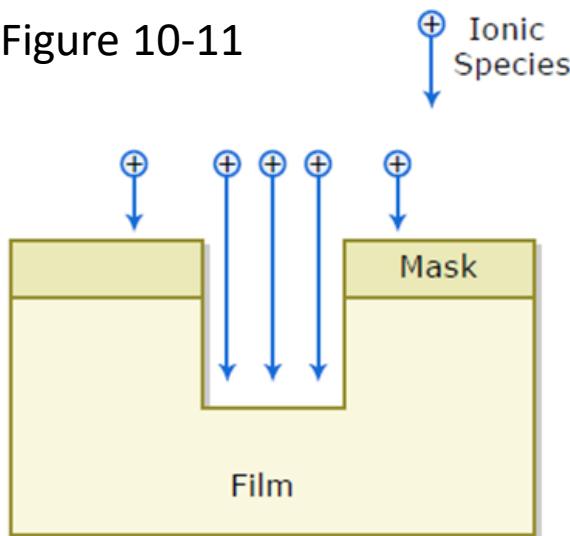


Figure 10-11

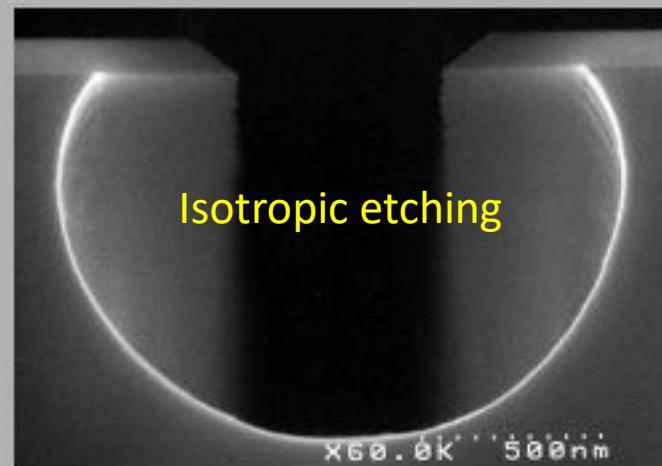
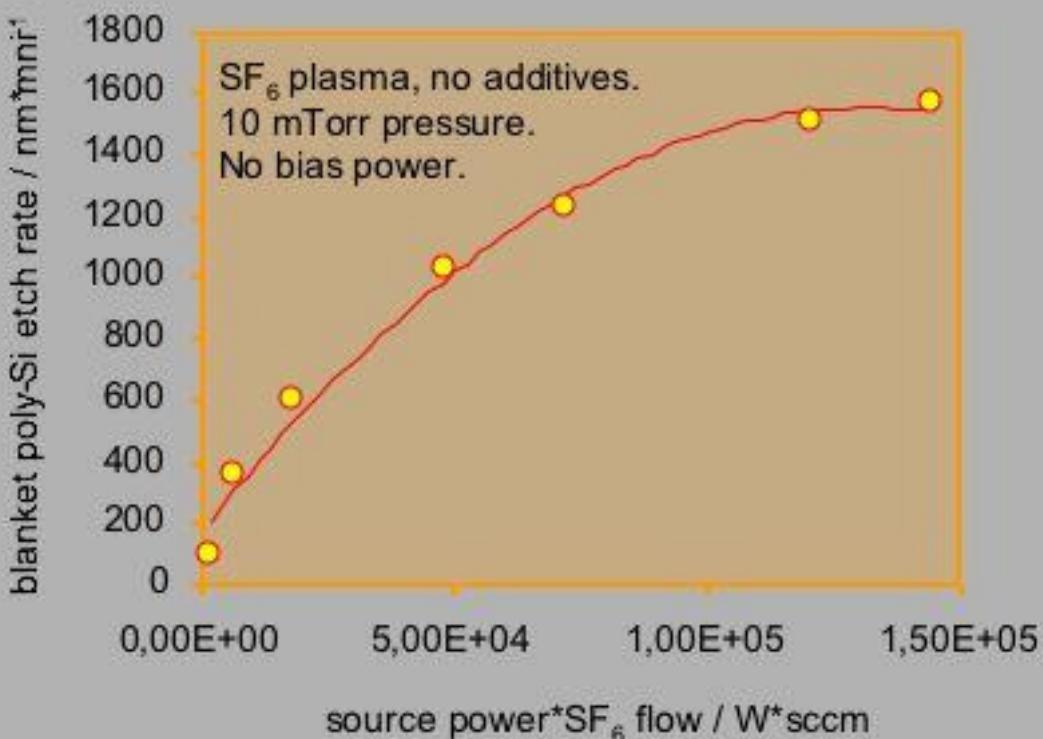


Anisotropic
 $S \approx 1$

"Reaction", here momentum transfer by *physical* bombardment, takes place at every shot, usually sputter off a few atoms, so $S \sim 1$.

Spontaneous Etching Reactions in Plasma Etching

- Silicon Etching with Fluorine Radicals -



Spontaneous or chemical etch processes are typically isotropic because the lack of ion bombardment does not allow the formation of a sidewall passivation.

Silicon etch rate is driven by radical concentration in the gas phase.
Higher source power leads to deeper SF₆ dissociation and more free fluorine.
Higher SF₆ flow reduces the concentration of reaction products in the gas phase.
For very high source power / SF₆ flow combinations, the etch rate saturates, indicating that surface processes are becoming rate limiting.



Etch byproducts should have low boiling point

Low boiling point means very volatile, so it can be pumped away.

This is not necessary for *physical* etching/sputtering, where etch product is sputtered off that ideally doesn't fall on the other part of the wafer (re-deposition).

Boiling points of typical etch products

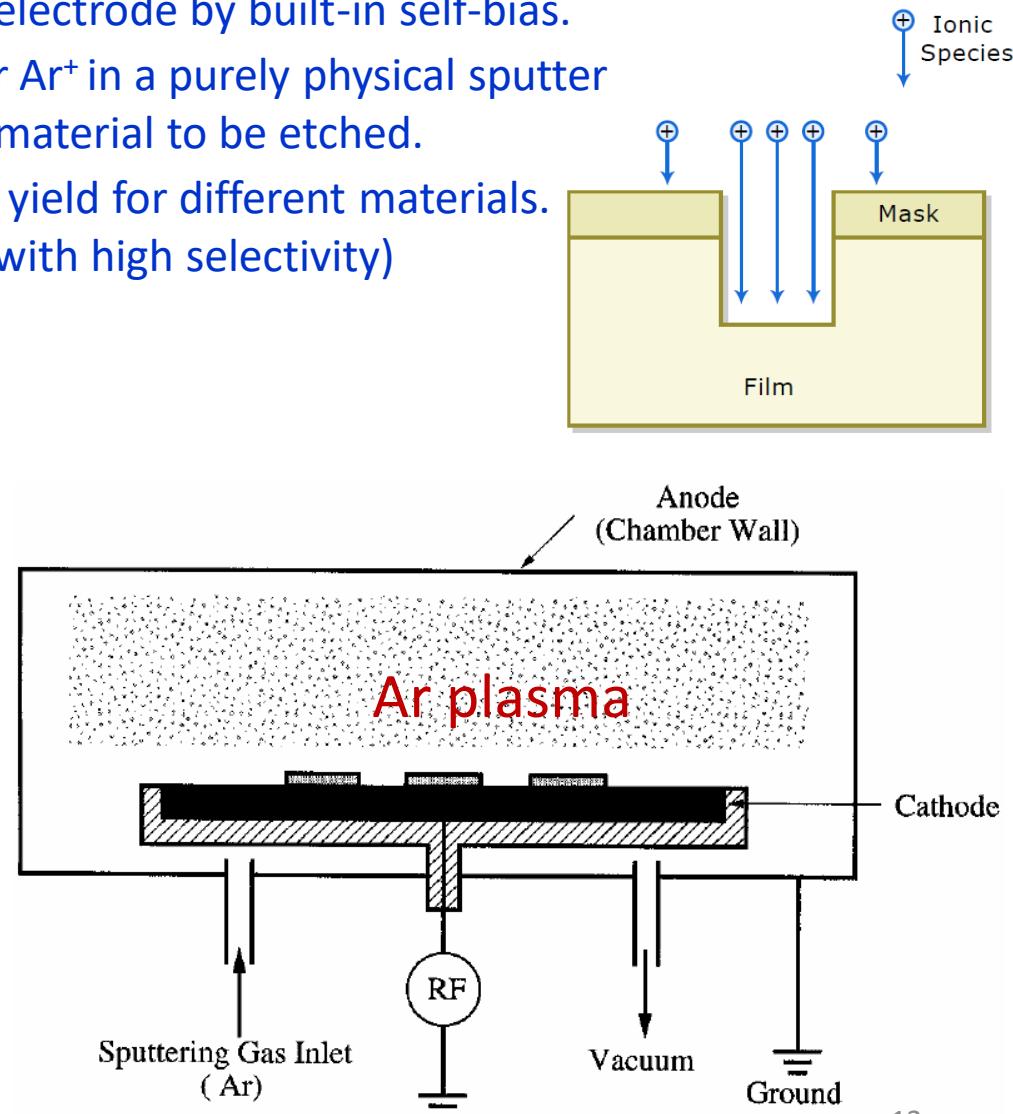
ELEMENT	CHLORIDES	BOILING POINT (°C)	FLUORIDES	BOILING POINT (°C)
Al	AlCl ₃	177.8 (subl.)	AlF ₃	1291 (subl.)
CU	CuCl	1490	CuF	1100 (subl.)
Si	SiCl ₄	57.6	SiF ₄	-86
Ti	TiCl ₃	136.4	TiF ₄	284 (subl.)
W	WCl ₆	347	WF ₆	17.5
	WCl ₅	276	WOF ₄	187.5
	WOCl ₄	227.5		

Physical etch component in a plasma etch system (much less important than chemical etch)

- Ionic species are accelerated toward each electrode by built-in self-bias.
- The ionic species such as Cl_2^+ , CF_4^+ , CF_3^+ (or Ar^+ in a purely physical sputter etch) strike wafer surface and remove the material to be etched.
- Directional, non-selective - similar sputter yield for different materials.
(But CF_3^+ can also etch Si *chemically*, then with high selectivity)
- It may result in significant re-deposition.

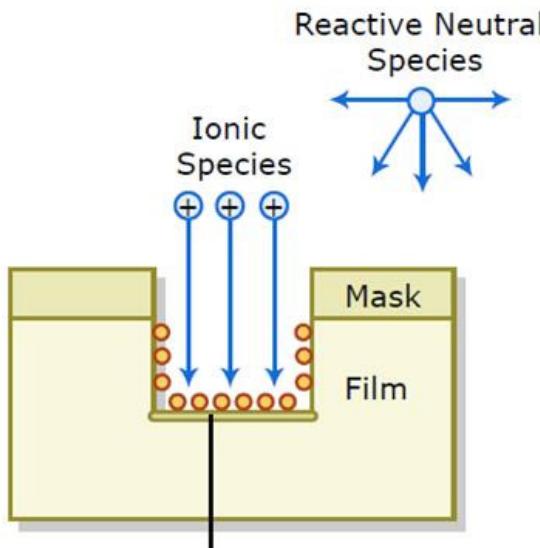
Pure physical etch: sputter etching system

- Self-bias few 100V, but low ion energy (order 10V) due to collision energy loss.
- Thus **very low milling rate** in a sputter system, often for surface cleaning only.
- Here is the case for sputter etching system with gas pressure order 10mTorr.
- For a dedicated ion milling system (no plasma, see later slides), the pressure is $\sim 10^{-4}$ Torr or even lower (cannot sustain a plasma), leading to large mean free path, high ion energy and high milling rate.

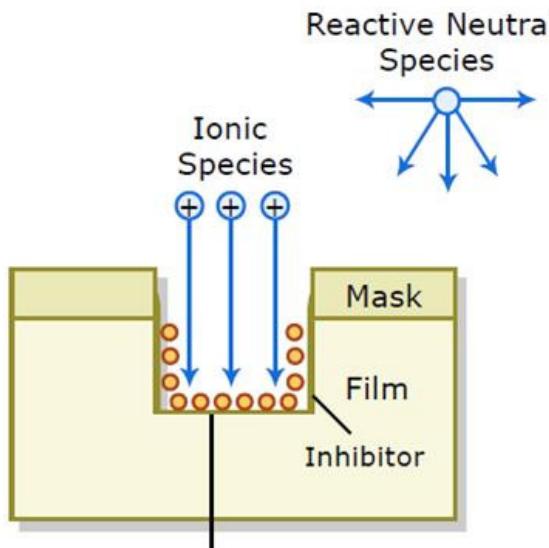


Ion enhanced etching (IEE): chemical etch assisted by physical bombardment

- IEE is an anisotropic (due to directional ion bombardment) and highly selective (due to chemical reaction) etching process.
- Reactive ion etch (**RIE**) is the most popular form of IEE.
- Ion bombardment can enhance one of the following steps during chemical etch: surface adsorption, **etching reaction** (by physically damaging/weakening the chemical bond of the material to be etched), **by-product (inhibitor layer) removal**, and removal of un-reacted etchants.



Chemical etch enhanced
by ion bombardment



Inhibitor removed by
ion bombardment

Inhibitor layer: e.g. fluorocarbon polymer formed from CHF_3 during etching of SiO_2 .

When removal rate \ll deposition rate, net deposition will occur, then the process becomes similar to PECVD!!

Indeed, the RIE and PECVD are pretty similar tools, except PECVD is typically heated.

Figure 10-13

IEE: first proof of etching mechanism

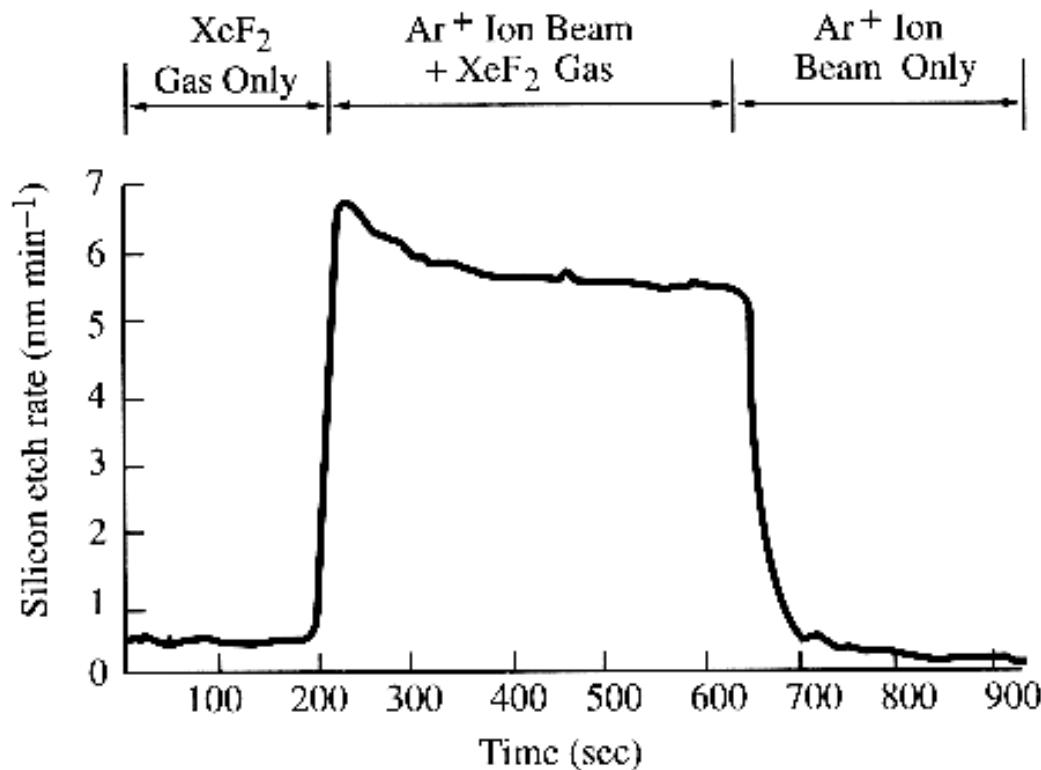


Figure 10-12 Etch rate of silicon as XeF₂ gas (not plasma) and Ar⁺ ions are introduced to the silicon surface. Only when both are present does appreciable etching occur, illustrating synergistic etching behavior between chemical and ionic species. (After [10.5, 10.6].)

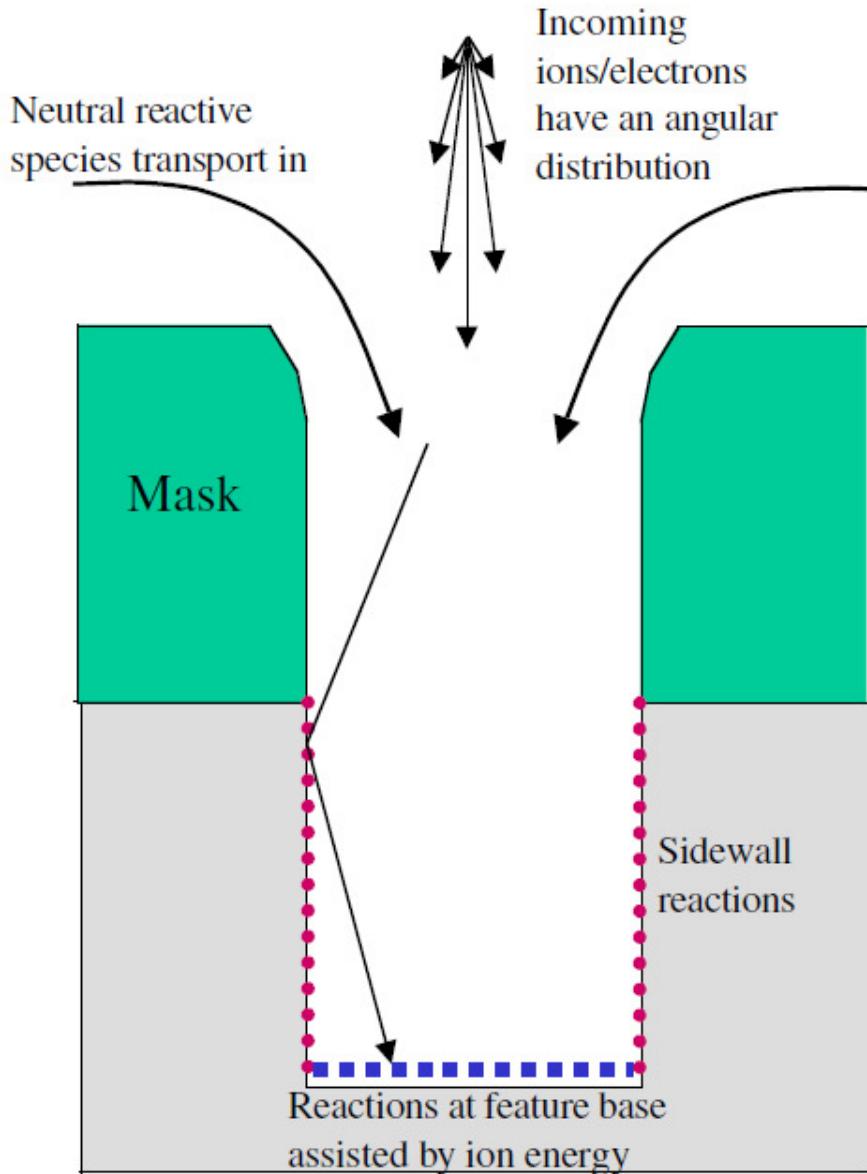
Gas phase etch, with or without the aid of Ar ion beam.
NO plasma.
Very slow etch when pure chemical or physical etch *alone*

The ion enhancement could be due to the damage/weakening of silicon lattice by ion bombardment, which makes the etching by XeF₂ easier.

The resulted profile will be anisotropic since the horizontal surfaces are much more bombarded than vertical ones.

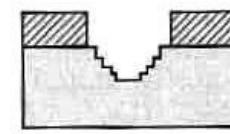
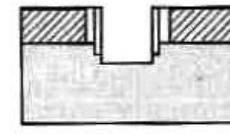
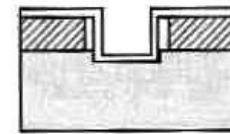
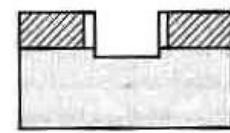
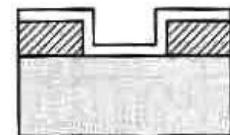
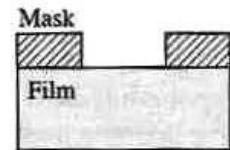
This is one example of CAIBE (chemically assisted ion beam etching), see later slides

Ion enhanced etching is highly anisotropic



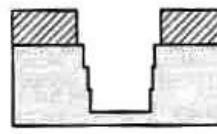
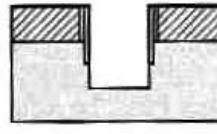
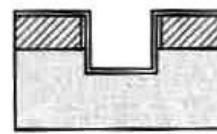
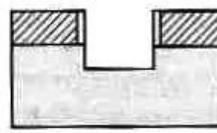
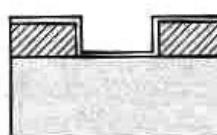
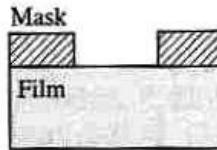
- Sidewall reactions can lead to an isotropic etch component.
- To prevent sidewall etching, one can build up a passivation (inhibitor – inhibit chemical reaction) layer that protects it.
- Then there is a competition between passivating and etching reaction.
- For the feature base/horizontal surfaces, etch rates tend to be temperature independent because of ion energy input (i.e. inhibitor sputtered away by ions).
- On sidewall, substrate temperature can play an important role as sidewall passivation depends on the volatility of the inhibitor that is controlled by temperature (cryo-etcher at below -100°C is available recently, then the sidewall passivation layer is not volatile).
- Even without sidewall passivation, lower temperature still increases anisotropy since chemical attack of sidewall is suppressed at low temperature. (Attack of horizontal surfaces are assisted by ion bombardment)

High inhibitor deposition rate



a. Inhibitor deposition rate fast compared to etch rate

Low inhibitor deposition rate

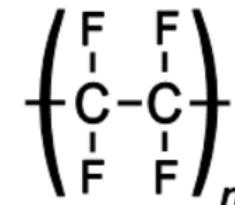


b. Inhibitor deposition rate relatively slow compared to etch rate

Example:

etching profile of Si or SiO_2

Teflon



- Fluoropolymer (like Teflon) in CHF_3 or $\text{CF}_4 + \text{H}_2$ RIE of Si or SiO_2 is the inhibitor.
- If Ar gas is added, inhibitor is mainly removed by ion bombardment. Thus less attack of inhibitor on sidewall.
- If O_2 gas is added, inhibitor on sidewall is removed at faster rate than Ar ion, but the etch of inhibitor at horizontal surface is even faster.
- Yet at very low temperature, inhibitor $\text{SiO}_{x,y}$ (not act as inhibitor at higher temperature when it is volatile) forms when O_2 is added, which is the mechanism for fast anisotropic etching of Si using cryo-etcher. (deep Si etcher, popular for MEMS – micro electro mechanical systems)

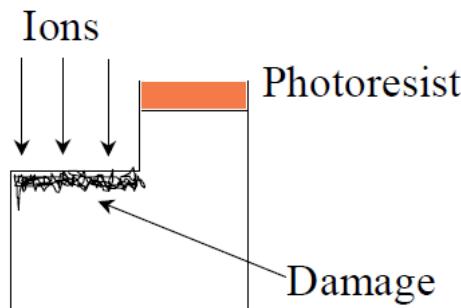
Figure 10-14

Anisotropy due to ion bombardment: summary

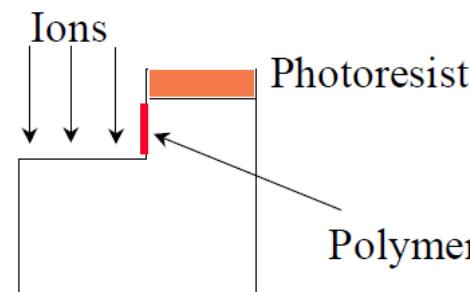
- Due to its extremely low density, ions don't contribute much to etching; neutral radicals do.
- So even with directional ion bombardment, the overall etching can still be pretty isotropic.
- For instance, SF_6 etch of Si is very isotropic with large undercut like wet etch.
- To achieve anisotropy, there are two mechanisms:
 - Energy-driven anisotropy: bombardment by ion disrupts an un-reactive substrate and causes damages such as dangling bonds and dislocations, resulting in a substrate more reactive towards etchant species (electron or photon can also induce surface activation).
 - Inhibitor-driven anisotropy: ion bombardment removes the inhibitor layer from horizontal surface (sidewall remain passivated), and reaction with neutrals proceed on these un-passivated surfaces only.

One may think that ions won't help much due to its much lower density than radicals. But ion has sticking coefficient $S \sim 1$ (every ion bombardment counts), whereas radicals $S \sim 0.01$ (most radicals hit the surface and left without doing anything).

Energy-driven anisotropy



speeds chemical reaction
on horizontal surfaces.



slows chemical reaction
on vertical surfaces.

Inhibitor-driven anisotropy

Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Plasma etching in barrel etchers

Barrel etcher:

- Chemical etching only, isotropic and selective like pure wet etch.
- Use plasma shield to keep ion bombardment from wafers, thus very little damage.
- Poor uniformity edge to center.
- Used in non-critical steps such as photoresist removal by O₂ plasma (Barrel “asher”
Polymer + O → CO₂ + H₂O).

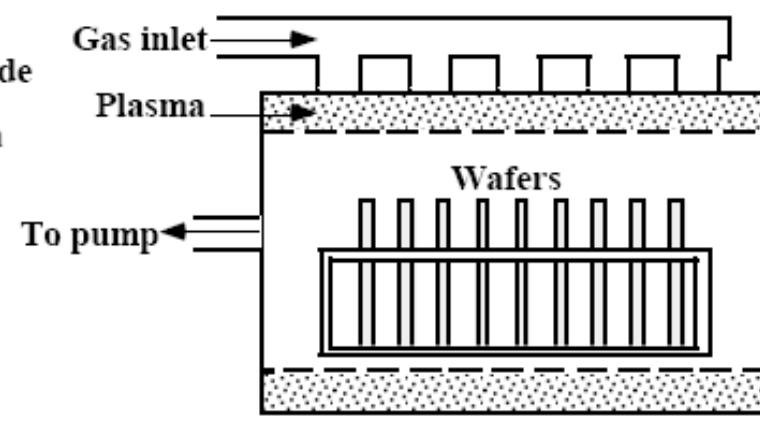
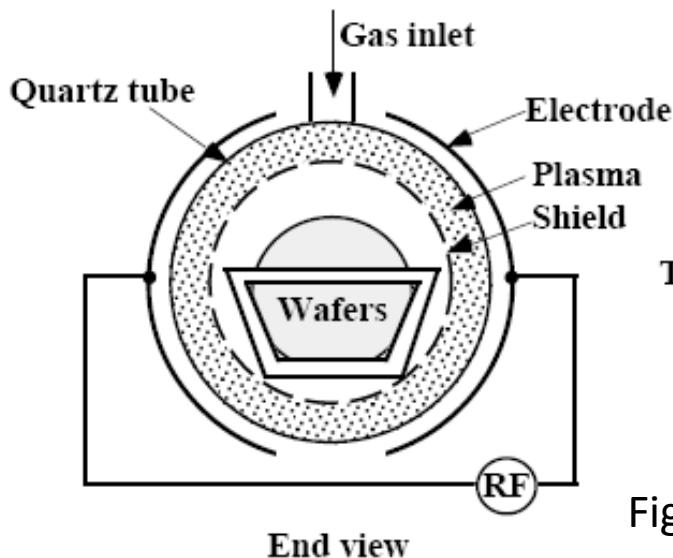
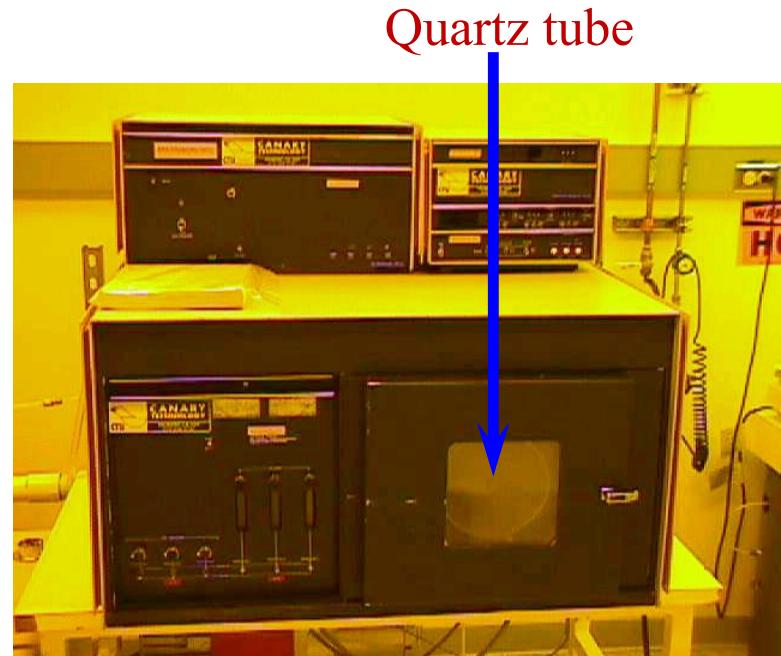
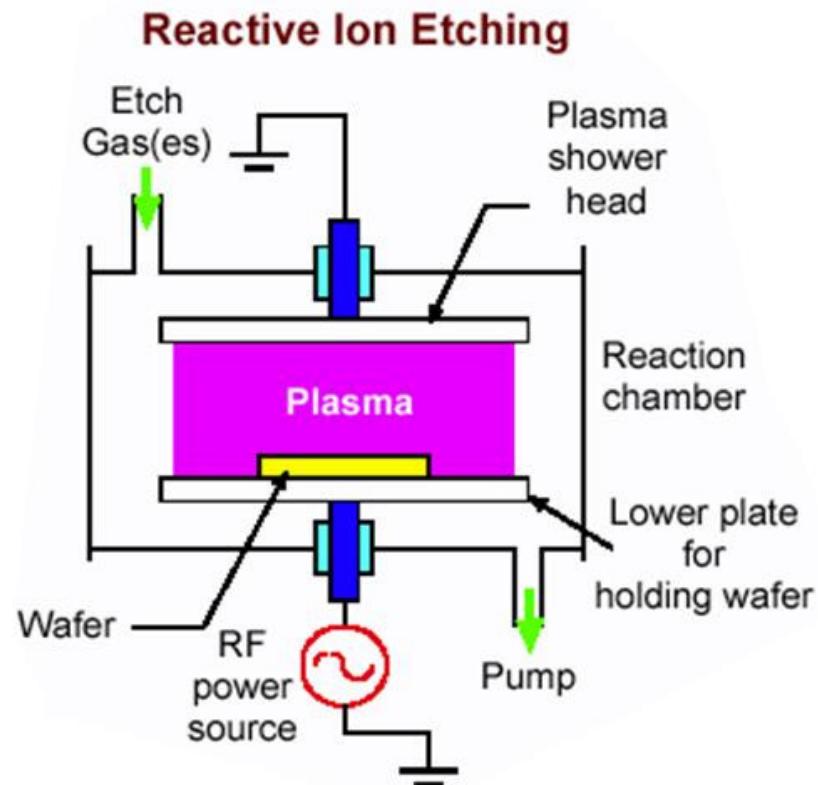


Figure 10-15

Side view

Parallel plate etchers (regular RIE, low density plasma)

- Absolutely the most important form of dry etching, though recently ICP (see later slides) is becoming more and more popular.
- Smaller wafer electrode (*counter electrode* grounded to chamber wall), lower pressure (<100mTorr), more physical bombardment (voltage drop many 100V).
- Ion enhanced etching mechanism, (usually) directional/anisotropic and selective.



RIE using parallel plate setup is low density plasma system (ions 10^8 – $10^{10}/\text{cm}^3$), thus low etch rate.

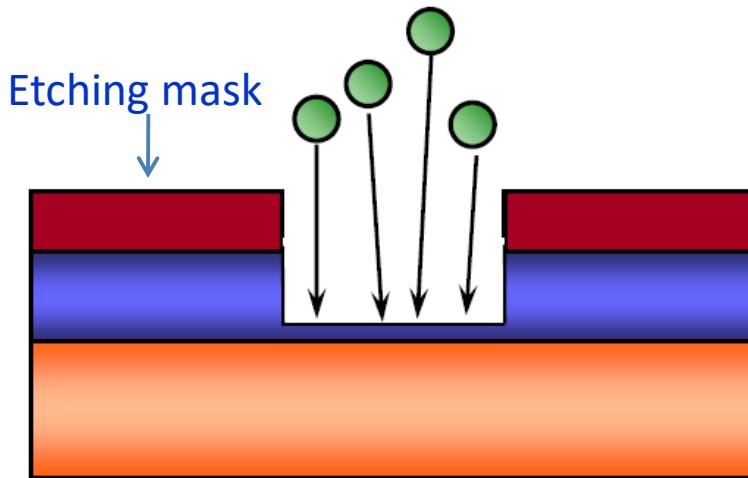
Here low (ion) density plasma also implies low density of free radicals. Thus low etching rate.

At a pressure of 20 mT the plasma consist of

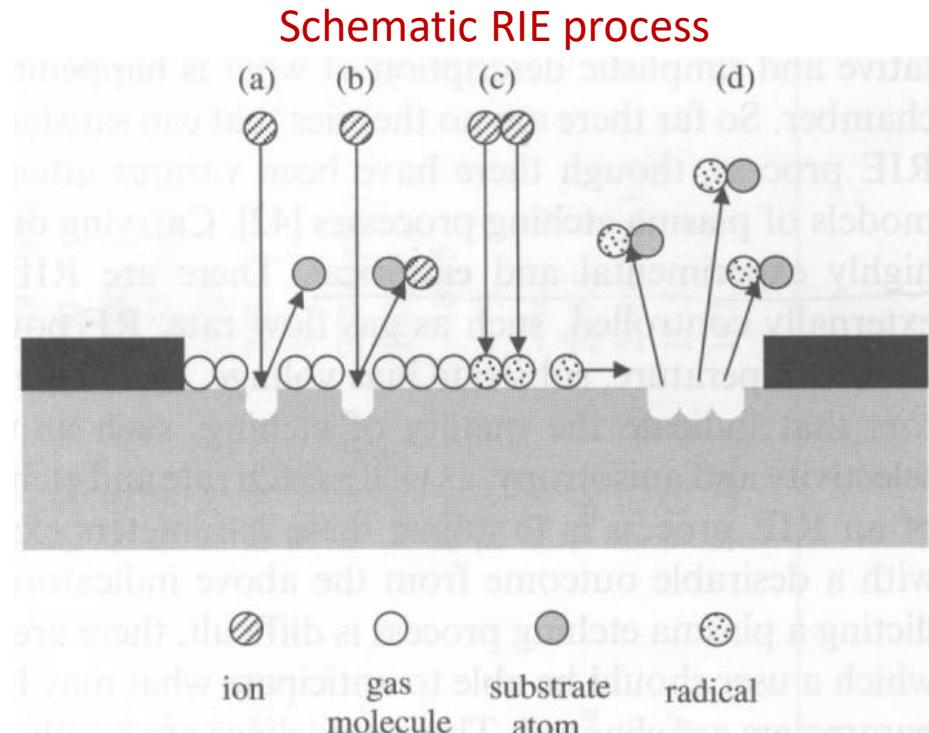
Neutral Species	10^{15} cm^{-3}
Reactive Neutral Species	$10^{12} - 10^{13} \text{ cm}^{-3}$
Ions and Electrons	$10^8 - 10^{12} \text{ cm}^{-3}$

VERY roughly, one can say that plasma consists of order 1% radicals (reactive neutral species) and 0.01% ions²¹

Reactive ion etch (RIE)



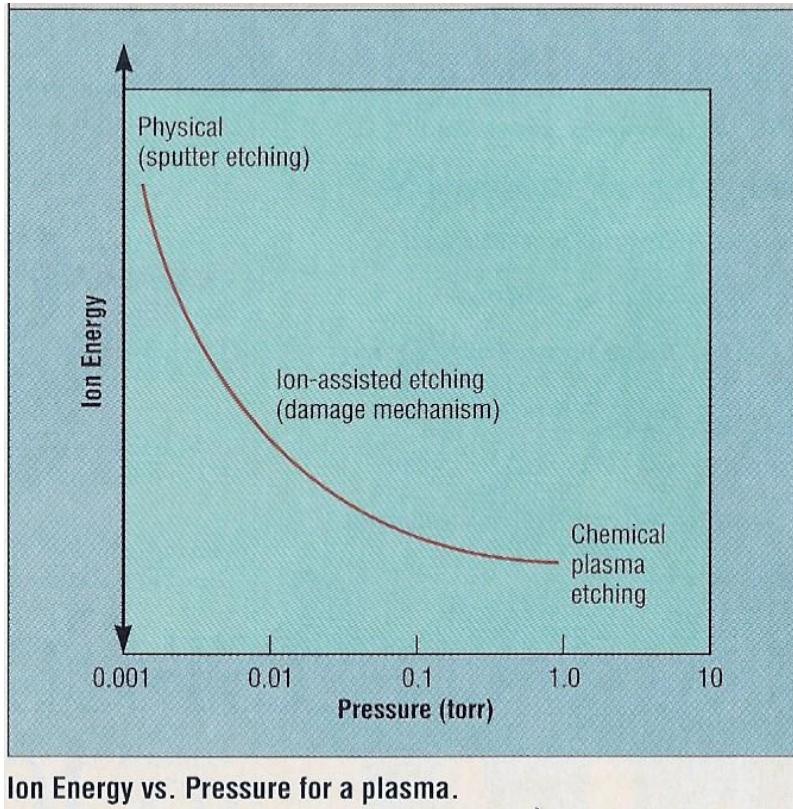
- Due to its simultaneous anisotropy and selectivity, RIE is intensively used.
- Works for most semiconductors and dielectrics.
- OK for few metals that form volatile etch products: Al (form AlCl_3), Ti (TiF_4 , TiCl_4), W (WF_6), Cr (CrO_2Cl_2).



- a) Ion sputtering, b) reactive ion etching, c) radical formation, d) radical etching (most important)

In RIE, ion energy is low (several 10s eV, << voltage drop near wafer surface, due to collision energy loss), and its number density is very low, thus negligible etching by ion bombardment. The name reactive “ion” etching is very misleading since ions don’t contribute *directly* to etching – it just “helps” chemical etching.

Ion energy vs. pressure for a plasma



- Lower pressure (<10mTorr) increases mean free path as well as voltage drop near wafer electrode, both of which leads to more energetic and directional ion bombardment, thus more anisotropic, but less selective and slower etching rate due to low ion/free radicals density.
- High pressure (>100mTorr), short mean free path, low voltage drop, more isotropic chemical etching.

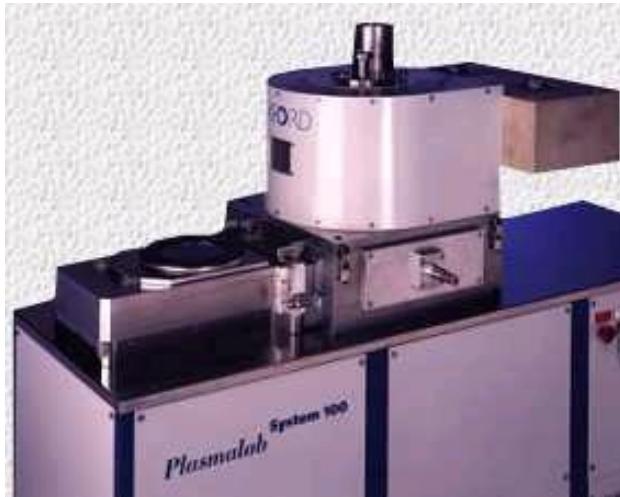
RIE mode: 10-100mTorr

Sputter etching: pressure as low as possible, as long as plasma can be sustained, but still very slow etching rate.

Etching in high density plasma (HDP) systems

- Ion flux and ion bombarding energy can be independently controlled. For regular RIE, they are tightly coupled (e.g. higher power increases both).
- High plasma (ion) density ($> 10^{11}$) enhances etch rate.
- Since ionization is much more efficient, can operate at lower pressure, which leads to less ion collision, so more directional/anisotropic, thus enhances profile control.
- As ion energy is independently controlled, it can be kept low if desirable.
- Then the extent and amount of damage will be reduced, without sacrificing etching rate that is still high for high density plasma.
- Currently HDP represents an optimum compromise in high etch rates, good selectivity, good directionality, while low ion energy and damage.

Electron cyclotron resonance (ECR) and inductively coupled plasma (ICP)



ECR was introduced in 1985.

ICP was introduced much later (1991- 1995).

Dual plasma source:

Top one (ECR or ICP RF power) generates HDP, determines ion density/current.

Bottom one (CCP RF power) generates bias voltage like regular RIE, determines ion energy.

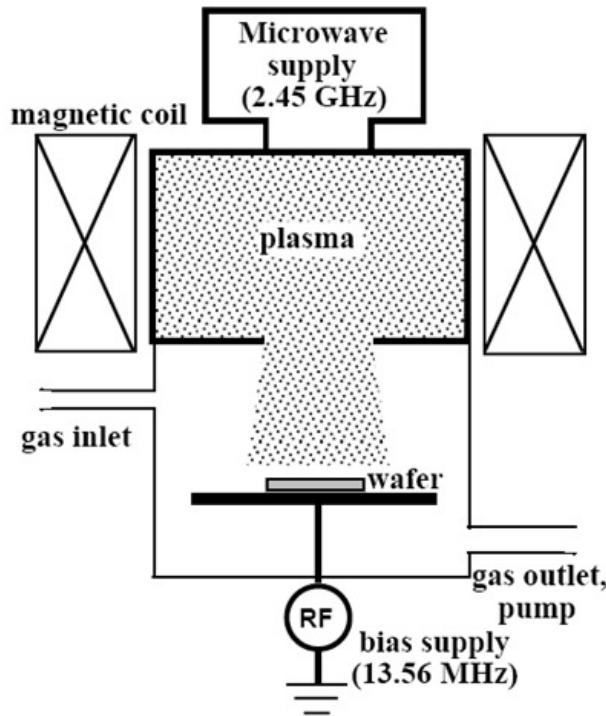
Typical parameters for HDP and conventional plasma etcher

Parameters	RF Etcher	HDP Etcher
Pressure (mT)	50~1000	0.5~50
Plasma Density (cm^{-3}) ions/ cm^3	$10^9\sim10^{11}$	$10^{10}\sim10^{12}$
Ion Energy (eV) should be lower	200~1000	Controllable 20~500
Ion Current Density (mA/cm^2)	0.1~2	1~20
Electron Temperature (eV)	1~5	2~7
Ionization rate	$10^{-6}\sim10^{-3}$	$10^{-4}\sim10^{-1}$

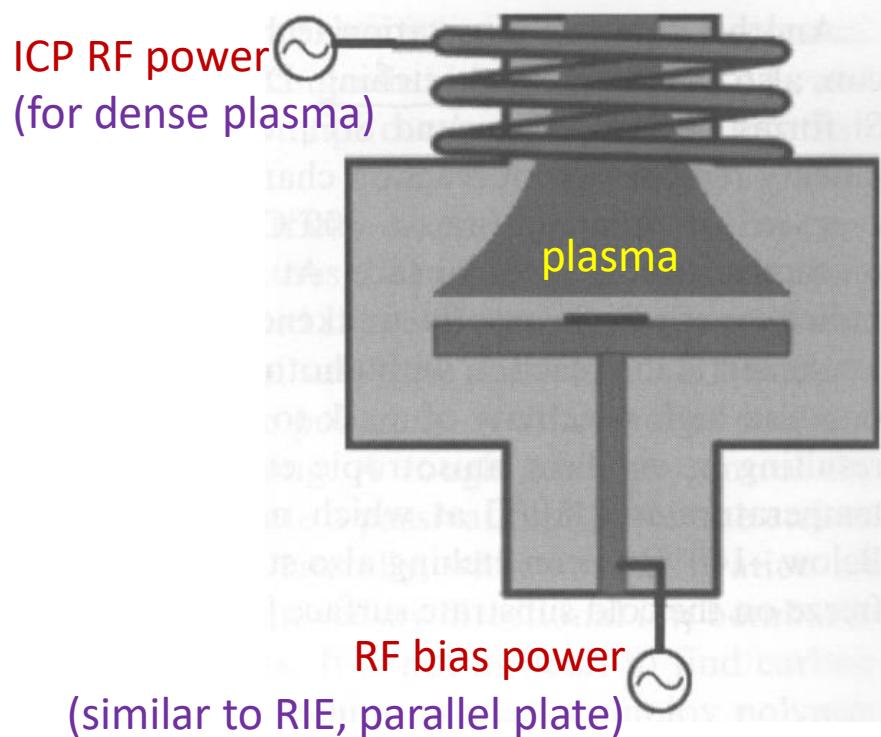
CCP: capacitively coupled plasma, parallel plate, used for conventional regular RIE.

ECR and ICP

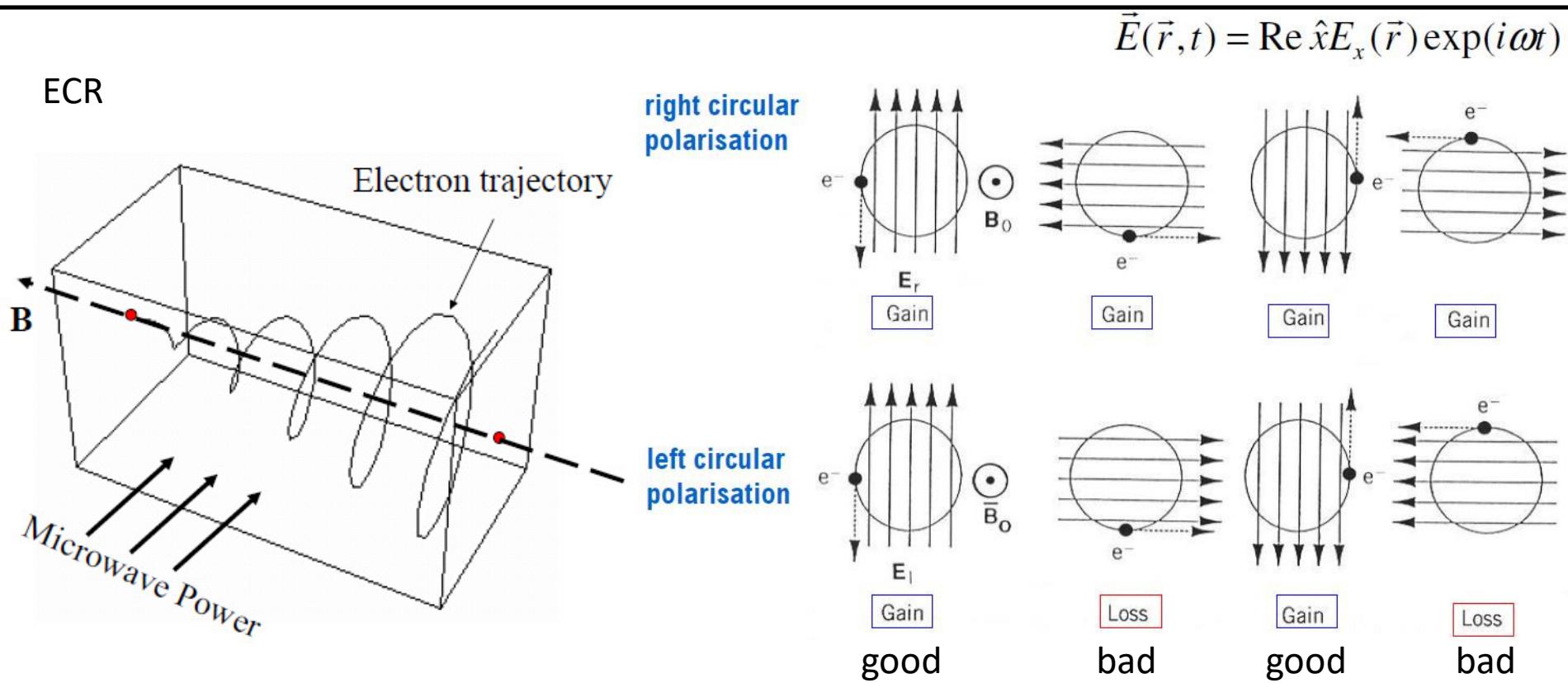
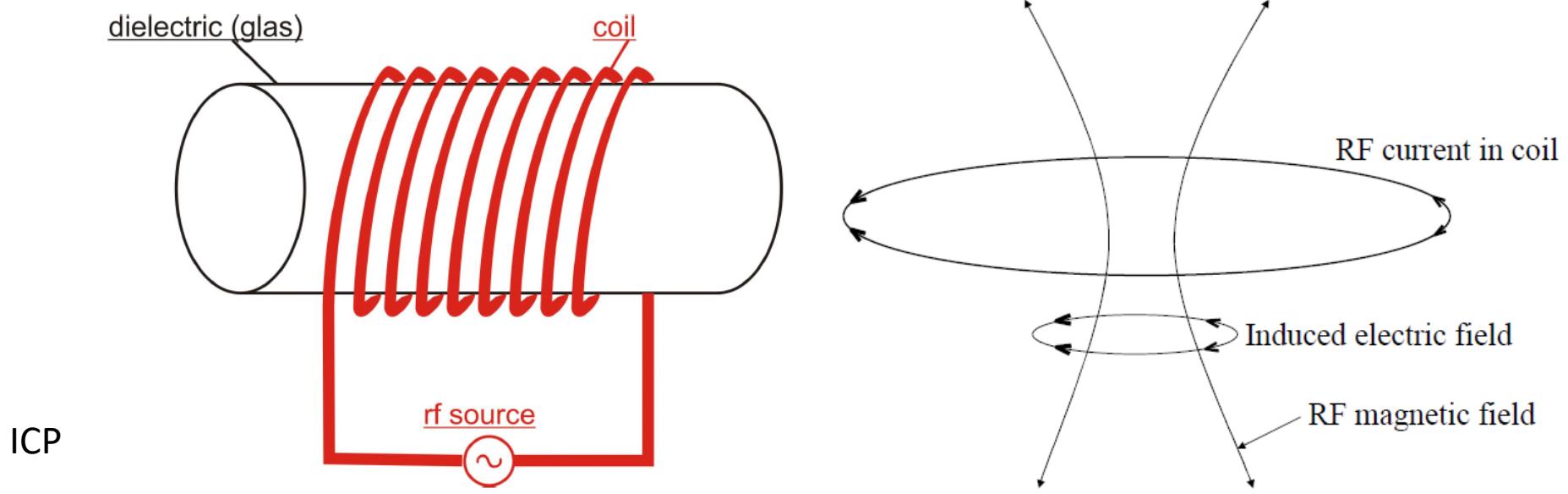
Electron cyclotron resonance plasma (less common nowadays)



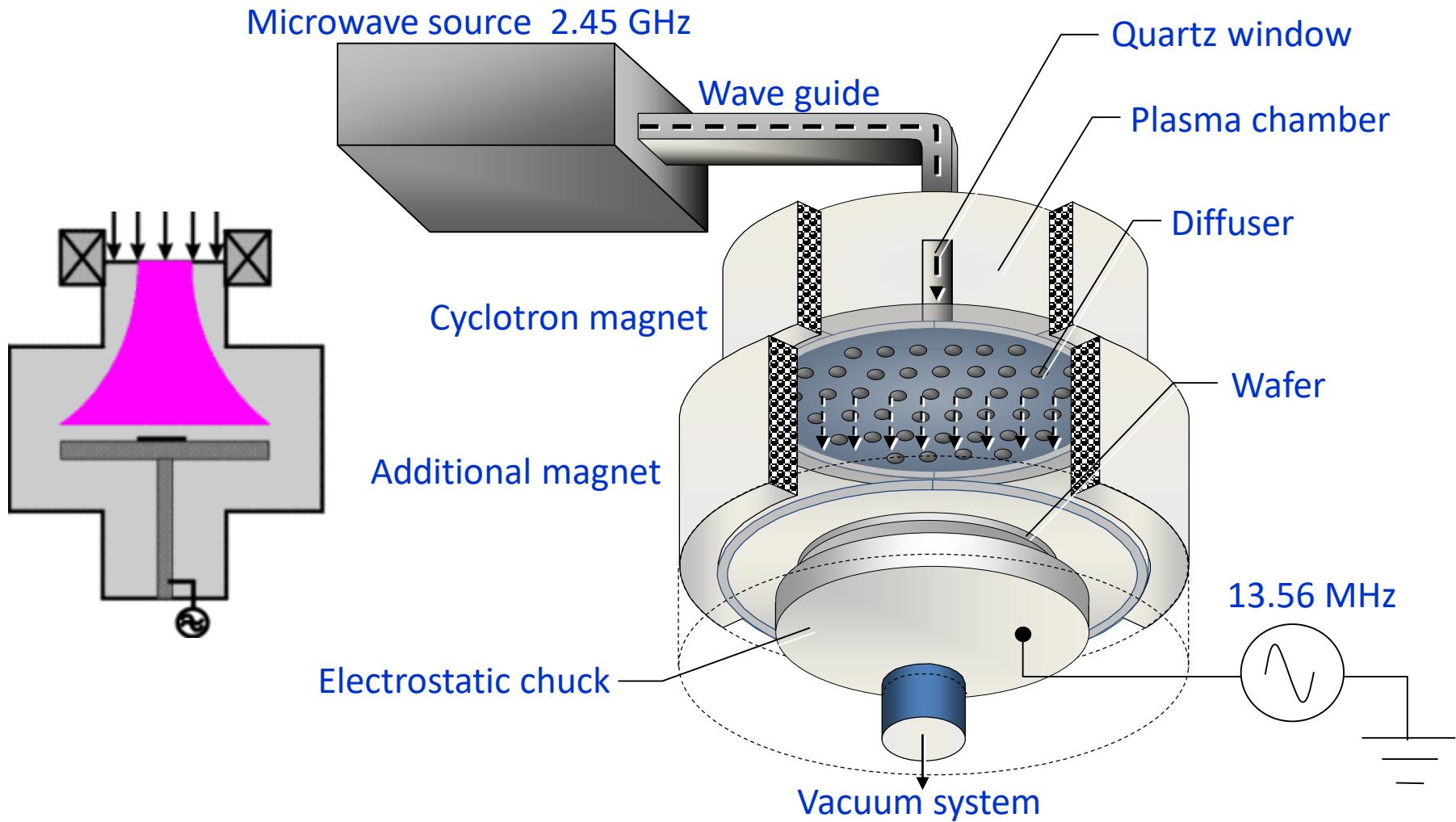
Inductively coupled plasma (ICP) (four systems at Waterloo)



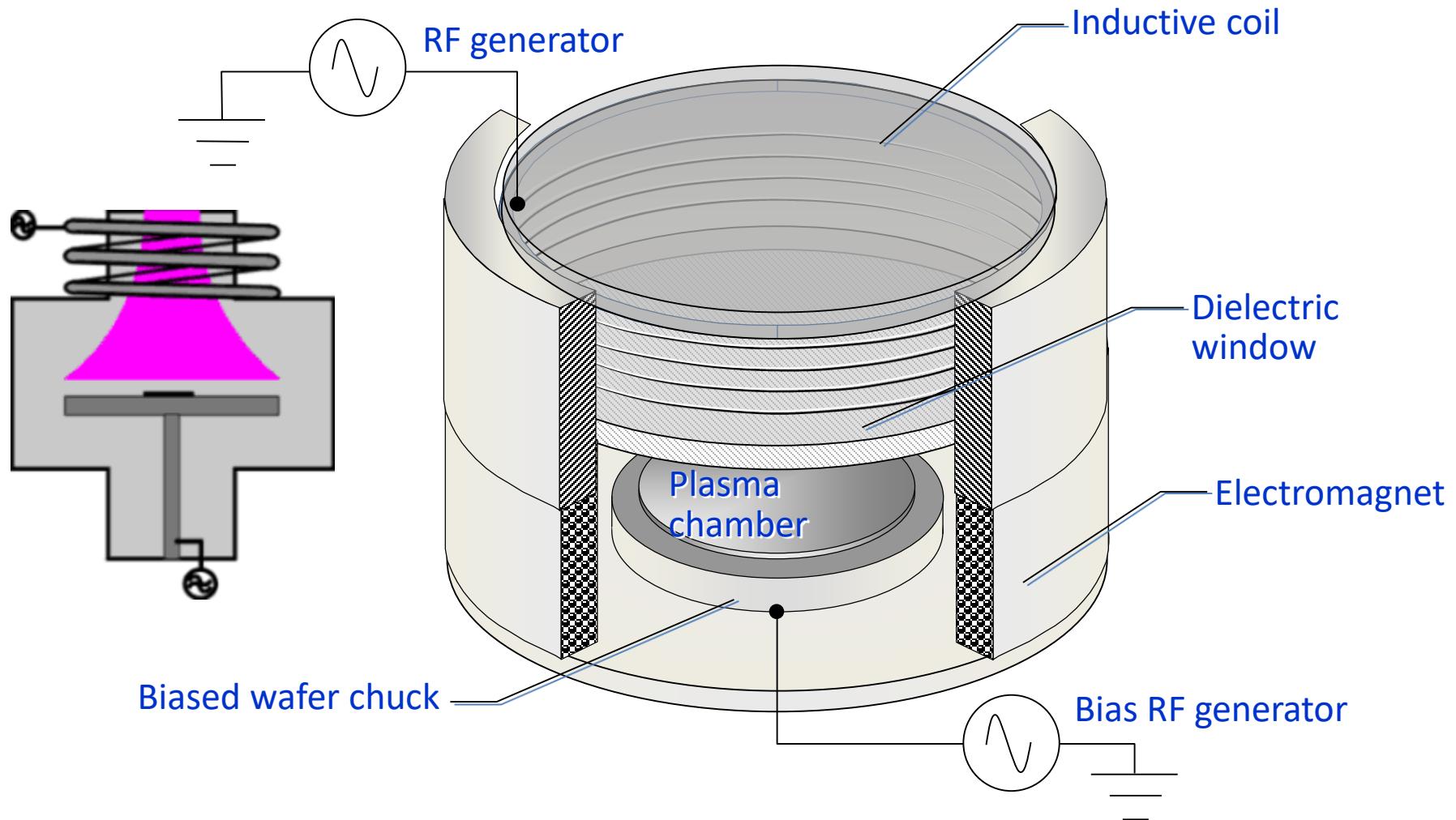
- High magnetic field in the coil, so electrons move in circles with long path, leading to higher collision and ionization probability, and much less electron loss to chamber wall and the bottom plate where sit the wafer. Moreover,
- For ICP, **AC** magnetic field induces circular electrical field, which accelerates electrons.
- For ECR, **DC** magnetic field, electron cyclotron $\omega = qB/m$; electrons accelerated if this frequency matches the microwave frequency.



Schematic of ECR etcher



Schematic of ICP etcher

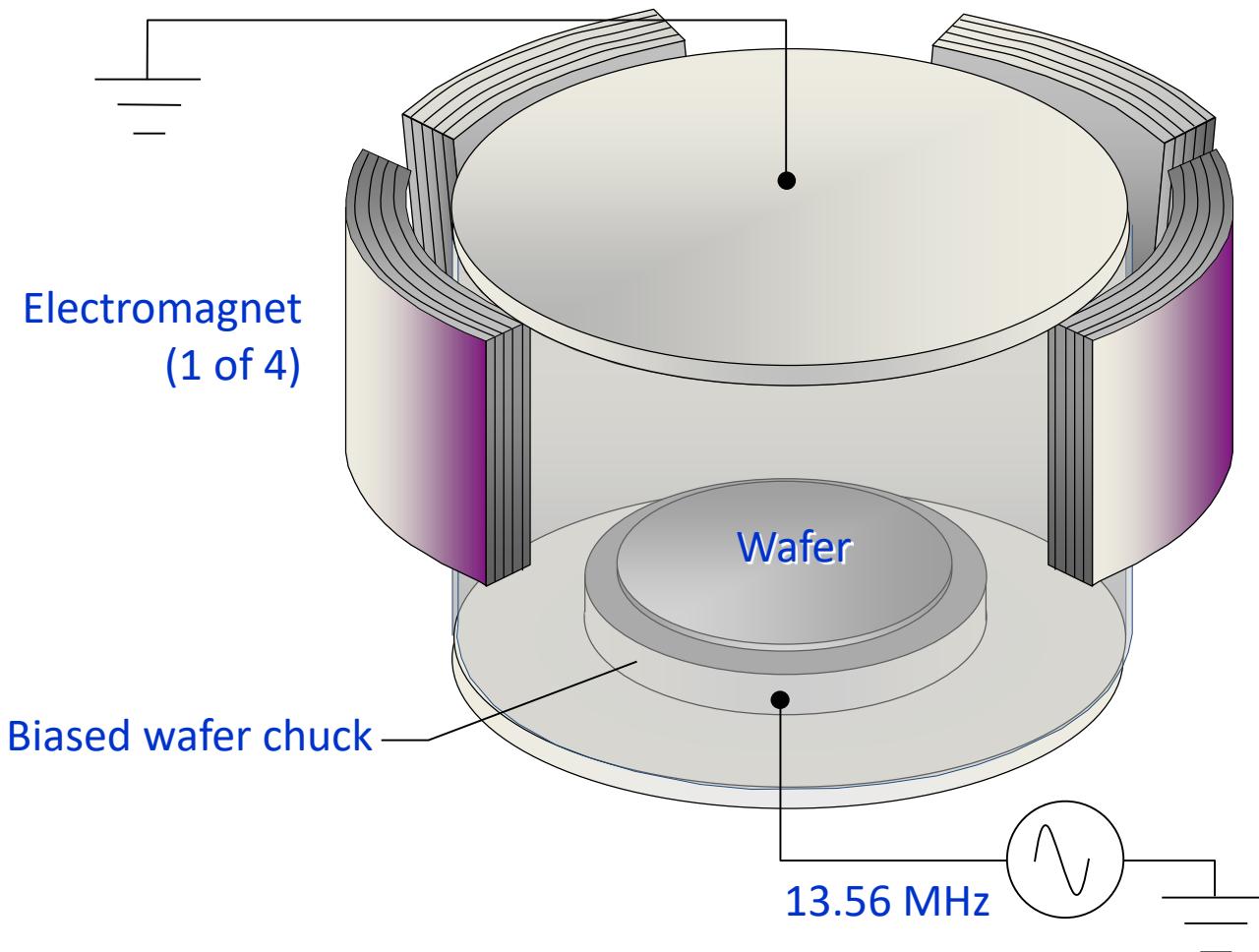


As you see, there is practically no top plate as in parallel plate regular RIE.
The wafer sees the ICP power – the two power sources are not physically separated.

Magnetically enhanced reactive ion etch (MERIE)

Like regular parallel plate RIE, but magnetic field forces electron to go circles, increasing collision with gas molecules and decreasing loss to chamber walls or top/bottom plates.

However, now that electrons don't loss to bottom plate, no or little bias voltage – need to apply an external bias to accelerate ions.



I haven't seen any MERIE, so it is not so popular.

On the contrary, magnetron sputtering is very popular.

This is probably because there are many ways to increase etching rate; but sputter without magnetron is always very slow:

few nm/min, vs. 10s to 100s nm/min RIE etching rate.

Sputter etching and ion milling

Sputter etching: (etch *inside plasma*)

- The etch mechanism is purely physical.
- Very similar in principle to sputter deposition, but now the target becomes substrate to etch.
- Poor selectivity (few:1), very anisotropic.
- Sputtering rate depends on sputter yields which can be a function of incident angle.
- Problems include faceting (sputter yield is a function of incident angle), trenching, re-deposition, charging and ion path distortion, radiation damage.
- Not popular, etches too slow, though reactive gas (CF_4 , CCl_4 , O_2) can be added to slightly improve selectivity and etching rate.

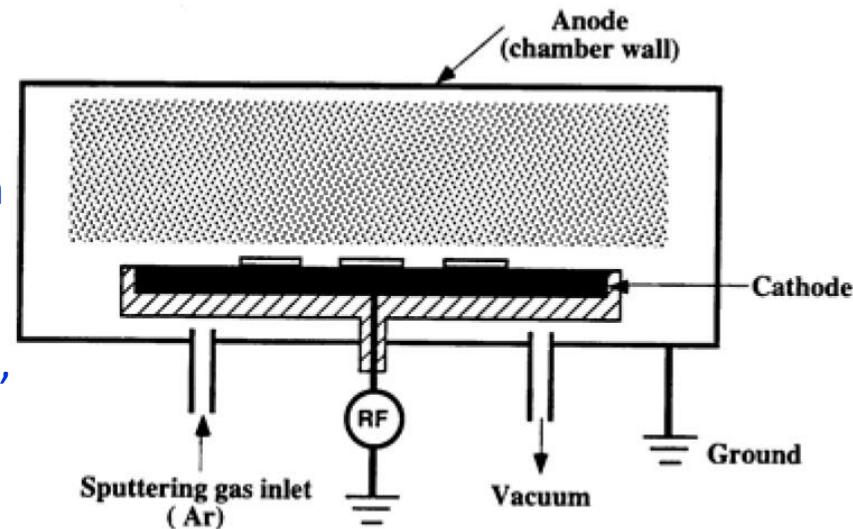
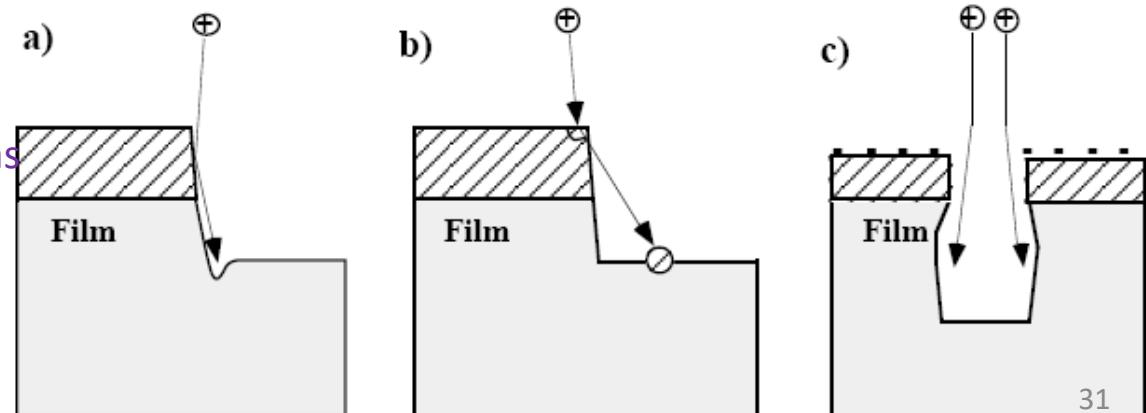


Figure 10-8 Problems associated with sputter etching (or any etching that has a high degree of physical/ionic etching): a) trenching at bottom of sidewalls; b) redeposition of photoresist and other materials; c) charging and ion path distortion.



Ion milling or ion beam etching (IBE)

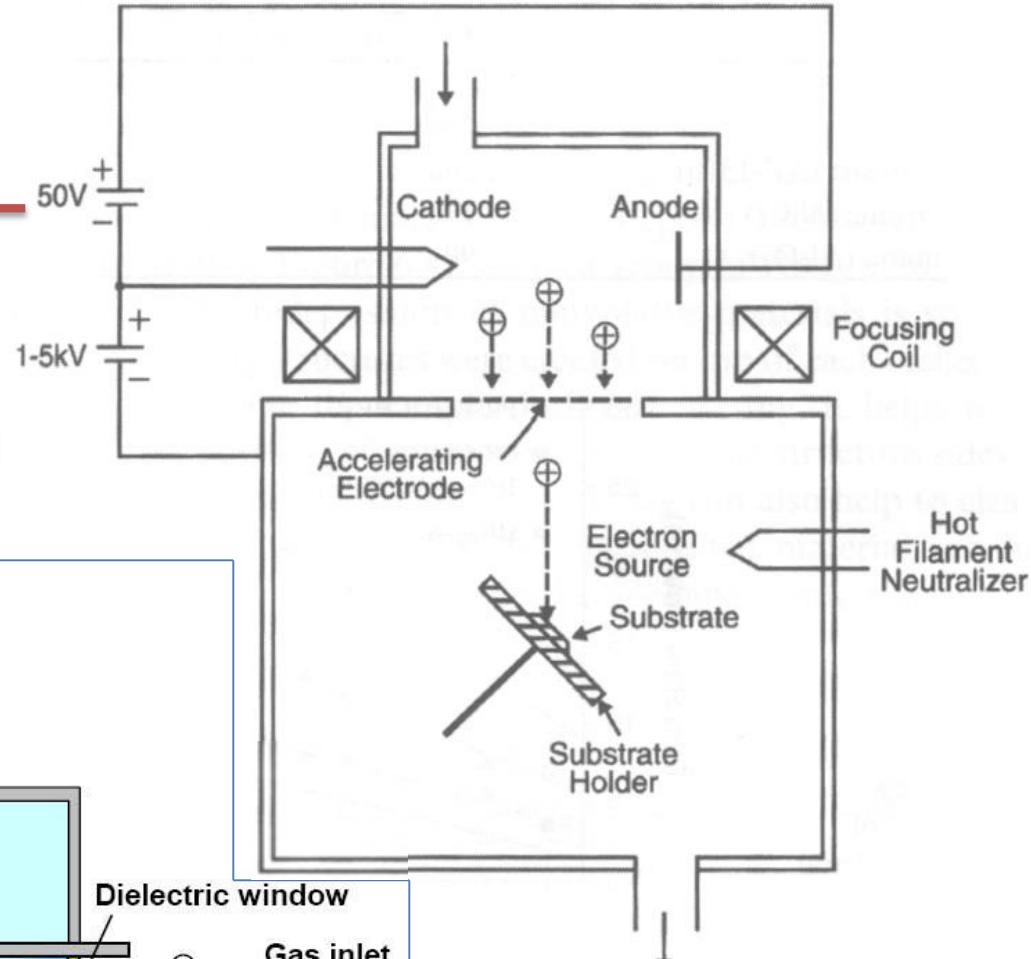
Used to call ion milling, seems now more called as ion beam etching.

- Physical milling when using heavy inert gases (Ar).
- Plasma is used to generate ion beam (Ar^+), which is extracted and accelerated to etch the sample. (i.e. sample *outside* of plasma)
- Thus the ion density (determined by plasma source) and ion energy (determined by DC acceleration voltage – bias by applied DC voltage, not by RF bias as in high density plasma etching system), can be controlled independently.
- Low pressure $\sim 10^{-4}$ Torr (>1 order lower than RIE), so large mean free path and less energy loss due to collision. (such low pressure cannot sustain a plasma, so ion milling is not plasma etching)
- High acceleration voltage ($>1\text{kV}$), leading to mill rate $\sim 10\text{-}30\text{nm/min}$.
- Despite the high voltage and low pressure, such a rate is still $<$ typical RIE rate where chemical etching dominates.
- Used whenever RIE is not possible (due to the lack of volatile species formation). Usually employed to etch Cu, Ni, Au, superconducting materials containing metals...

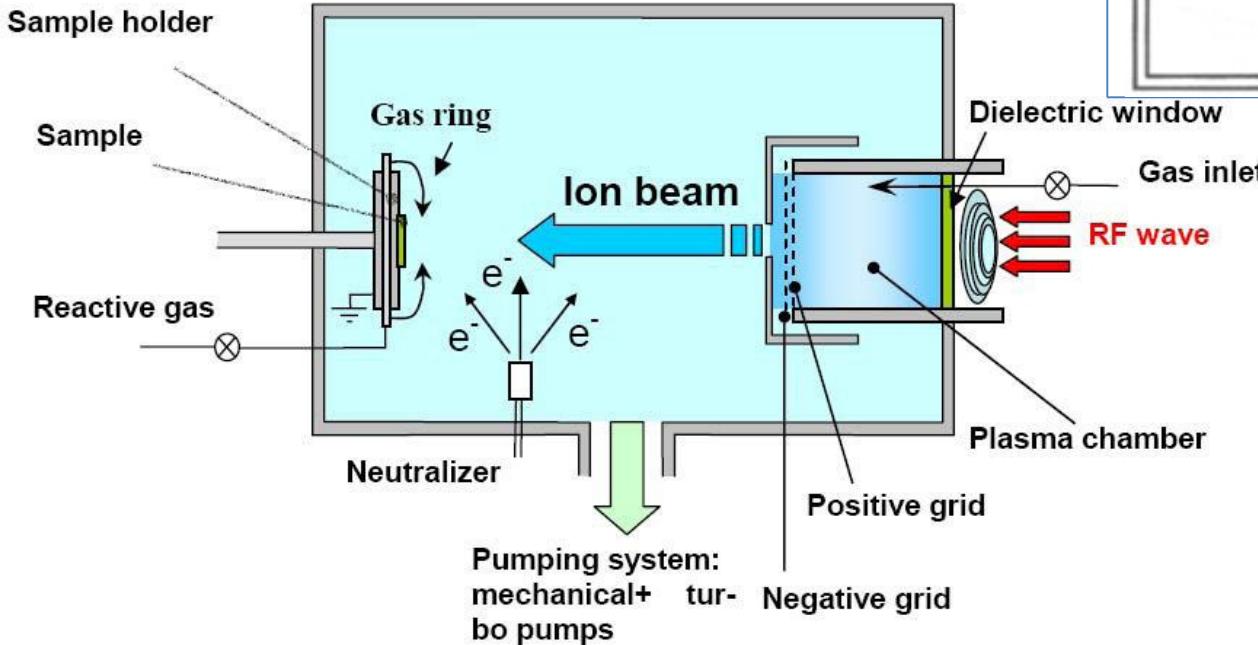
Ion beam etching system: triode configuration

Electron beam is first generated by hot filament.

Ions are generated by electron bombardment, then accelerated to bombard the substrate.



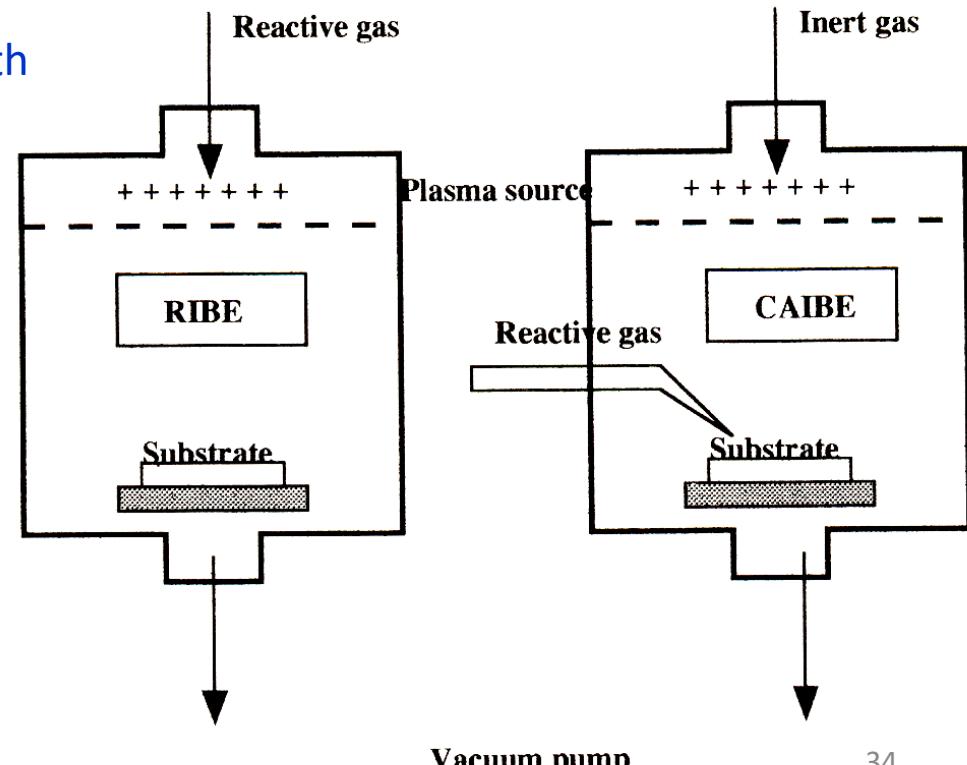
RF plasma ion beam source
(here reactive gas added, so it is actually a CAIBE, see next slide)



DC ion beam source
Electrons sprayed to sample to neutralize ions.
Tilted sample to greatly increase milling rate.
But then shadowing may become a problem when milling high aspect ratio structures.

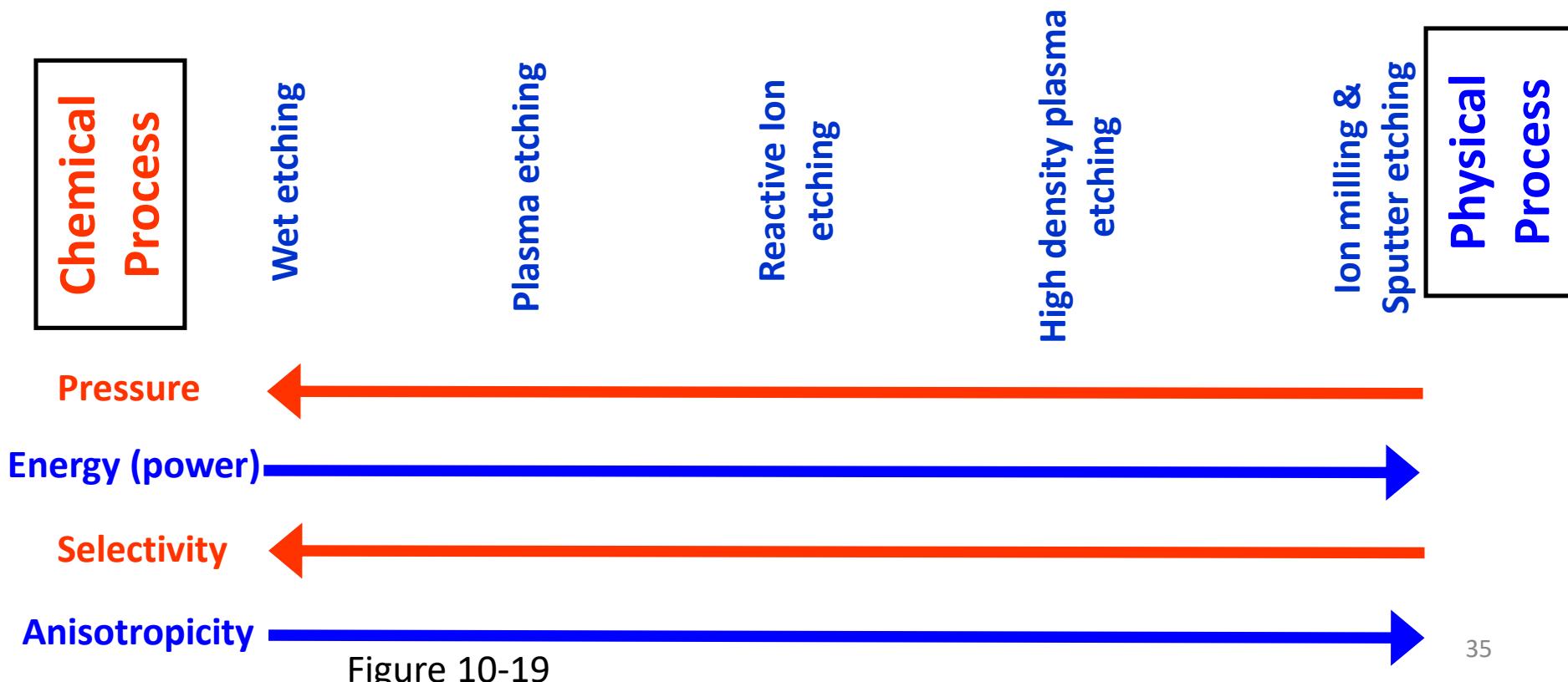
Chemically assisted ion beam etching system

- Adding reactive gases (CF_4 , CCl_4 , O_2 , Cl_2) to increase etching rate and selectivity.
 - Usually physical etching still dominates, no need of volatile etch product.
 - Here it is chemically assisted physical etching, different from RIE that is a kind of physically assisted chemical etching.
-
- **RIBE:** reactive IBE, reactive gases are introduced into plasma region together with Ar gas, so they are ionized. Here the same ion has both a physical (ion impact) and chemical (reactive etching) component.
 - **CAIBE:** chemically assisted ion beam etching, inert Ar ion, neutral reactive gas is introduced into lower chamber, so it is not ionized. A better name could be ion-beam-assisted chemical etching, if chemical etching dominates physical etching.

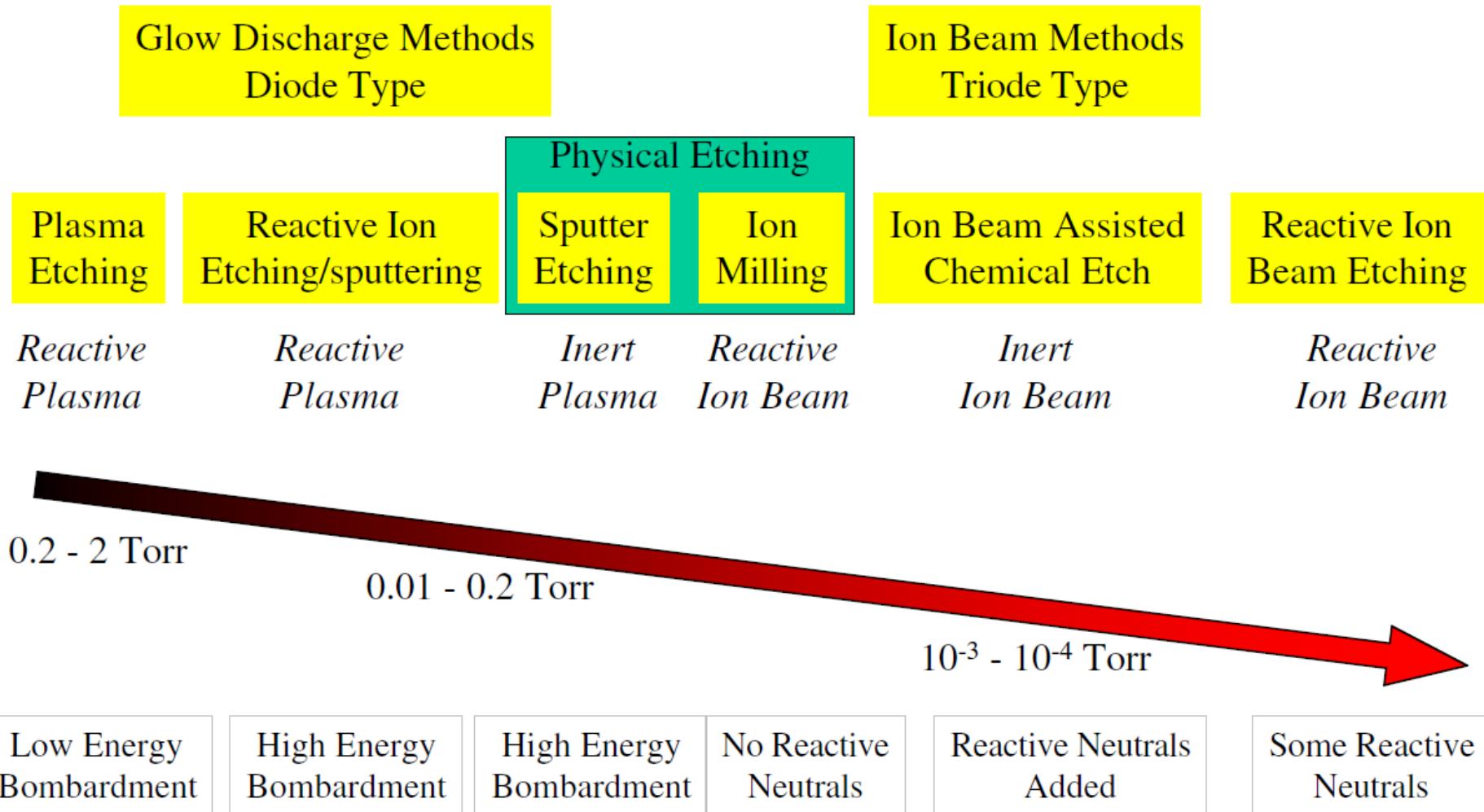


Summary: plasma etching mechanism

- Chemical etching: free radicals react with material to be removed.
- Physical etching or sputtering: ionic species, accelerated by the built-in electric field (self-bias), bombard the materials to be removed.
- Ion enhanced etching: combined chemical and physical process, higher material removal rate than each process alone. E.g. reactive ion etching (RIE), which is the most widely used dry etching technique.



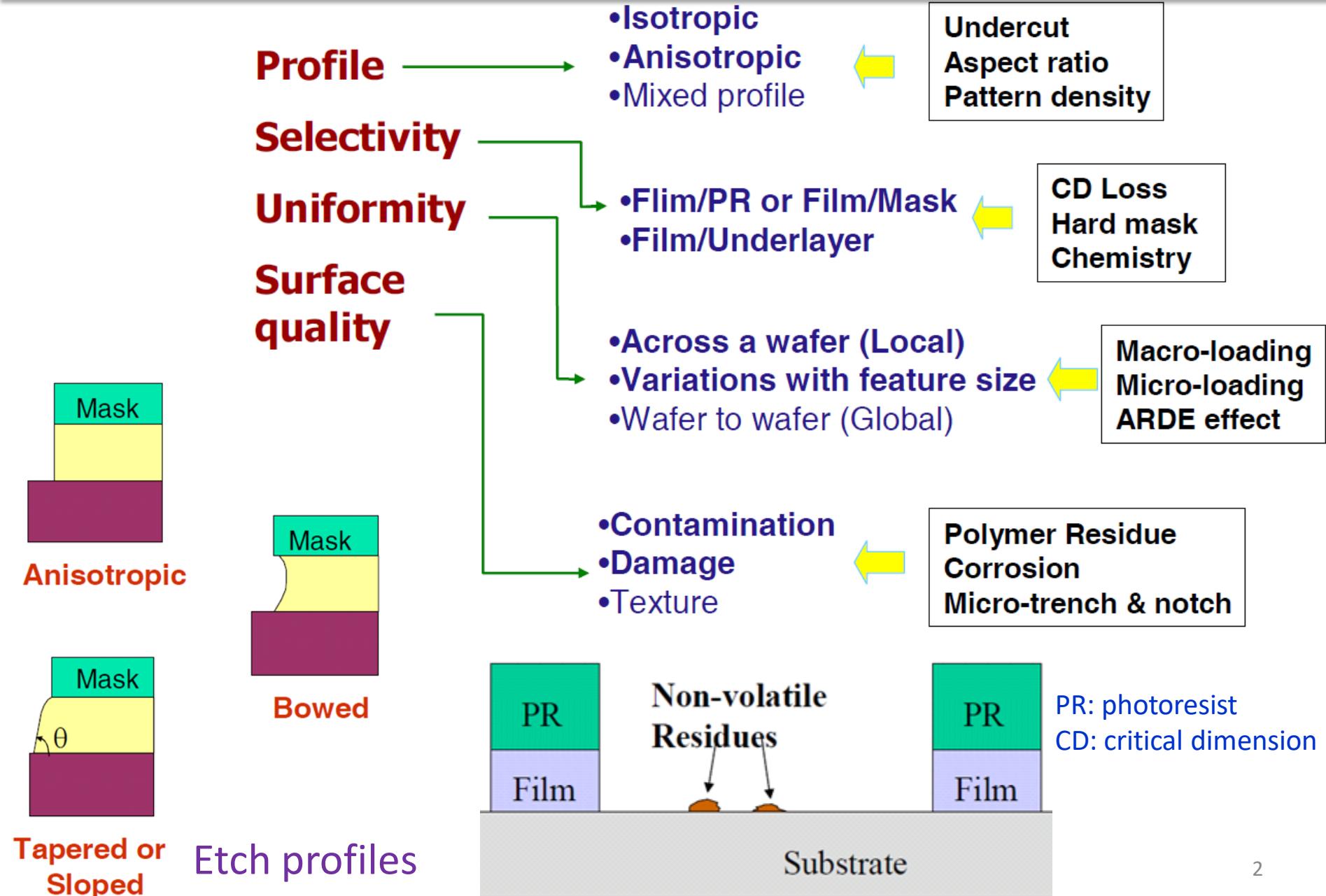
Dry etching techniques: summary



Chapter 10 Etching

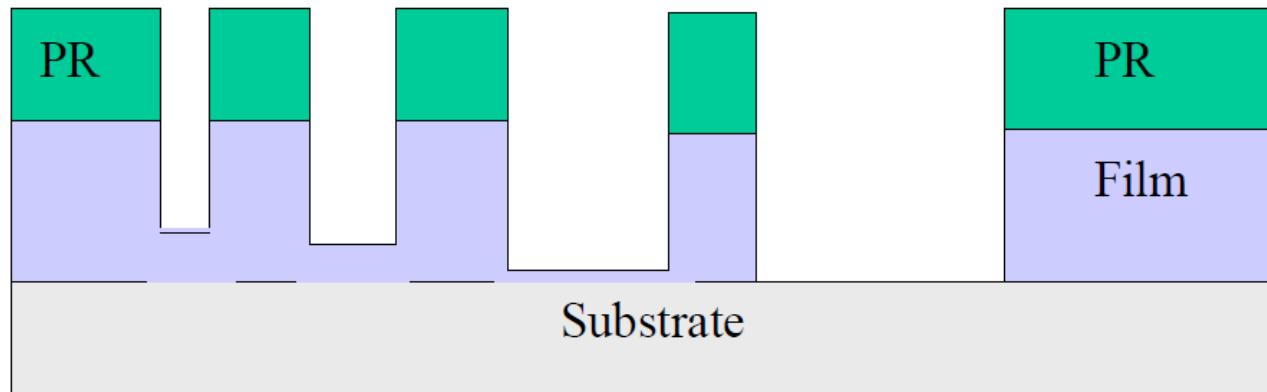
1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Main issues in plasma etching

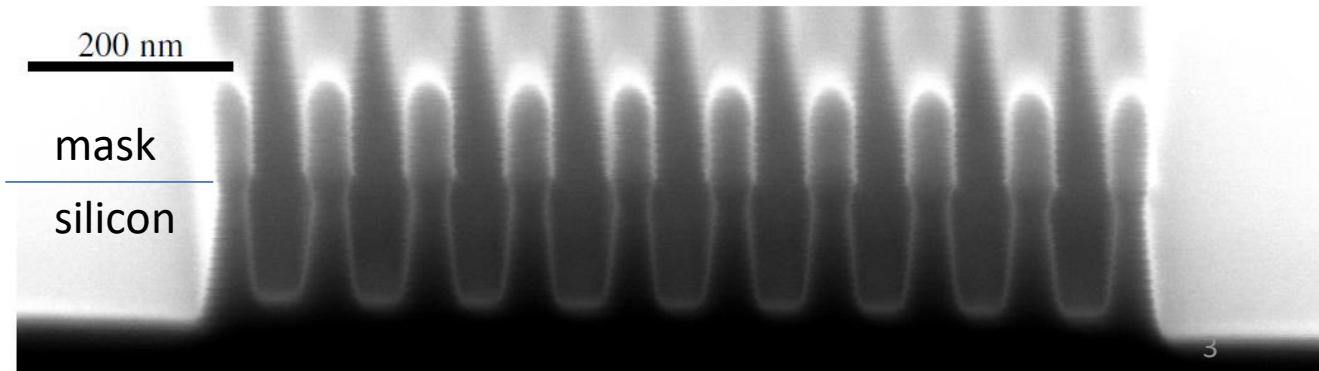


Aspect ratio and micro-loading effect

- Micro-loading: etch rate depends on local pattern density.
- Aspect-ratio dependent etching, RIE-lag: smaller features etch more slowly, because:
 - Etchants are more difficult to pass through the smaller hole.
 - Etch byproducts are harder to diffuse out.
- Lower pressure can minimize the effect: more directional and longer mean free path, easier for etchants to reach the trench/hole bottom and for etch byproducts to get out.



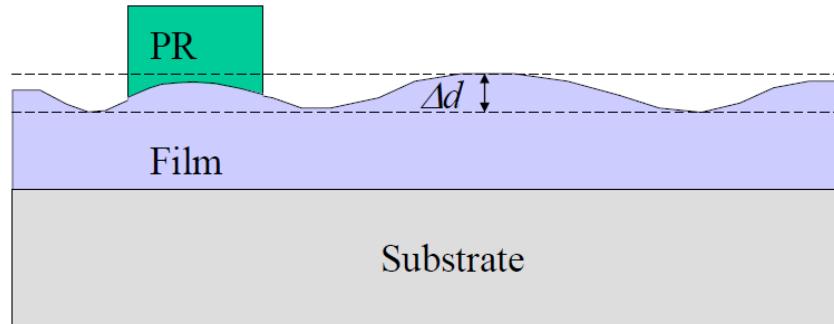
Micro-loading and sidewall bowing in Si etched with HSQ/AZPN114 bi-layer mask in Cl₂/HBr plasma.



When to stop etching: main etch & over etch

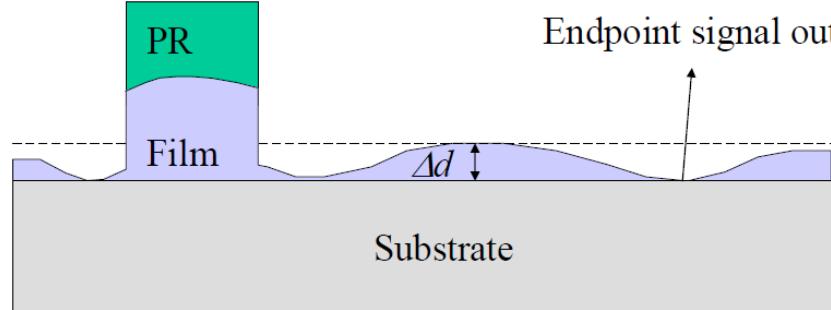
- Film thickness and etch rate is not uniform, needs over-etch to remove possible leftover film.
- Selectivity of etched film and substrate is very important.
- RIE uses **optical endpoint detection** to switch from main etch to over etch.

Start main etch

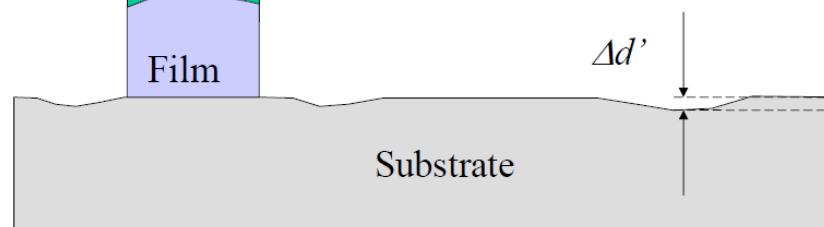


For R&D tool, usually no endpoint detection installed, thus often one does a lot of over-etch to make sure etching is finished.

Before over etch



After over etch



Etching endpoint detection by optical emission spectroscopy

- Each atom/molecule has its own emission wavelength.
- Color of plasma changes when etch different materials
- Optical sensors can be used to detect the change and indicate the endpoint for plasma etch processes.

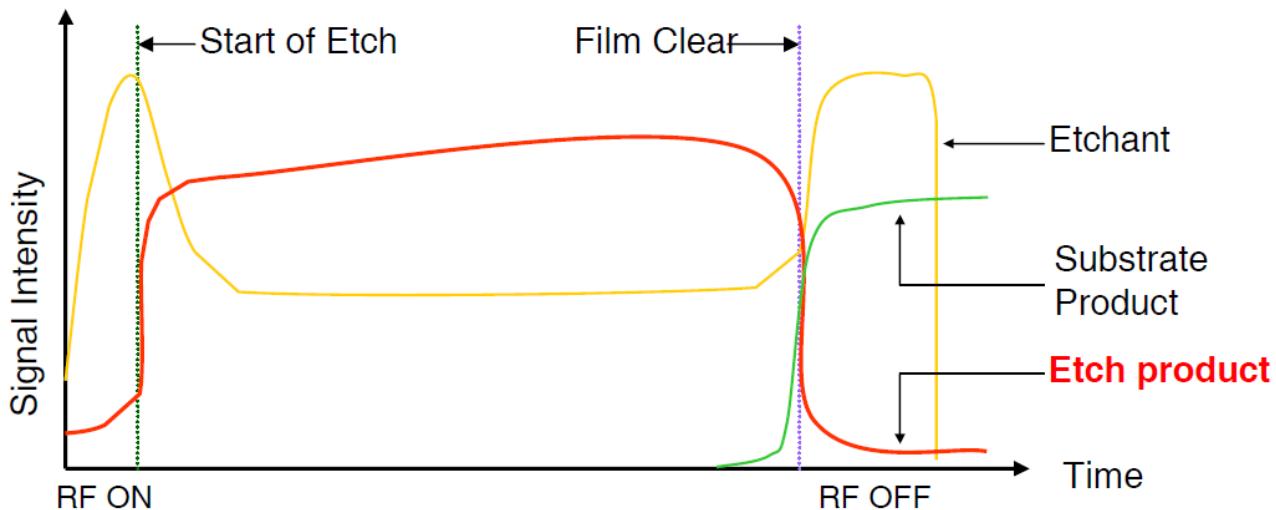
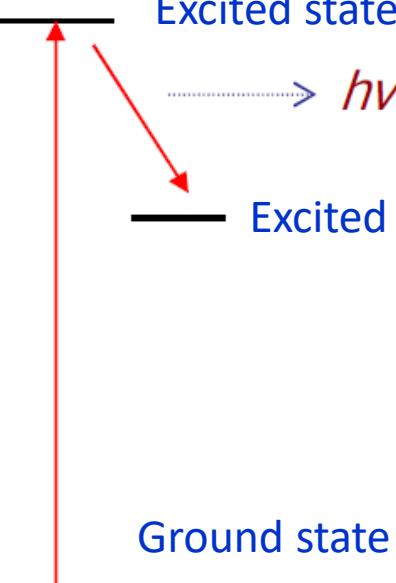
$$E = h\nu = h \frac{c}{\lambda}$$

$$\lambda \text{ (nm)} = \frac{1240}{E \text{ (eV)}}$$

Excited state I

$> h\nu$

Excited state II



Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch system.
7. Dry etching issues.
8. Dry etching method for various films.
9. Deep Si etching (can etch through a wafer).

Plasma etch methods for various films: overview

- Most reactant gasses contain halogens: F, Cl, Br, or I
- Exact choice of reactant gasses to etch each specific film depends on
 - Ability to form volatile by-products that can be removed by pumping
 - Selectivity and anisotropy.
- Boiling points are good indicators of volatility of species
 - Lower boiling point, higher tendency to evaporate.
 - High boiling point may need etching at elevated temperatures.

Table 6.3 Volatility of etched compounds with different gas chemistries (Reproduced from [48] with permission)

Elements	Fluorides	Boiling temperature (°C)	Chlorides	Boiling temperature (°C)	Bromides	Boiling temperature (°C)	Hydrides, trimethyls	Boiling temperature (°C)
Al	AlF ₃	1297 (subl.)	AlCl ₃	178 (subl.)	AlBr ₃	263		
As	AsF ₃	-63	AsCl ₃	130.2	AsBr ₃	221	AsH ₃	-55
	AsF ₅	-53			AsBr ₅			
C	CF ₄	-128	CCl ₄	77	CBr ₄	189	CH ₄	-164
Cr	CrF ₂	>1300	CrO ₂ Cl ₂	117	CrBr ₂	842		
Cu	CuF	1100 (subl.)	CuCl	1490	CuBr	1345		
	CuF ₂	950	CuCl ₂	993			CuH	55–60
Ga	GaF ₃	1000	GaCl ₃	201.3	GaBr ₃	278.8	Ga(CH ₃) ₃	134
Ge	GeF ₄	-37 (subl.)	GeCl ₄	84	GeBr ₄	186.5	GeH ₄	-88.5
In	InF ₃	>1200	InCl ₃	300 (subl.)			In(CH ₃) ₃	55.7
Mo	MoF ₅	213.6	MoCl ₅	268				
	MoF ₆	35	MoOCl ₃	100 (subl.)				
	MoO ₂ F ₂	270 (subl.)						
	MoOF ₄	180						
P	PF ₃	-101.5	PCl ₃	75	PBr ₃	172.9	PH ₃	-87.7
	PF ₅	-75	PCl ₅	162 (subl.)	PBr ₅	106		
Si	SiF ₄	-86	SiCl ₄	57.6	SiBr ₄	154	SiH ₄	-111.8
Ta	TaF ₅	229.5	TaCl ₅	242	TaBr ₅	348.8		
Ti	TiF ₄	284 (subl.)	TiCl ₄	136.4	TiBr ₄	230		
W	WF ₆	17.5	WCl ₆	346.7	WBr ₅	333		
	WOF ₄	187.5	WCl ₅	275.6	WOBr ₄	327		
			WOCl ₄	227.5				

Typical or representative plasma etch gases for films used in IC fabrication

Table 10-3

Material	Etchant	Comments
Polysilicon	SF ₆ , CF ₄ CF ₄ /H ₂ , CHF ₃ CF ₄ /O ₂ HBr, Cl ₂ , Cl ₂ /HBr/O ₂	Isotropic or near isotropic (significant undercutting); poor or no selectivity over SiO ₂ Very anisotropic, non-selective over SiO ₂ Isotropic, more selective over SiO ₂ Very anisotropic, most selective over SiO ₂
Single crystal Si	same etchants as polysilicon	
SiO ₂	SF ₆ , NF ₃ , CF ₄ /O ₂ , CF ₄ CF ₄ /H ₂ , CHF ₃ /O ₂ , C ₂ F ₆ , C ₃ F ₈ CHF ₃ /C ₄ F ₈ /CO	Can be near isotropic (significant undercutting); anisotropy can be improved with higher ion energy and lower pressure; poor or no selectivity over Si Very anisotropic, selective over Si Anisotropic, selective over Si ₃ N ₄
Si ₃ N ₄	CF ₄ /O ₂ CF ₄ /H ₂ CHF ₃ /O ₂ , CH ₂ F ₂	Isotropic, selective over SiO ₂ but not over Si Very anisotropic, selective over Si but not over SiO ₂ Very anisotropic, selective over Si and SiO ₂
Al	Cl ₂ Cl ₂ /CHCl ₃ , Cl ₂ /N ₂	Near isotropic (significant undercutting) Very anisotropic; BCl ₃ often added to scavenge oxygen.
W	CF ₄ , SF ₆ Cl ₂	High etch rate, non-selective over SiO ₂ Selective over SiO ₂
Ti	Cl ₂ , Cl ₂ /CHCl ₃ , CF ₄	
TiN	Cl ₂ , Cl ₂ /CHCl ₃ , CF ₄	
TiSi ₂	Cl ₂ , Cl ₂ /CHCl ₃ , CF ₄ /O ₂	
Photoresist	O ₂	Very selective over other films ⁸

RIE/plasma etch gases: from another book

Table 6.2 Materials and corresponding RIE gases

Materials to be etched	Chemical gases (multi choices)
Single-crystal silicon	CF ₃ Br, HBr/NF ₃ , SF ₆ /O ₂
Polysilicon	SiCl ₄ /Cl ₂ , BCl ₃ /Cl ₂ , HBr/Cl ₂ /O ₂ , HBr/O ₂ , Br ₂ /SF ₆
Al	SiCl ₄ /Cl ₂ , BCl ₃ /Cl ₂ , HBr/Cl ₂
Al-Si-Cu, Al-Cu	BCl ₃ /Cl ₂ + N ₂
W	SF ₆ , NF ₃ /Cl ₂
TiW	SF ₆
WSi ₂ , TiSi ₂ , CoSi ₂	CCl ₂ F ₂ /NF ₃ , CF ₄ /Cl ₂
SiO ₂	CCl ₂ F ₂ , CHF ₃ /CF ₄ , CHF ₃ /O ₂ , CH ₃ CHF ₂
Si ₃ N ₄	CF ₄ /O ₂ , CF ₄ /H ₂ , CHF ₃ , CH ₃ CHF ₂
GaAs	SiCl ₄ /SF ₆ , SiCl ₄ /NF ₃ , SiCl ₄ /CF ₄
InP	CH ₄ /H ₂
Photoresists	O ₂

Most lab systems have only fluorine-based gases (SF₆, CF₄, CHF₃) since they are relatively safe; chlorine-based gases are corrosive. Most RIE has Ar and O₂ gas, some has H₂ and He.

Plasma etching of silicon dioxide (using CF_4)



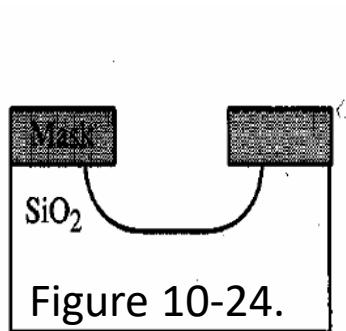
- Etching is mostly done by reactive neutral species - free radicals such as F.
- Additives like O_2 can be used that react with CF_3 (and CF_2 , form CO and CO_2) and reduce $CF_3 + F$ recombination, thus higher F concentration and slightly higher etch rate.
- However, too much O_2 reduces etch rate due to dilution effect, competition of O with F for surface sites, and reversal of the second reaction (form oxide with SiF_4).
- These processes are chemical (spontaneous, can etch without ion bombardment) and are therefore pretty isotropic and selective.
- Nonetheless, for low pressure where ion energy is high, damage-driven ion-enhanced attack dominates, thus anisotropic etching.
- The situation is similar for SF_6 and NF_3 gas.
- The key to promote anisotropy as well as increasing selectivity over Si is to reduce F and increase C, because then:
 - Etching is more done by CF_3 or CF_2 (than by F), which is strongly ion-enhanced (appreciable etching happens *only* where ion bombardment occurs).
 - Polymer inhibitor formation is increased, which protects the sidewall.
- This is realized by adding H_2 that reacts with/consumes F to form HF.
- Or by using fluorine-poor and carbon-rich gases such as CHF_3 , C_3F_8 and C_2F_6 .

Effect of H₂ addition

- The addition of H₂ removes F from the system by forming stable HF (not very reactive) gas, which reduces F/C ratio and slows down the formation of SiF₄ and thus the etching rate.
- With less F, the chance of CF_{2,3} to combine with F to form CF₄ is lower, thus higher CF_{2,3} concentration, which promotes the formation of fluorocarbon polymer.
- Addition of H₂ increases selectivity of SiO₂ with respect to silicon. Because:
 - Increased polymerization tends to inhibit etching. However, on horizontal surfaces , ionic bombardment provides enough energy, causing the carbon/hydrogen (F-C polymer on oxide surface) to combine with surface oxygen, releasing CO and H₂O.
 - Once O in Si-O is gone, Si has a dangling bond, which is removed by combining with released fluorine radicals.
- Silicon will not be etched (if H₂ is high enough) because of the absence of oxygen at the surface.

Effect of C/F ratio

Inhibitor: Teflon like fluorocarbon polymer

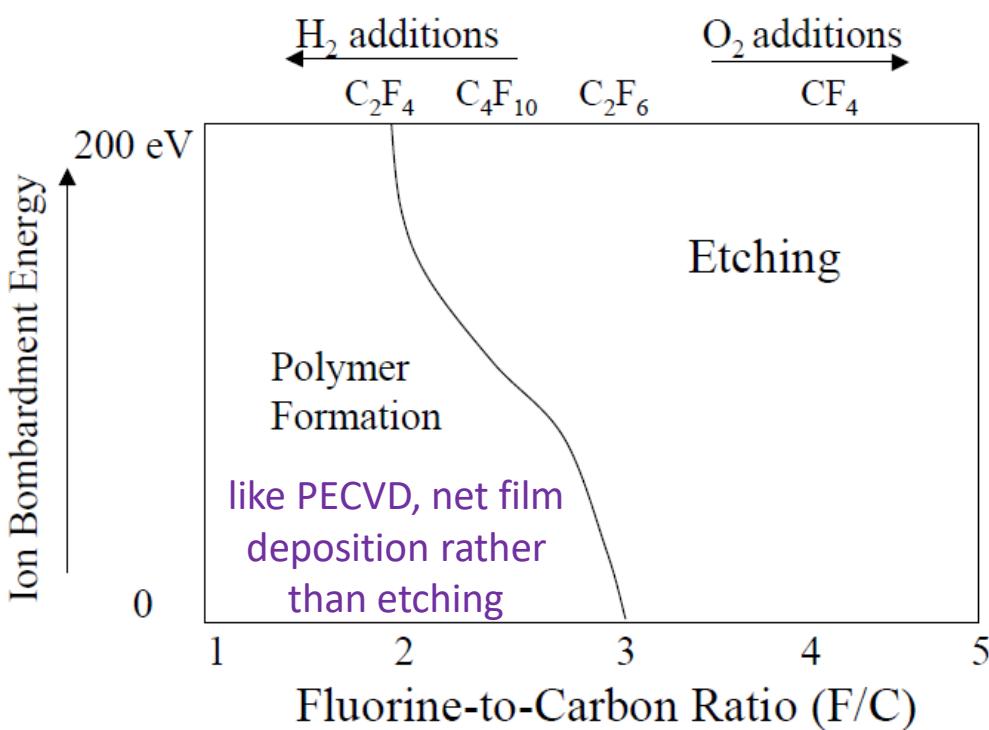


(a) Little or No Inhibitor Deposition
(e.g. $\text{CF}_4 + \text{O}_2$)

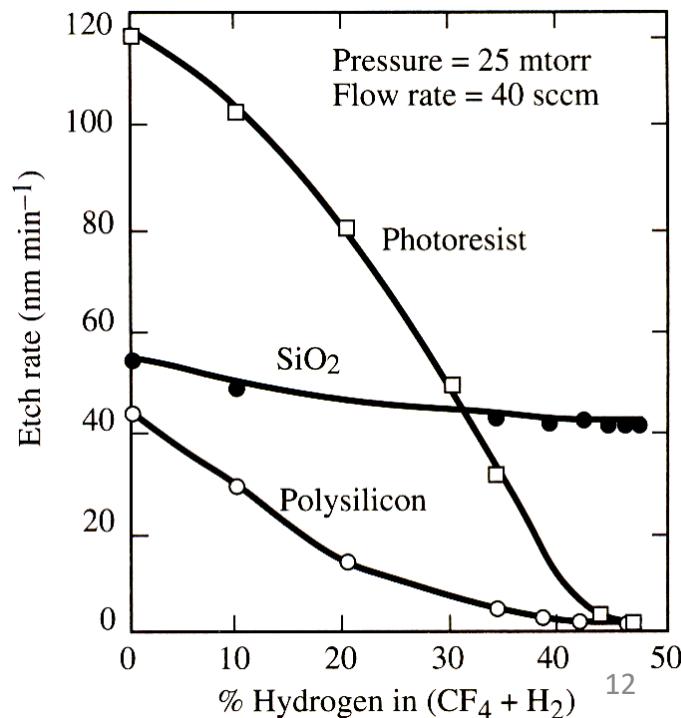
(b) Inhibitor Deposition
(e.g. $\text{CF}_4 + \text{H}_2$,
or $\text{CHF}_3 + \text{O}_2$)

(c) Excess Inhibitor Deposition
(e.g. CHF_3)

Higher C/F



Fluorocarbon polymer

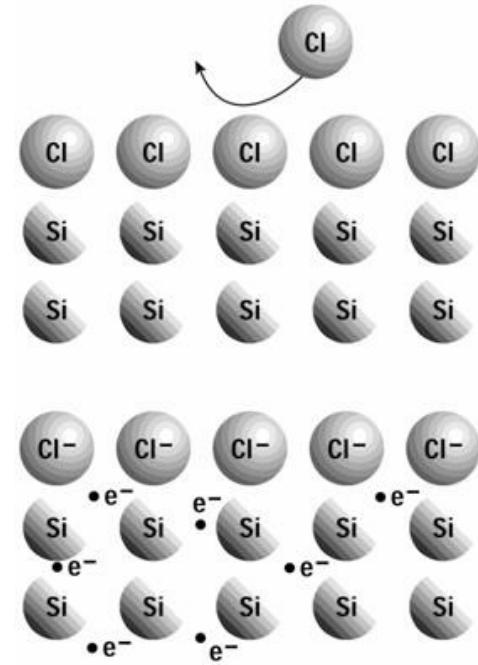


Plasma/RIE etching of silicon

- F-based chemistry (CF_4 , SF_6 , NF_3) tend to be isotropic.
 - When anisotropy is not important, SF_6/O_2 is a good chemistry for high selectivity.
 - When anisotropy is desired, one can use CF_4/H_2 (or CHF_3), but undercutting may still occur.
 - Or just use low pressure to promote ion-damage enhanced etching.
 - Hard to achieve using F simultaneously high anisotropy and selectivity over SiO_2 .
- Cl-based chemistry (Cl_2 , HCl , SiCl_4 , BCl_3) result in anisotropic and selective etching (yet etch rate lower than F chemistry).
 - Etch needs to be ion bombardment-enhanced.
 - Thus anisotropic without polymer inhibitor formation.
 - Selectivity to oxide is very high (100:1).
 - Anisotropy enhanced by adding small amount of O_2 (form SiO_2 inhibitor on sidewall, those formed on bottom bombarded away)
- Br-based chemistry (HBr , Br_2) are similar to chlorine based etchants (etch rate slower than F or Cl).
 - More anisotropic and selective to oxide than F- and Cl-based RIE without polymer inhibitor
 - Adding O_2 promotes inhibitor formation (forming SiO_2 from Si) and increase selectivity over oxide (due to removal of C from resist erosion).

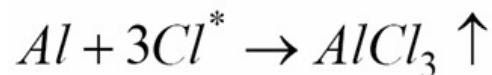
Chlorine Si RIE: doping is also very important

- Un-doped Si etch slowly in Cl/Cl₂ ambient without addition of ion bombardment.
- Heavily n-type doped Si etches without bombardment in Cl, but not in Cl₂.
- Doping implies that Cl etching involves electron transfer from substrate.
- Atomic Cl chemisorbs on Si. Once surface Cl becomes negatively charged, it can bond ionically with the substrate. This frees additional chemisorption sites and increases probability that Cl atoms penetrate surface and produce volatile silicon chlorides.
- Cl penetration is increased by ion bombardment, leading to energy-driven anisotropy.
- Charge transfer produces isotropic etch in heavily doped layers. Yet one can obtain anisotropic etch by sidewall polymerization.
- Polymerization is done by adjusting relative concentrations of Cl₂ and BCl₃, CCl₄ or SiCl₄. Like CF₄ plasma, in Cl₂/CCl₄ gas, etching is mainly by atom (Cl), and polymerization by molecule radial (CCl₃, CCl₂).



For un-doped Si (top), once a monolayer of Cl builds on surface, it impedes any Cl addition.

Plasma/RIE etching aluminum



Al etching using CCl_4 ,
(Cl_2 is more popular)

- Fluorine is not used because AlF_3 is not volatile.
- Presence of native oxide Al_2O_3 on Al surface requires a breakthrough etch before the main etch: Ar sputter, or use BCl_3 , $SiCl_4$, CCl_4 or BBr_3 to remove Al_2O_3 by both physical and chemical means.
- Cl_2 etches Al isotropically (pure chemical, no need of ion bombardment for etching).
- For anisotropic etching, need form sidewall inhibitor: add $CHCl_3$, $CFCl_3$, CCl_4 , BCl_3 or N_2 (form CN polymer from resist erosion).
- Al/Cu alloys are used in interconnects but Cu does not etch in Cl.
 - Etch requires ion bombardment or high temperature
- Corrosion of Al line occurs when exposed to ambient because Cl or $AlCl_3$ on sidewall and resist react with water to form HCl which etches Al.
- To passivate Al surface after etch before exposure to atmosphere:
 - Heat wafer to 100-150°C to drive-off Cl.
 - Bury Cl with CHF_3 plasma-induced C-F polymer and remove the polymer later.
 - Expose to F ambient such as CF_4 plasma to replace Cl with F.
 - Then O_2 plasma to remove photoresist, followed by DI water rinse.

Chapter 10 Etching

1. Introduction to etching.
2. Wet chemical etching: isotropic.
3. Anisotropic etching of crystalline Si.
4. Dry etching overview.
5. Plasma etching mechanism.
6. Types of plasma etch systems.
7. Dry etching issues.
8. Dry etching methods for various films.
9. Deep Si etching (can etch through a wafer).

Why deep RIE (DRIE)?

- Plasma etching can produce deeper trenches than wet etching, but with tapered angles.
- Tapered trenches are not desirable in many applications such as resonators that involve pairs of “centipedes-like” micro-devices with overlapped “fingers”.
- DRIE process may produce deep high aspect ratio structures with vertical sidewall ($\theta \approx 0^\circ$).
- It is the most important breakthrough in drying etching in recent years, popular for MEMS (micro electro mechanical systems) fabrication.

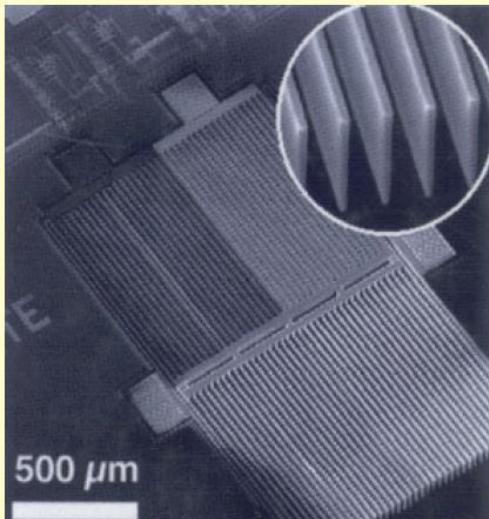
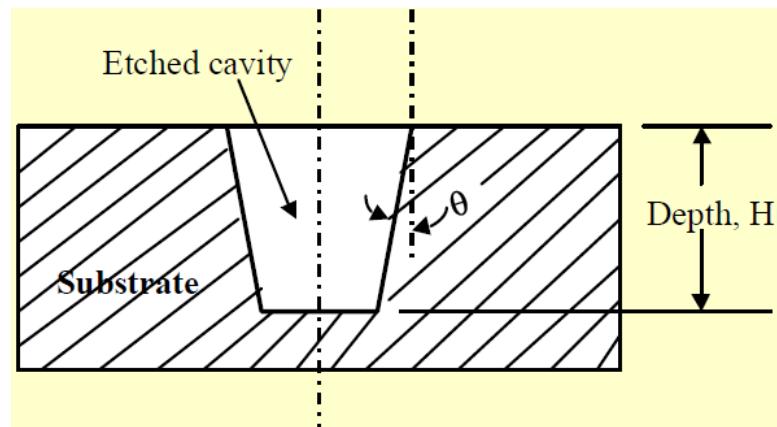
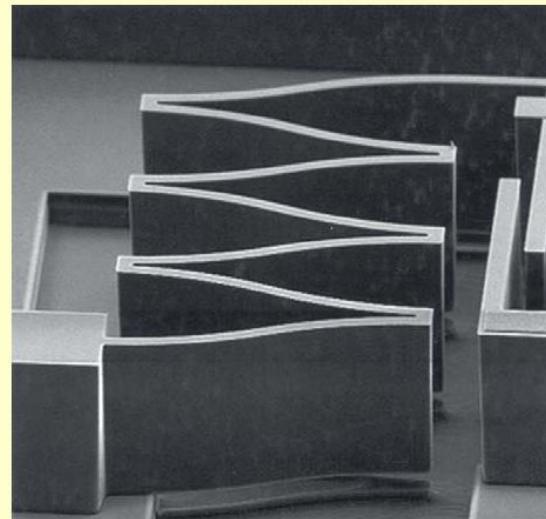


Plate- resonators



Spring resonator

Working principle:

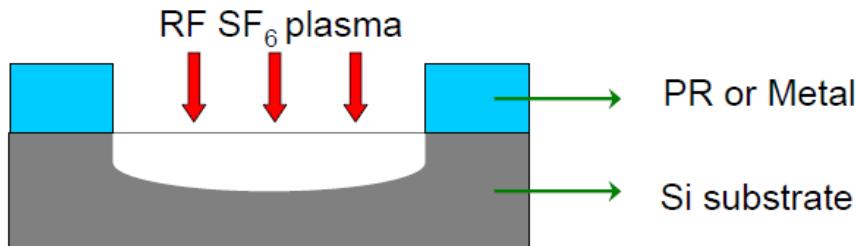
- The DRIE process provides thin film of a few microns protective coating on the sidewalls during the etching process.
- It involves the use of a high-density plasma source.
- The process allows alternating process of plasma (ion) etching of the substrate material and the deposition of etching-protective material on the sidewalls.

Deep Si etch: ICP - “Bosch” process

(a) Etch Step

ICP: inductively coupled plasma

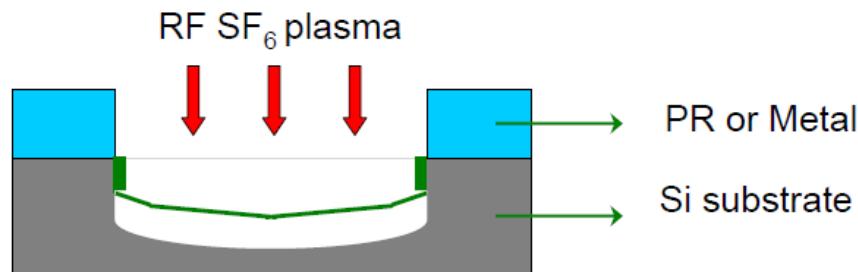
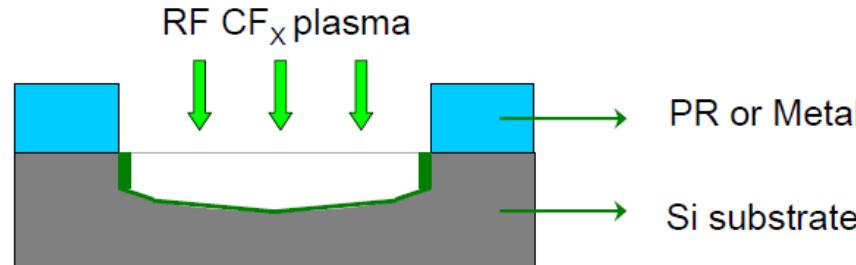
ECR: electron cyclotron resonance



(b) Passivate Step

Besides Bosch process, the other very popular deep Si etch is *cryo-etch* (i.e. at very low temperature, order -100°C, SiO_xF_y as inhibitor). Often a deep RIE tool can do both processes.

(a') Etch Step



- Uses high density plasma (ICP is used, but ECR also works) to alternatively etch silicon and deposit an etch-resistant polymer on sidewalls.
- SF₆ etch 5-13 sec; followed by C₄F₈ fluorocarbon polymer deposition 5-10 sec.
- Etch rate several $\mu\text{m}/\text{min}$, capable of etching several hundred μm with vertical walls.
- Sidewall is rough, depending on cycle times (longer cycle, more zigzag).
- Process recipe depends on geometry (aspect ratio...).
- More popular for MEMS, less common for nano-fabrication due to sidewall zigzag.

Deep Si etch - Bosch process

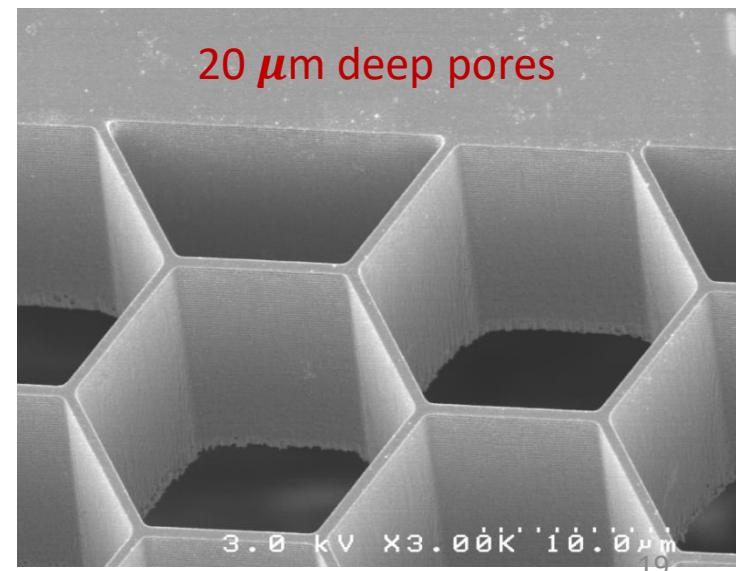
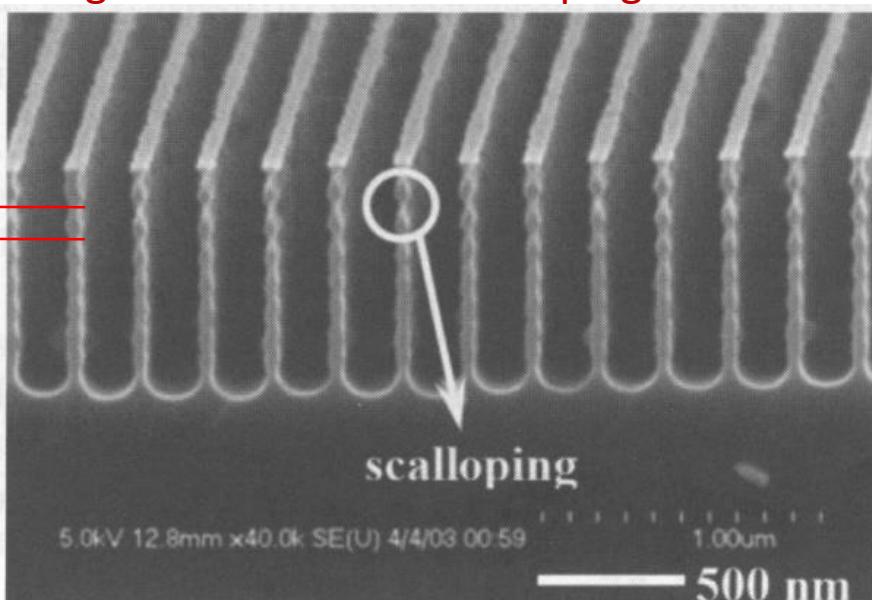
Table 6.4 Typical “Bosch” process conditions

Process parameters	Passivation	Etching
C ₄ F ₈	85 sccm	0 sccm
SF ₆	0 sccm	130 sccm
RF power at stage	0 W	12 W
RF power from coil	600 W	600 W
Cycle time	7.0 s	9.0 s
Delay time	0.5 s	0.5 s
Etch rate		1.5–3 $\mu\text{m min}^{-1}$

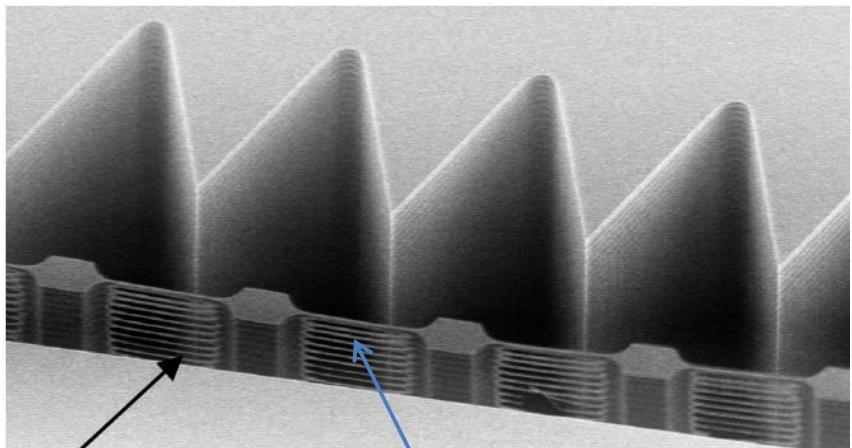
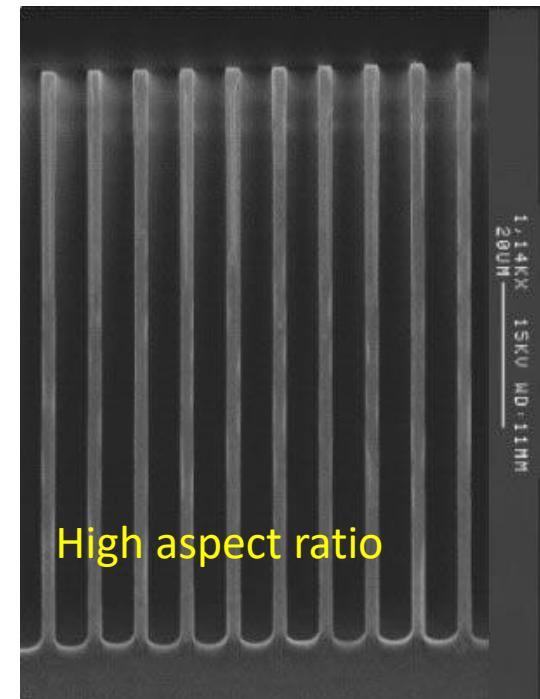
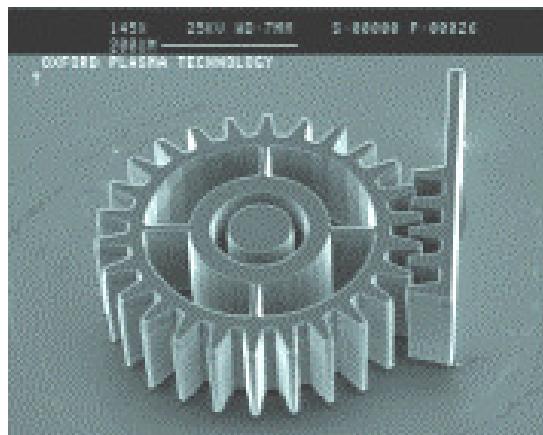
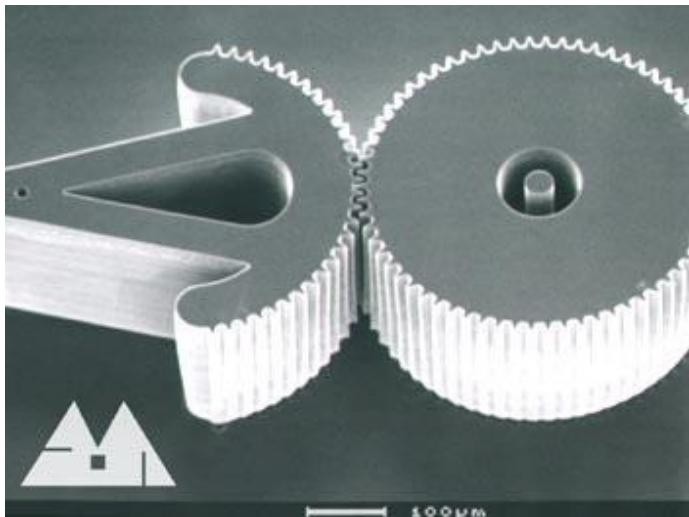
sccm: standard cubic centimeter per minute

DRIE uses lower energy ions → less damage and higher selectivity.

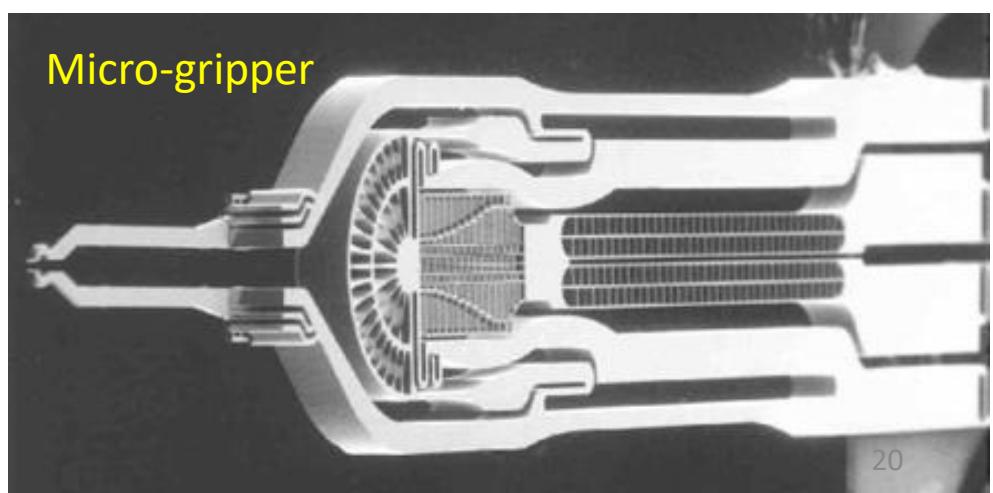
Plasma maintained at 0.5 to 3mTorr.



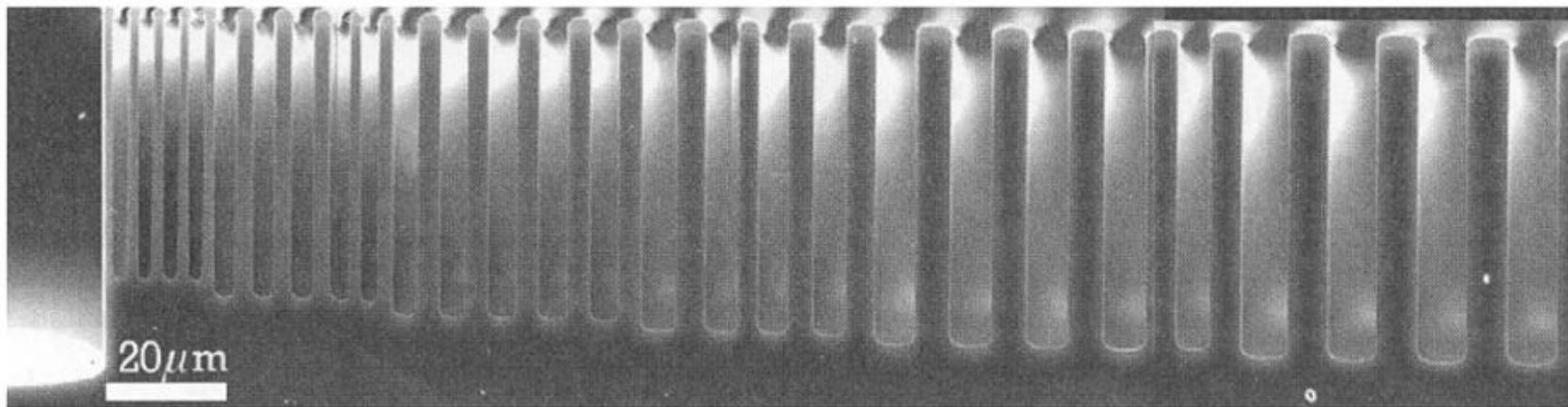
More examples of deep Si etch



2.13 K X 3μm
EHT = 3.10 kV Signal A = SE2
WD = 10 mm Photo No. = 499
Date : 25 Apr 2000
Time : 20:54



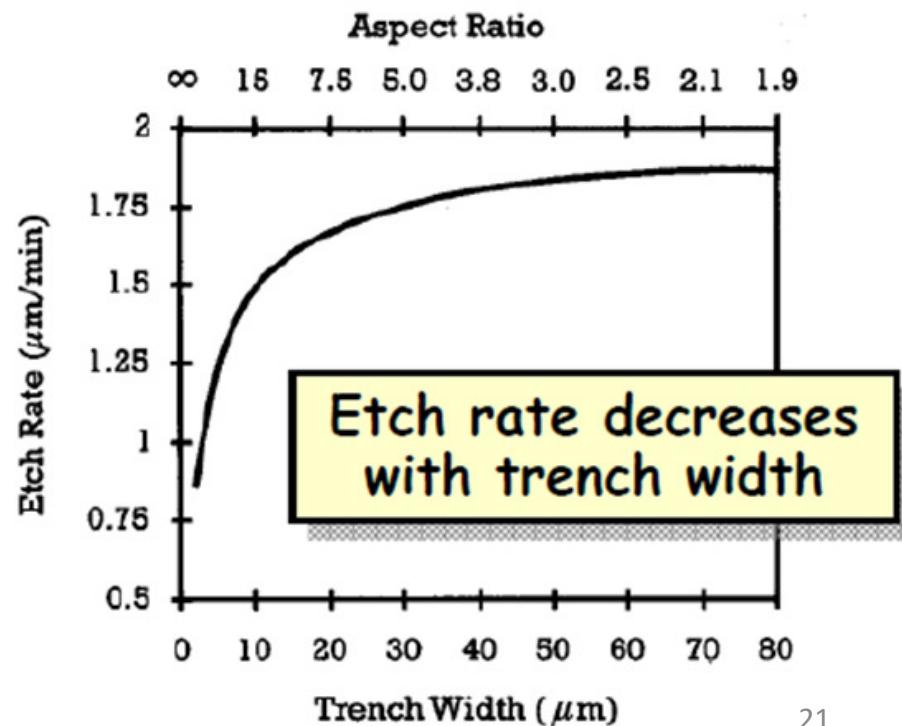
DRIE issues: etch rate variance (micro-loading effect)



Etch rate is diffusion-limited and drops for narrow trenches.

One can adjust mask layout (add “dummy” features) to eliminate disparities.

Or adjust process parameters (slow down the etch rate to that governed by slowest feature).



Laser-assisted chemical etching

Laser creates Cl radicals from Cl_2 , which react with Si to form SiCl_4 .

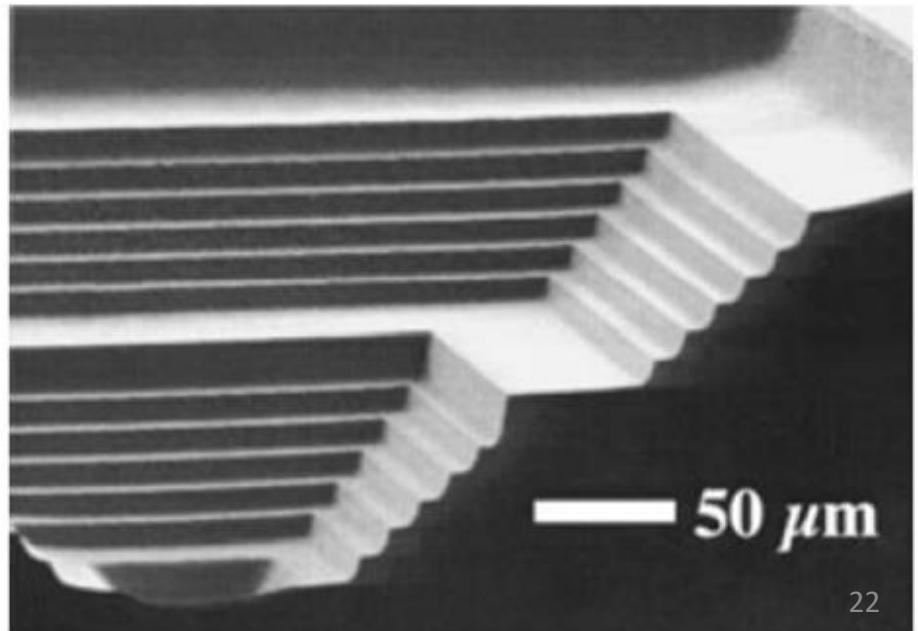
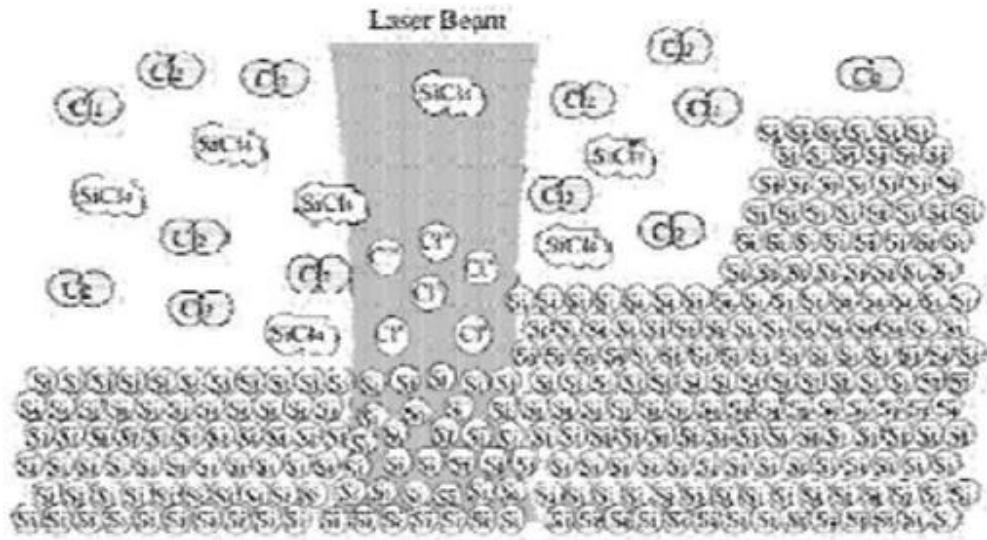
Etch rate: $100,000 \mu\text{m}^3/\text{sec}$. (3min to etch a $500 \times 500 \times 125 \mu\text{m}^3$ trench)

Surface roughness: 30nm rms.

Direct patterning, no need to pattern mask first.

(not popular – I have never seen this in cleanroom)

Laser assisted etching of a $500\mu\text{m} \times 500\mu\text{m}$ terraced silicon well, with each step $6\mu\text{m}$ deep.



Historical perspective: poly-silicon etch technology evolution

Geometry Requirements	Time Frame and Reactor Design	Chemistries	Strengths	Limitations and Problems	Controls
4 to 5 μm , isotropic etch	Pre-1977: wet etch	HF/HNO ₃ buffered with acetic acid or H ₂ O	Batch Process	Resist lift; bath aging; temperature sensitive	Operator judgement for endpoint
3 μm	1977: barrel etcher	CF ₄ /O ₂	Batch Process	Non-uniformity, isotropic etch, large undercut	Manometer and timer
2 μm	1981: single wafer etch	CF ₄ /O ₂	Single wafer; individual etch endpoint, improvement in repeatability	Low oxide selectivity; isotropic process	Endpoint detection
1.5 μm	1982: single wafer RIE	SF ₆ /Freon 11, SF ₆ /He	MFCs; independent pressure and gas flow control, improvement in repeatability	Low oxide selectivity; profile control	MFCs; separate gas flow and pressure control
To 0.5 μm	1983: variable gap; load-locked	CCl ₄ /He, Cl ₂ /He, Cl ₂ /HBr	Load-locked chamber; variable gap, improvement in repeatability	Microloading in high aspect ratios; profile control	Control of electrode gap; computer controls
To 0.25 μm and below	1991: inductively coupled plasma (ICP)	Cl ₂ , HBr	High-density plasma; low pressure; simple gas mixtures, improvement in repeatability	Complex tool; many variables	Independent RF control for plasma generation and wafer bias