6

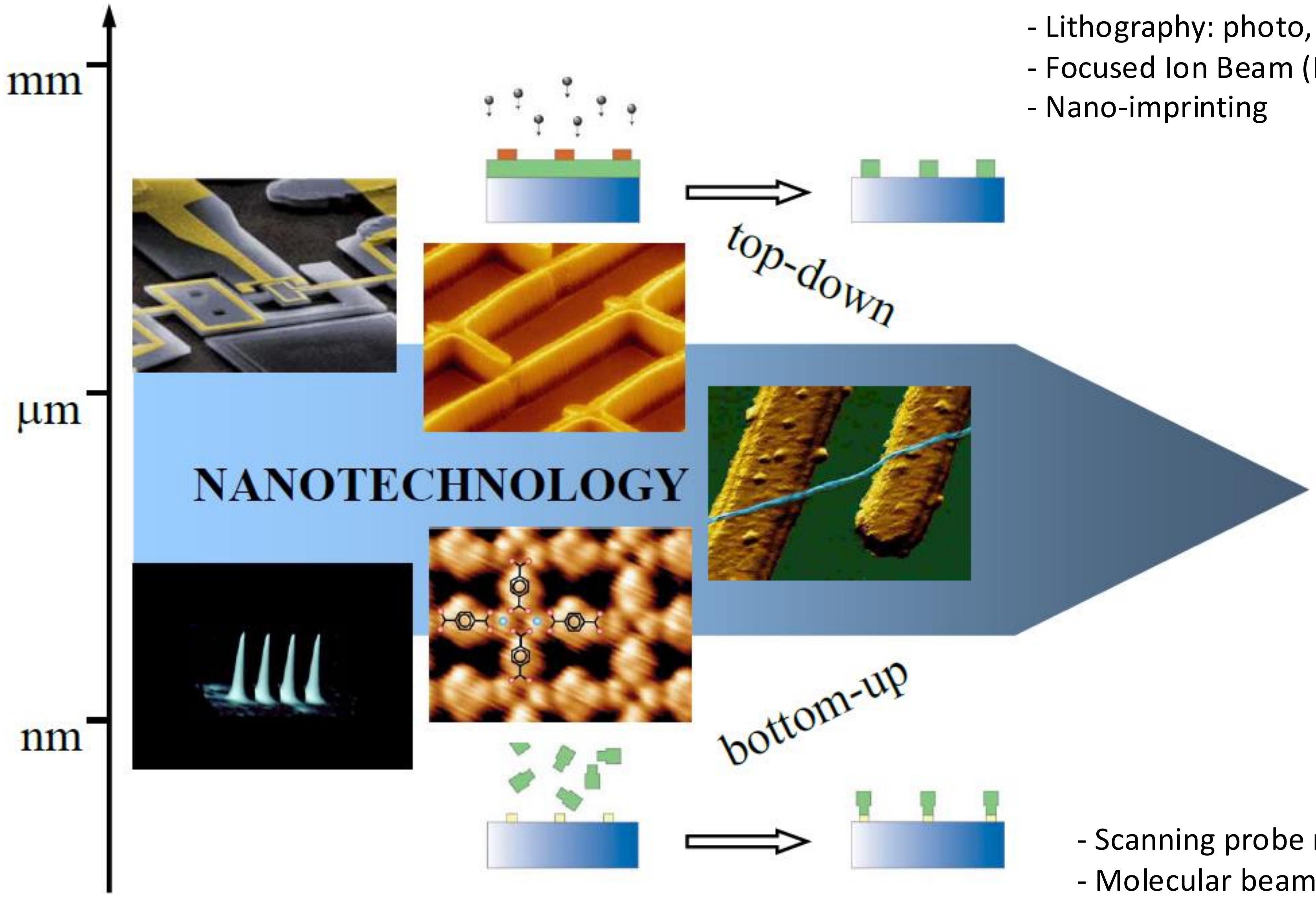
Top-down approach

Top-down

Nanofabrication
Principles, Capabilities and Limits
Zheng Cui
Springer
DOI: 10.1007/978-3-031-62546-6_1

Nanophysics and Nanotechnology
An Introduction to Modern Concepts in Nanoscience
E. L. Wolf
WILEY-VCH

How to make the nanostructures: two approaches



- Lithography: photo, UV, e-beam
- Focused Ion Beam (FIB)
- Nano-imprinting

- Scanning probe manipulation
- Molecular beam epitaxy
- Self-organized growth
- Self assembly

Top-down

The top-down approach is relevant in large-scale applications (devices), especially for parallel process techniques (lithography)

In general, it’s combined with other techniques (growth: MBE, ABE; etching, …)

Examples

Field Effect Transistor

and its **miniaturization**

Field Effect Transistor (FET)

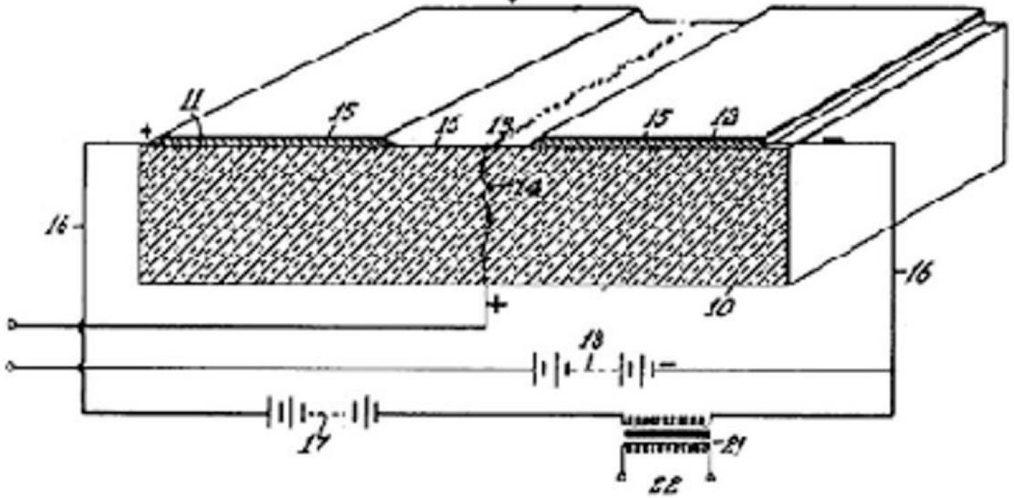
J.E. Lilienfeld, Bell Labs
US Patent (1926)
Semiconductor Field Effect Device

Jan. 28, 1930. J. E. LILIENFELD 1,745,175

METHOD AND APPARATUS FOR CONTROLLING ELECTRIC CURRENTS

Filed Oct. 8, 1926

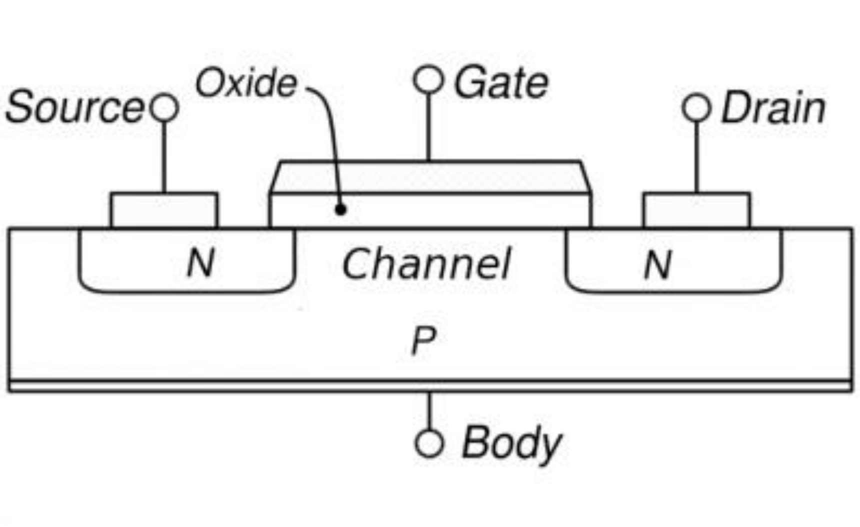
Fig.1.



Idea: make a switch based on semiconductor pn junctions; conduction between the two electrodes (source and drain) controlled by an electric field applied on a third electrode (gate)

It took time to realize it because of lack of knowledge of the surface properties of materials

MOSFET (Metal-Oxide - Semiconductor FET)



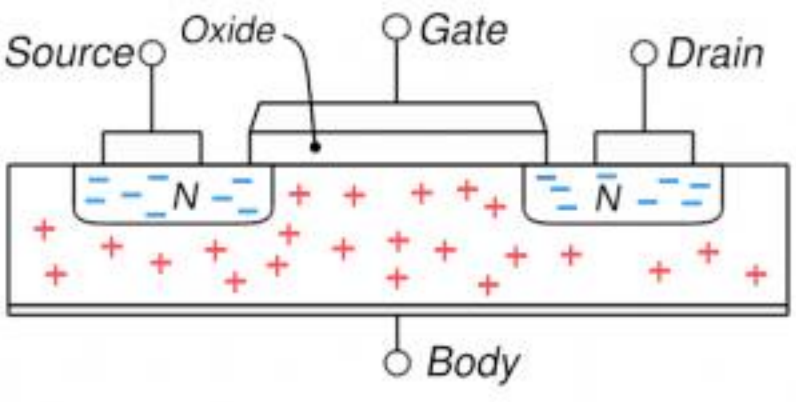
The source, drain, and channel are semiconductors.
Here, source and drain are n-type while the body is p-type; the converse is also possible

A transistor is a switch: conducting or insulating depending on the voltage on the gate

Basic operation: a voltage is applied across the source and drain to make current flow between them.
However, whether this happens or not depends on the voltage applied to the gate: the gate controls the conductive properties of the channel.
Changing the voltage on the gate changes the conductive properties of the channel and thus the current between source and drain.

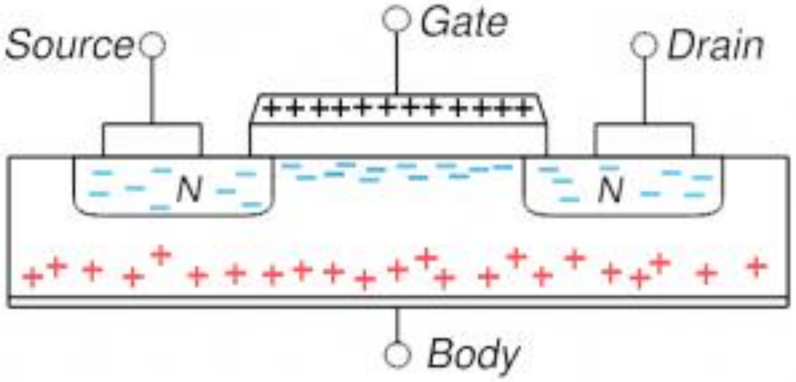
https://volga.eng.yale.edu/teaching-resources/semiconductors

MOSFET (Metal-Oxide - Semiconductor FET)



n-type source and drain: mobile electrons, p-type body: mobile holes.

- Equivalent to two diodes (two pn junctions) back-to-back: the first going from source to body and the next from body to drain, no current flow possible (OFF state)



Inversion

- High positive voltage on the gate: all holes driven away from the channel, and some electrons (from source and or drain) enter the channel region.
- The thin channel region is now n-type (has been inverted): n-type material connecting source and drain, no diode behavior and a very highly conductive state: current flows easily from source to drain when a voltage difference between the two is applied (ON state)

The gate voltage modulates the channel conductivity via electric field

Miniaturization - Moore's law

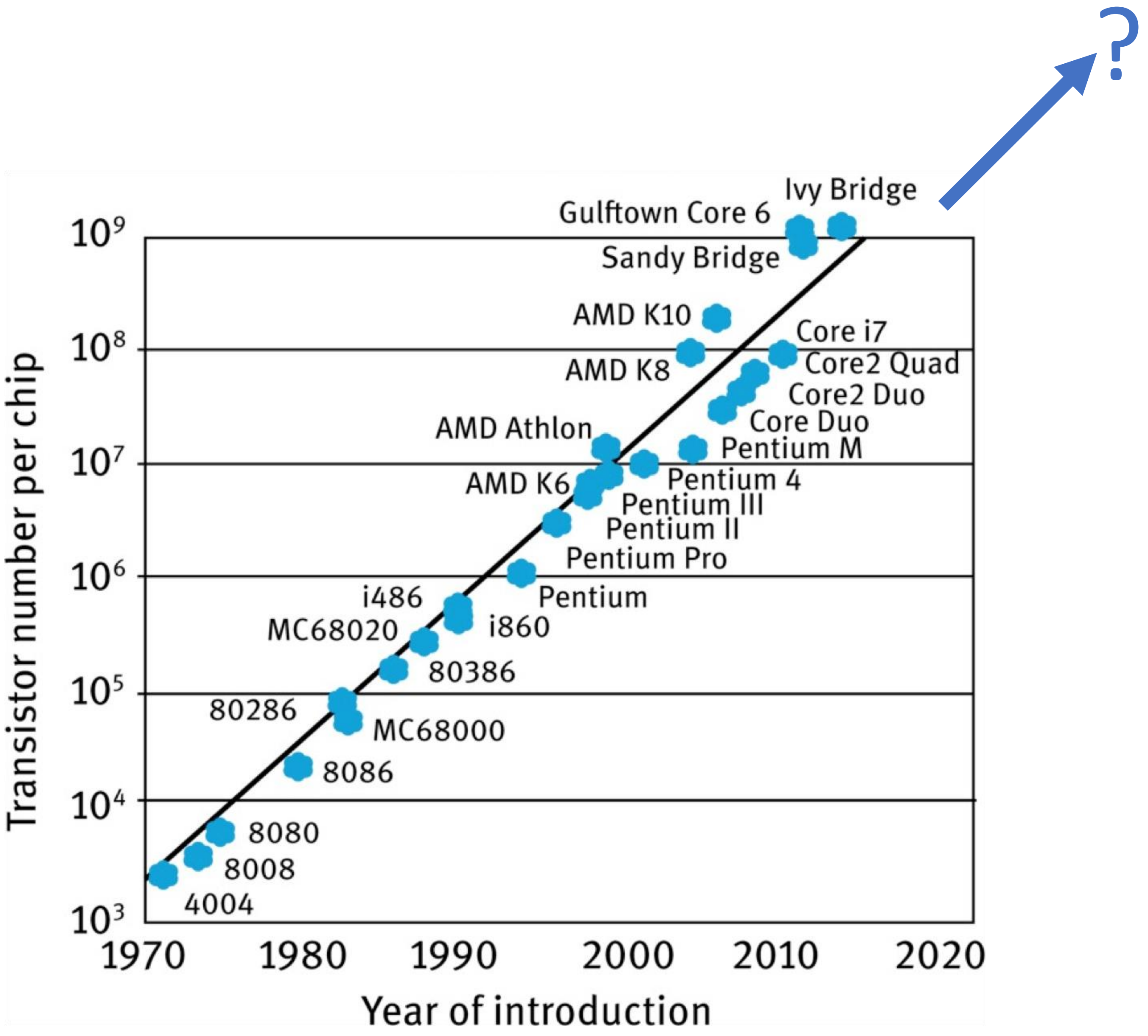
Gordon E. Moore
Co-founder, Intel Corporation.

Electronics, Volume 38, Number 8 (1965)

"The number of transistors on a piece of silicon will double every couple of years."

Questions:

1. How to make such transistors?
2. How small can they be?
3. What are the critical sizes?
4. How to increase the density beyond the size limitations?
5. …



Lithography

For modern devices:

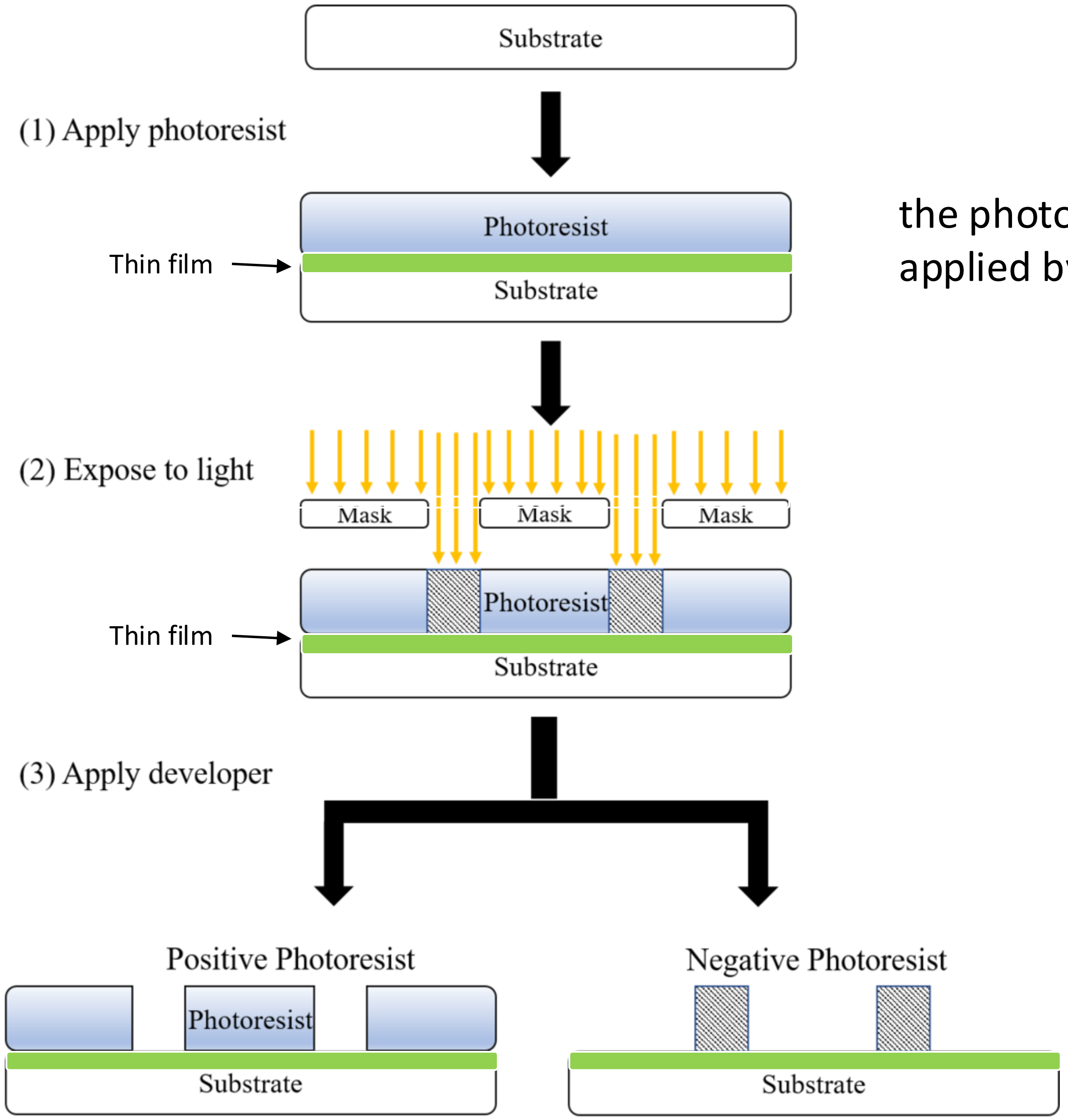
Lithography is the process of transferring a pattern into a reactive polymer film (a resist), which will subsequently be used to replicate that pattern into an underlying thin film or substrate.

Radiation sources: photons, X-rays, electrons, ions and neutral atoms.

Photolithography is the most widely used technique in microelectronic fabrication, particularly for mass production of integrated circuit.

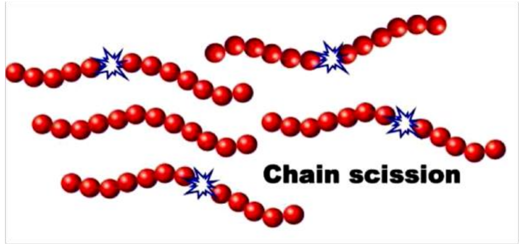
Lithographic stone

Projection photolithography

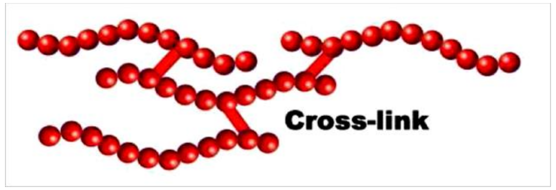


the photoresist is a polymer
applied by spin coating

positive resist:
light softens the resist



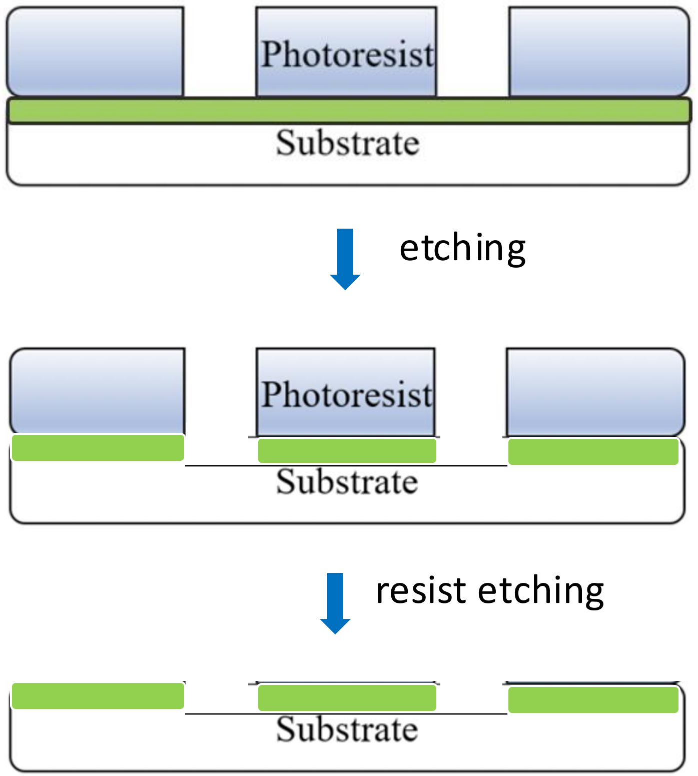
negative resist:
light hardens the resist



mask motif transferred
to the substrate

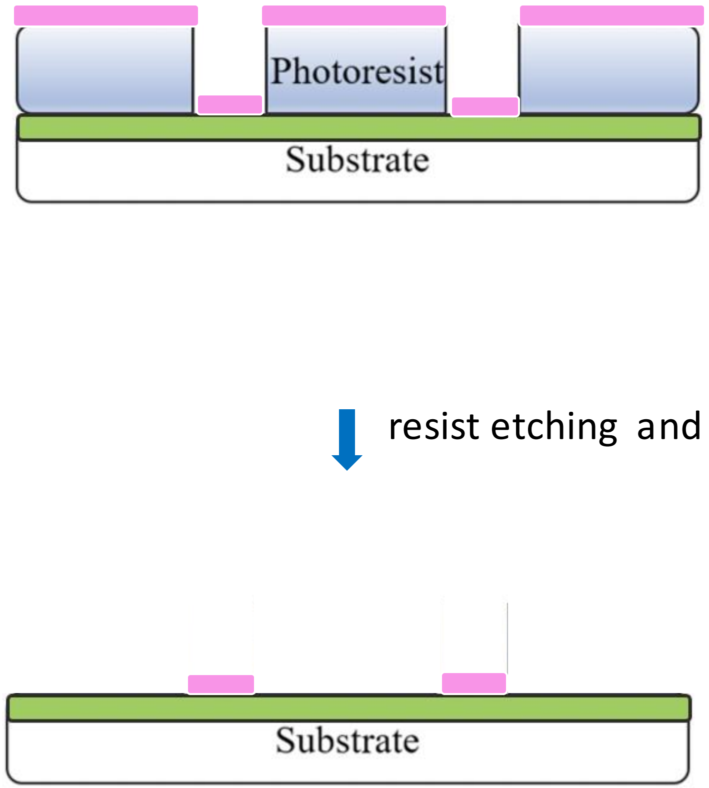
Projection photolithography

Examples of further steps:



etching

resist etching



growth of new layers
(MBE or ABE)

resist etching and lift off

Diffraction limit

Resolution limit of a projection photolithography system
i.e., size of the smallest feature

$$R \approx \frac{\lambda}{NA}$$

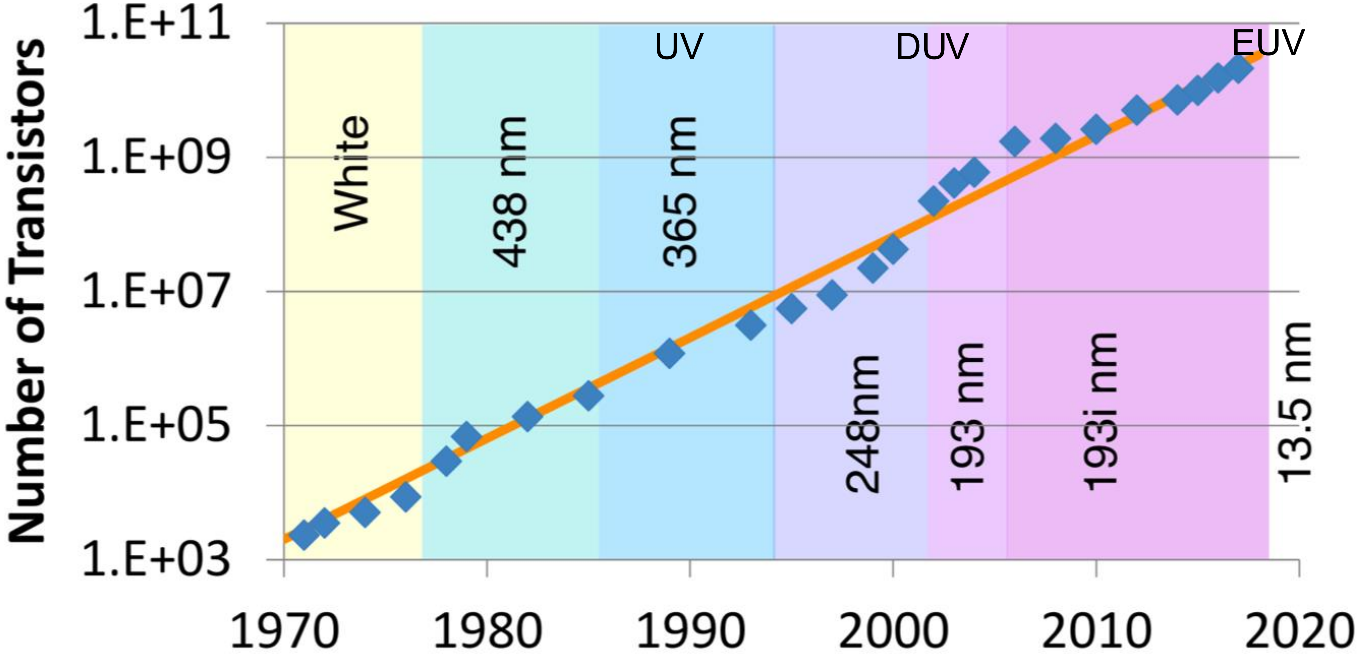
λ is the illumination wavelength
NA is the numerical aperture of the optical system

In first approximation: $R \approx \lambda$

The wavelenght limits the resolution and therefore the lateral size of the structures or the minimum distance between structures

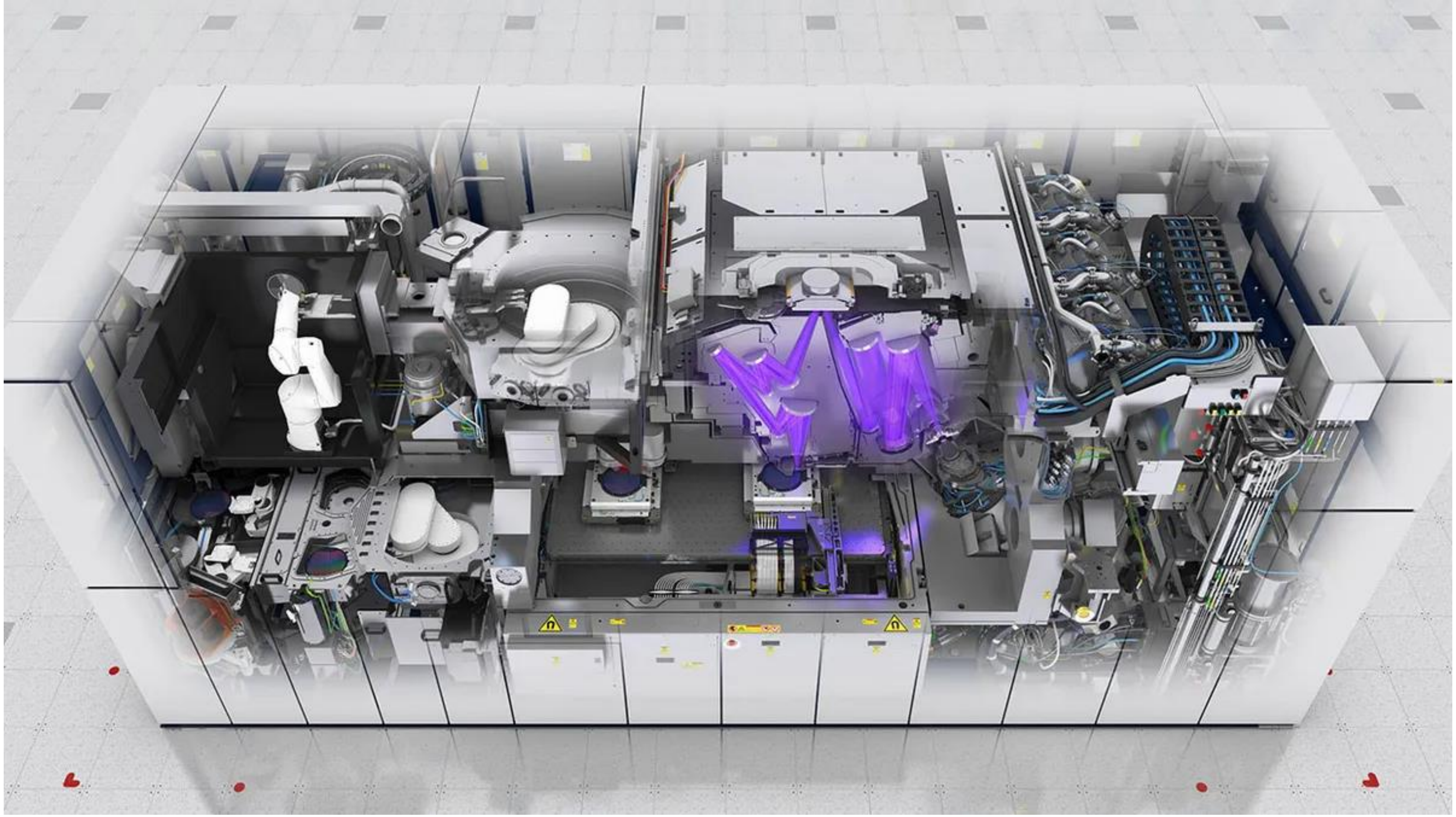
→

Reduce the wavelength
from visible to UV, to Deep UV, and to Extreme-UV (13.5 nm)



EUV lithography

https://www.asml.com/en/products/euv-lithography-systems



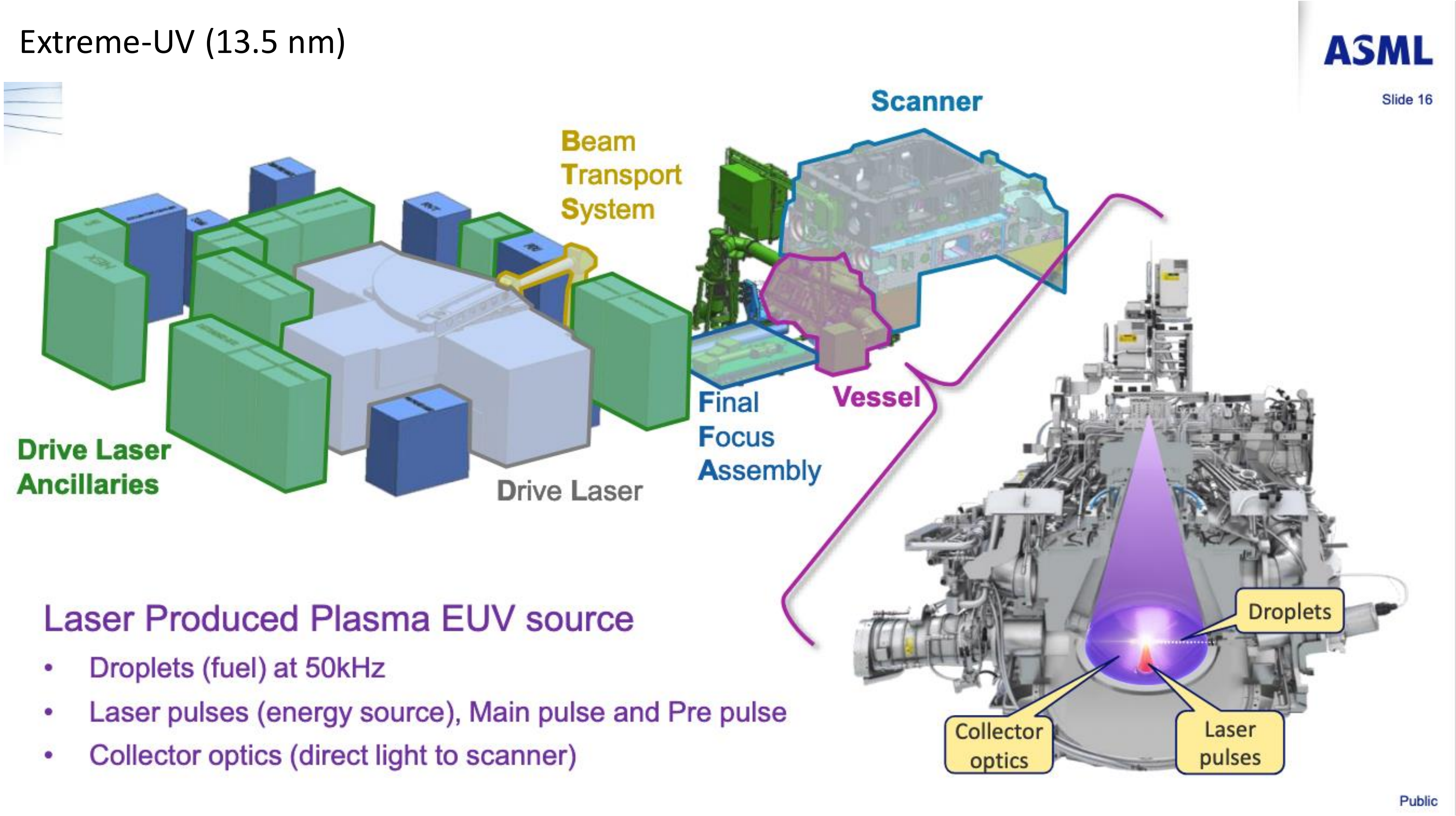
motorized high-precision positioning of the wafer (sub-nanometer) (ASML)

atomically-flat mirror optics (ZEISS)

13.5 nm EUV source: Laser-pulsed Sn droplet plasma (TRUMPF)

EUV lithography - Source

Extreme-UV (13.5 nm)



Laser Produced Plasma EUV source

- Droplets (fuel) at 50kHz
- Laser pulses (energy source), Main pulse and Pre pulse
- Collector optics (direct light to scanner)

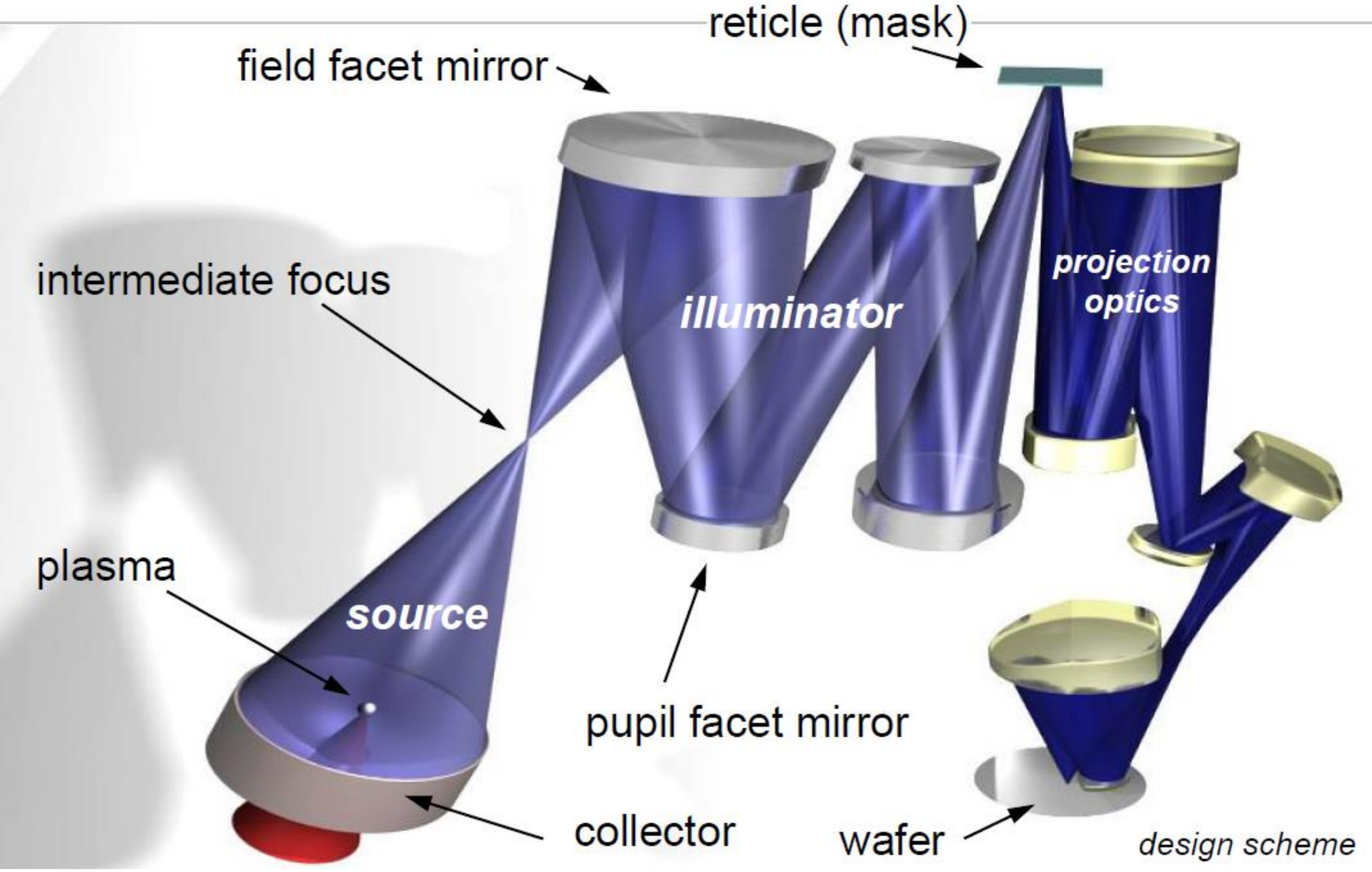
- tin (Sn) droplets in vacuum vessel
- high-power pulsed laser impacts the passing Sn droplets
- Sn atoms ionized in very high excited state, a plasma is created
- a collector mirror captures the EUV radiation emitted by the plasma
- and transfers it to the lithography system

https://www.asml.com/en/technology/lithography-principles/light-and-lasers

https://www.trumpf.com/en_CA/solutions/applications/euv-lithography/

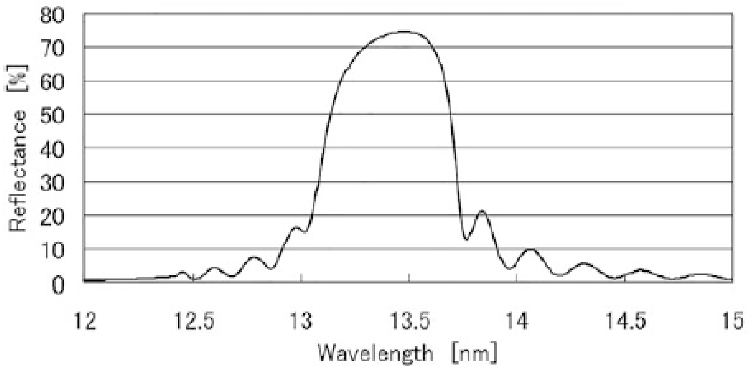
Handbook of Laser Micro- and Nano-Engineering, EUV Sources
https://link.springer.com/referenceworkentry/10.1007/978-3-030-63647-0_54

EUV lithography - Optics

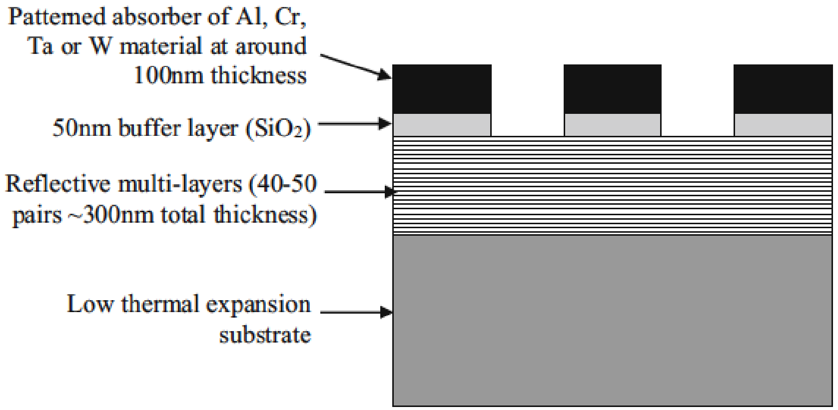


Theoretical reflectivity of a mirror consisting of Mo-Si multilayer at normal incidence of EUV radiation

mirrors in place of lenses



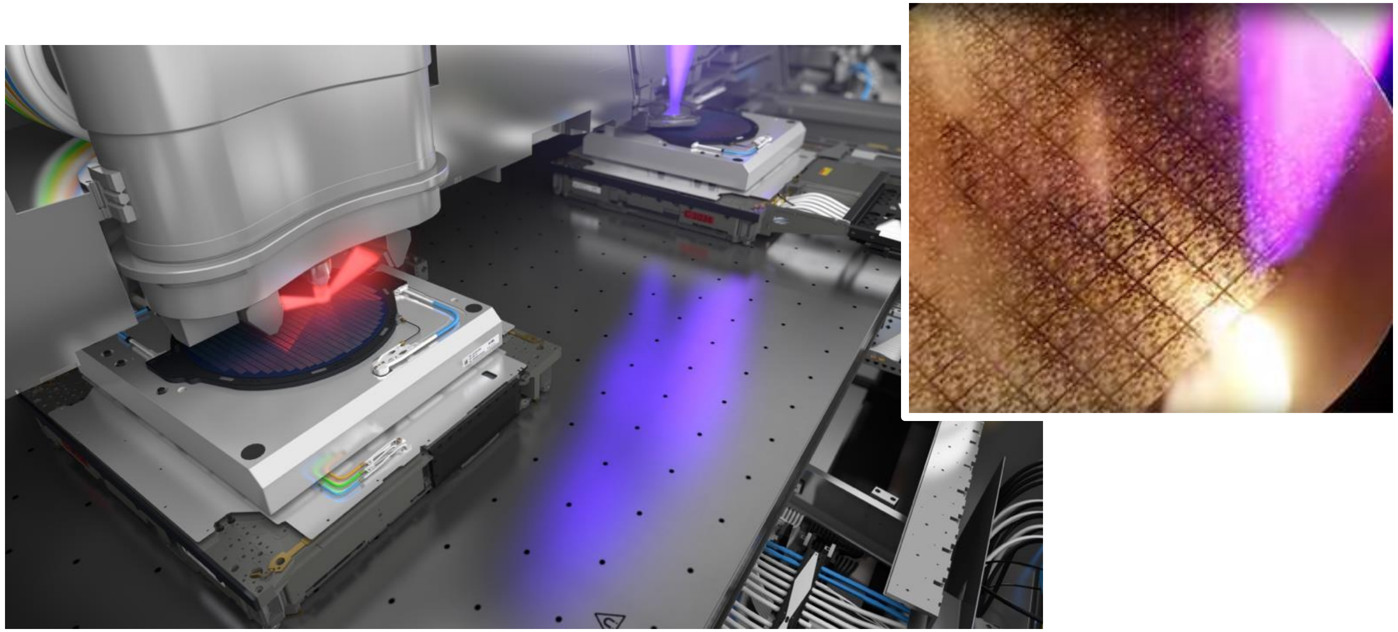
mask in reflection



https://www.euvlitho.com/2018/P22.pdf

EUV lithography - Positioning

https://www.asml.com/en/technology/lithography-principles/mechanics-and-mechatronics

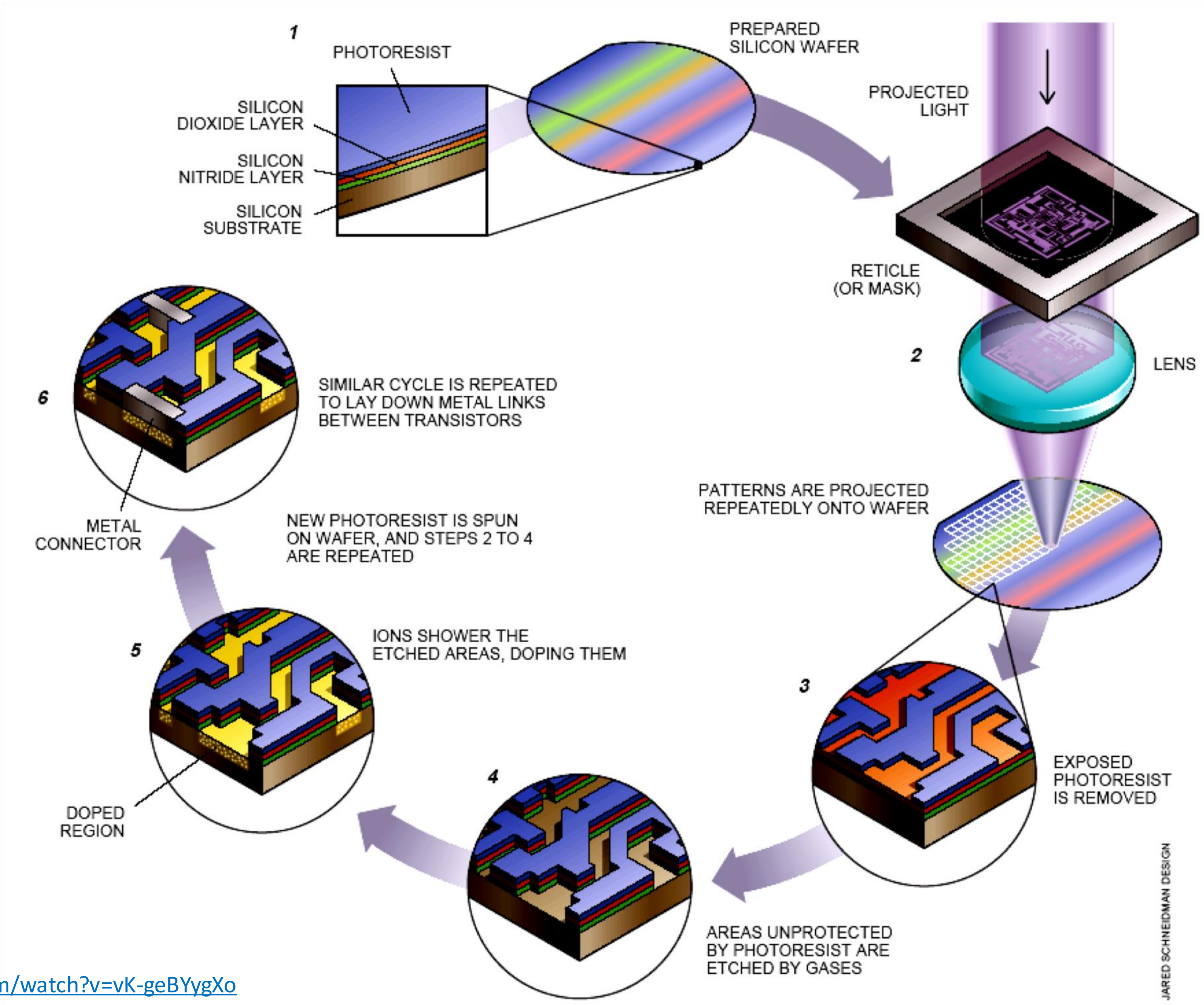


measure positioning 20’000 times per second with an accuracy of 60 pm

loading the wafer into the system, printing the pattern in almost 100 different places and then unloading the wafer as many as 275 times an hour;

magnetically levitating wafer tables

Integrated Circuit Fabrication



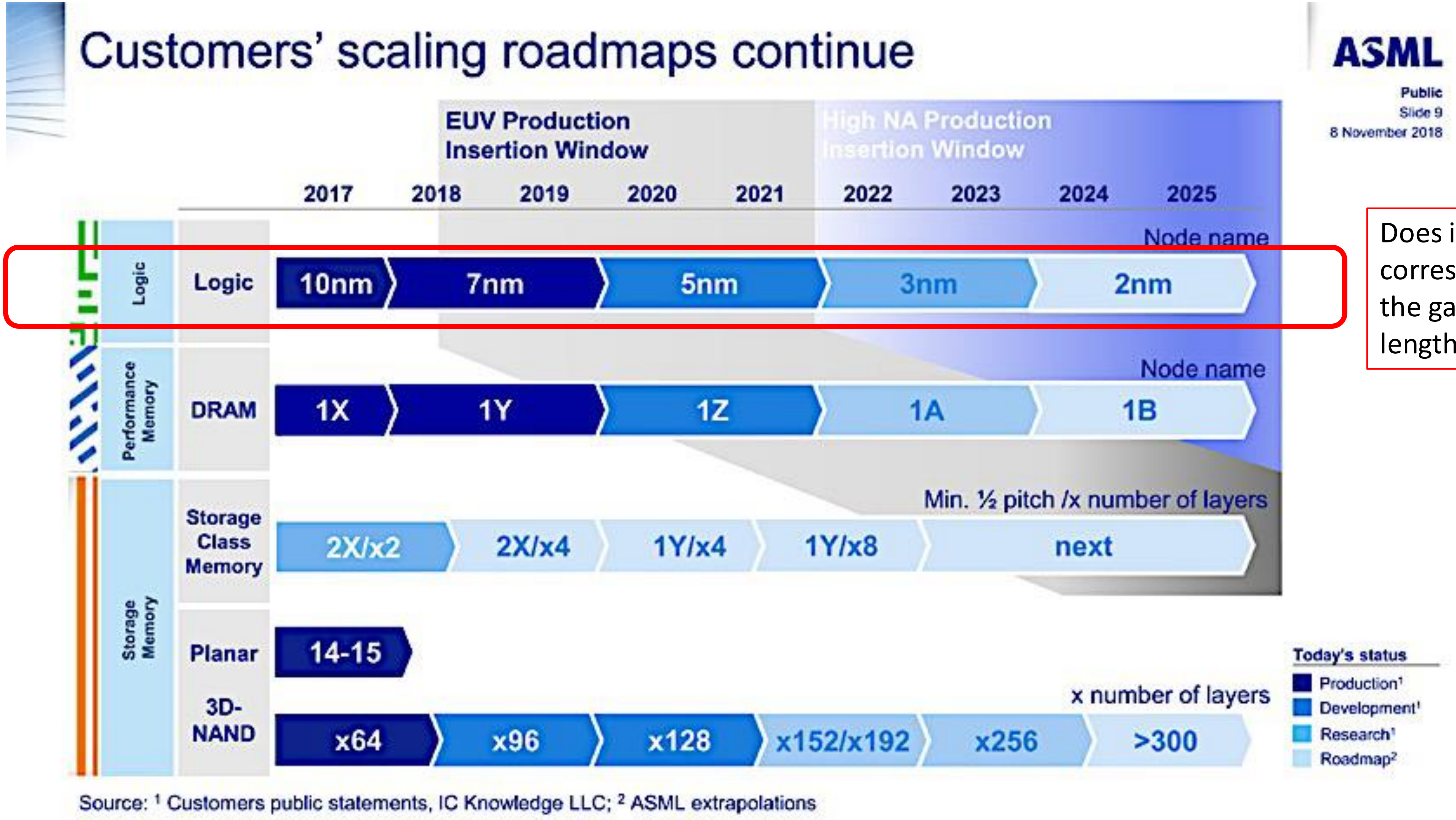
In this figure: Photolithography

Same for EUV lithography

Fabrication requires tens or even hundreds of successive steps

https://www.youtube.com/watch?v=vK-geBYygXo

Fabrication technology roadmaps



Customers' scaling roadmaps continue

ASML

Public
Slide 9
8 November 2018

EUV Production Insertion Window

High NA Production Insertion Window

2017 2018 2019 2020 2021 2022 2023 2024 2025

Logic

Logic — Node name: 10nm → 7nm → 5nm → 3nm → 2nm

Performance Memory

DRAM — Node name: 1X → 1Y → 1Z → 1A → 1B

Storage Memory

Storage Class Memory — Min. ½ pitch /x number of layers: 2X/x2 → 2X/x4 → 1Y/x4 → 1Y/x8 → next

Planar: 14-15

3D-NAND — x number of layers: x64 → x96 → x128 → x152/x192 → x256 → >300

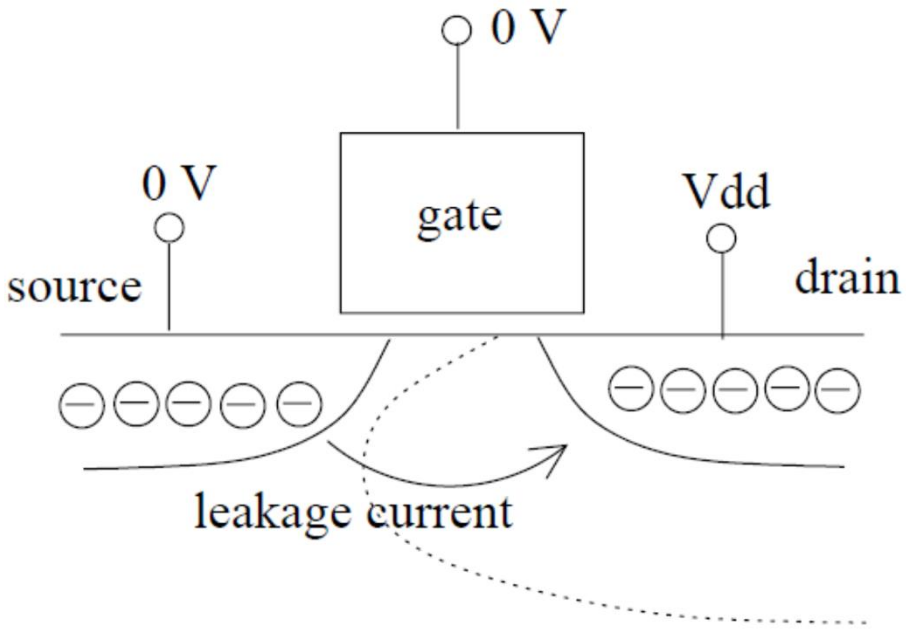
Today's status
- 🟦 Production[1]
- 🟦 Development[1]
- 🟦 Research[1]
- 🟦 Roadmap[2]

Source: [1] Customers public statements, IC Knowledge LLC; [2] ASML extrapolations

Does it always correspond to the gate length?

Scaling issues: Short-channel effect

Short-channel effect: **leakage between source and drain** caused by the overlap between depletion layers and drain voltage

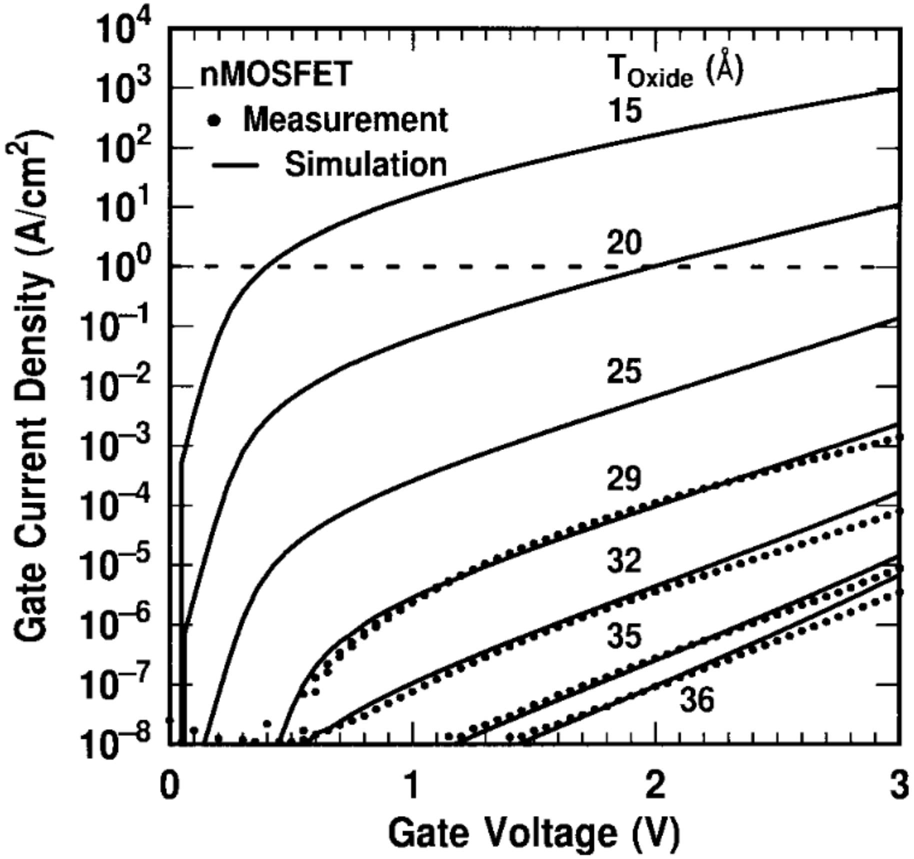


Scaling issues: Gate insulating layer

Scaling theory: reducing the lateral size implies reducing also the insulating layer thickness

- Silicon dioxide (SiO_2) is widely used as gate insulating layer: large band gap, can be formed by Si oxidation
- SiO_2 thickness: minimum is ~0.7 nm (full band gap)
- For layers thinner than 1-2nm, direct tunneling of charge carriers takes place → **gate leakage**

While the gate leakage current may be at a level negligible compared with the ON-state current, it will affect the chip standby power (need of an applied voltage for OFF-state)



acceptable value

IEEE ELECTRON DEVICE LETTERS, **18** (1997)
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=568766&tag=1

Scaling issues: Gate insulating layer

As the device area is scaled down, a minimum gate capacitance must be maintained to have an equivalent effect on the carriers in the channel:
loss of capacitor area counterbalanced by increase of capacitance density

$$C = \kappa \epsilon_0 \frac{A}{t}$$

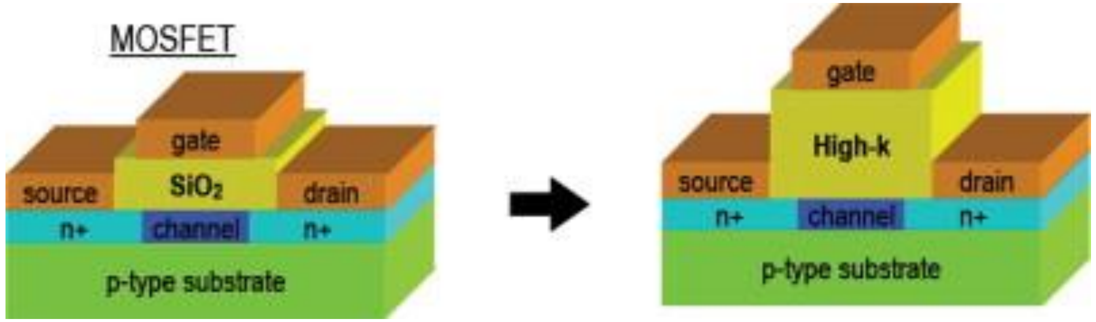
capacitance C
dielectric constant κ
area A
thickness t

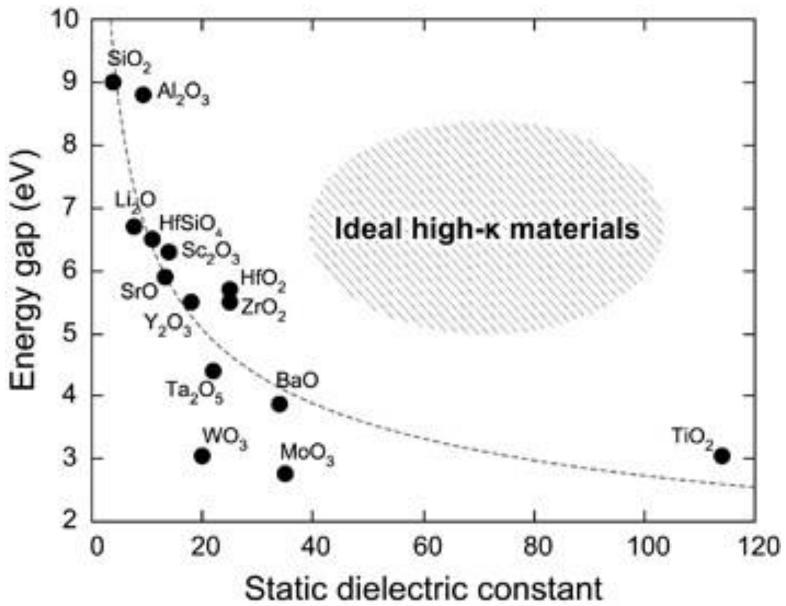
In order to maintain the capacitance:

- decrease the thickness of SiO_2 (but induces current leaks)

or

- use a material with a larger dielectric constant (higher κ)





Also important:

- band structure (band gap, band alignment)
- growth possibilities, interface quality

DOI: 0.1016/j.pmatsci.2011.01.012

https://www.sciencedirect.com/science/article/pii/S1369702104000525

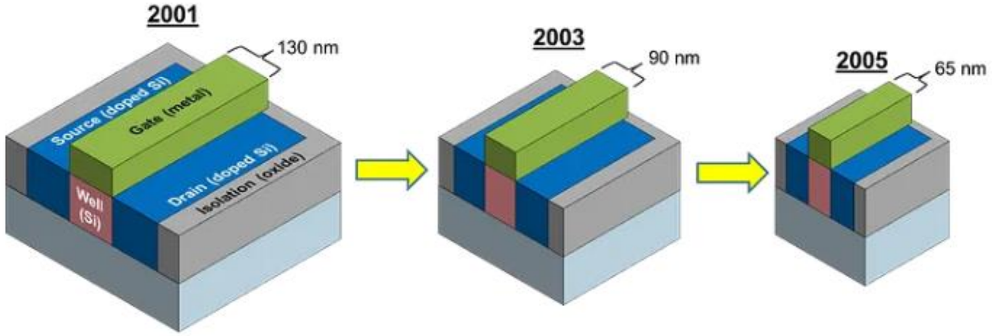
FET fabrication technology roadmaps - Nodes

In early days until 2008:

MOSFETs are purely planar, the node name (technology or process node) corresponds the size of the smallest feature on a processor, usually the **gate length** (coinciding with the channel length in planar FETS)

50 μm mid 1960's
1 μm late 1960's
350 nm late 1990's

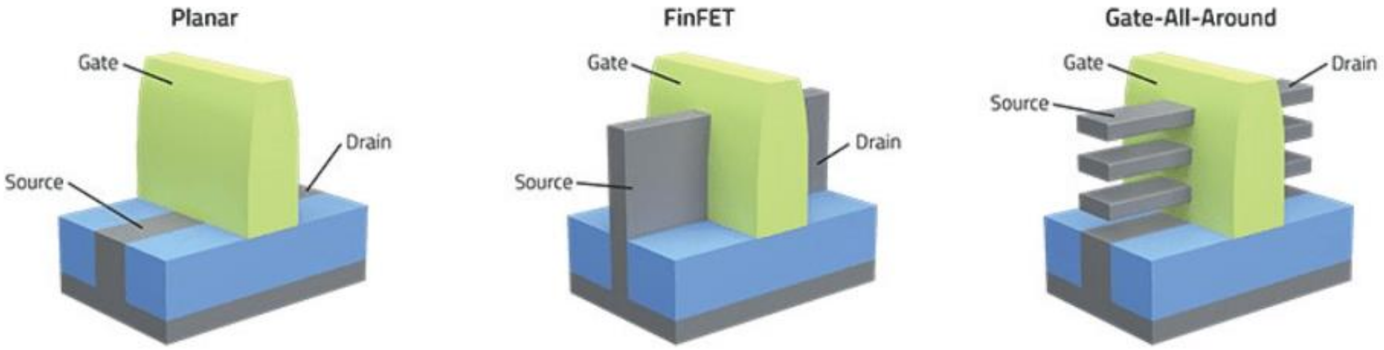
130 nm 90 nm 65 nm



After 2008:

It's not possible to reduce further the gate length without current leakage, meaning that the on/off states would not be well defined.

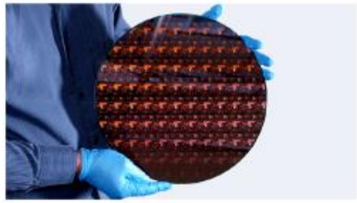
Other ways to increase the density (shape, stacking in third dimension...)



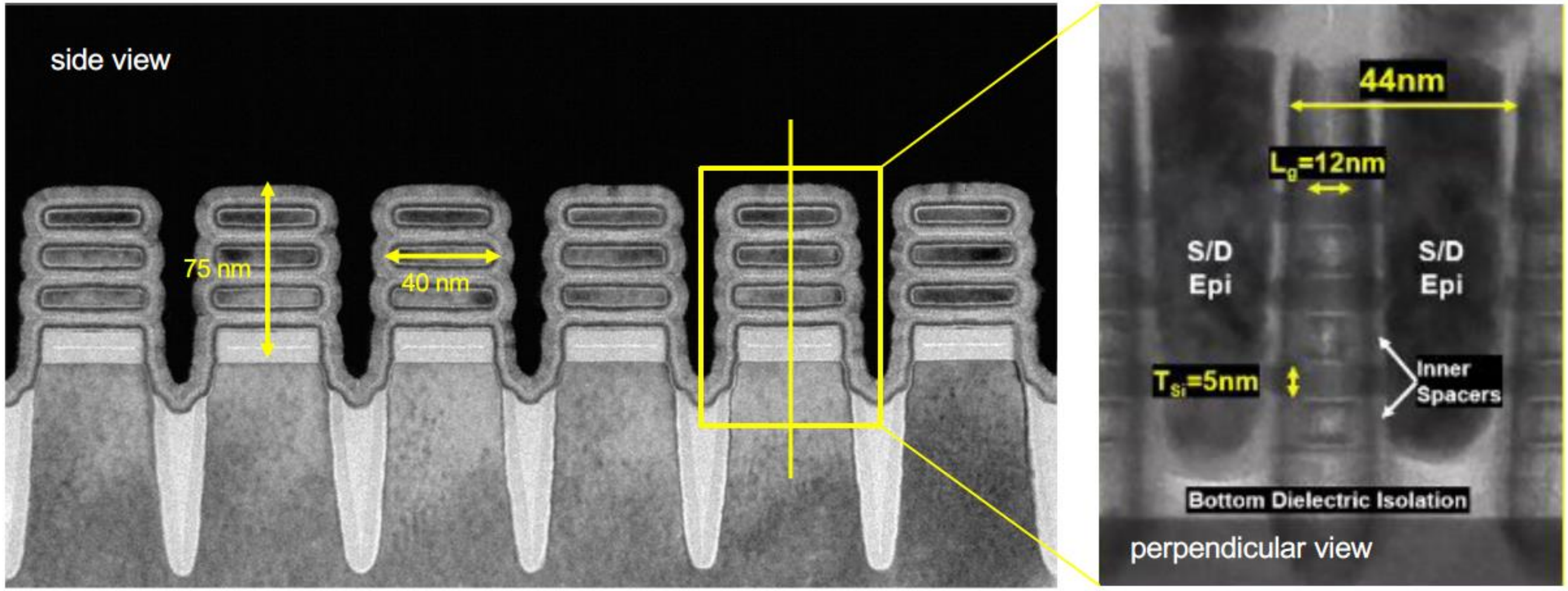
→ recover longer effective channel length (lower leakage current, better electrostatic control)

The node name is not anymore related to feature sizes; it rather indicates the **equivalent transistor density.**

Example



IBM – 2 nm Nanosheet Transistor

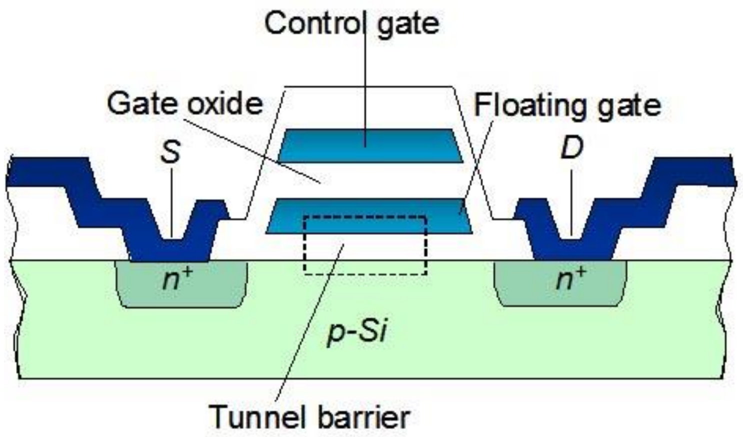


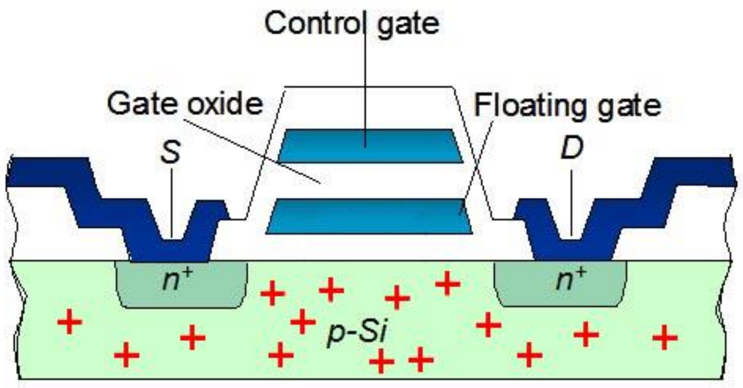
12 nm gate length!

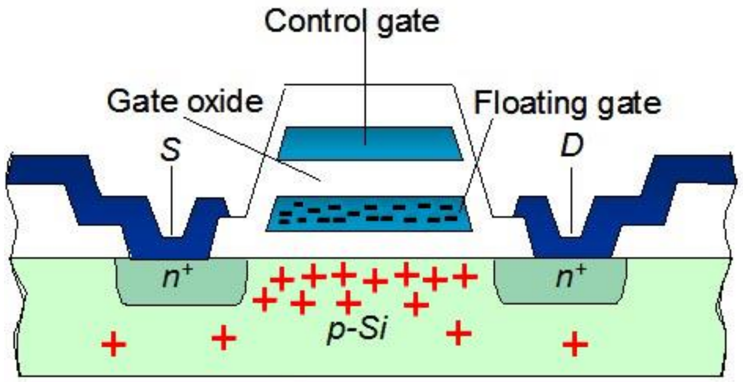
Floating gate transistor (FGT)

MOSFET is volatile: it forgets its state if power is removed and the gate voltage is not maintained
→
MOSFET can not be used to store information permanently without continuous power

modification: floating gate transistor (NAND, Flash memory, Solid State Devices SSD)



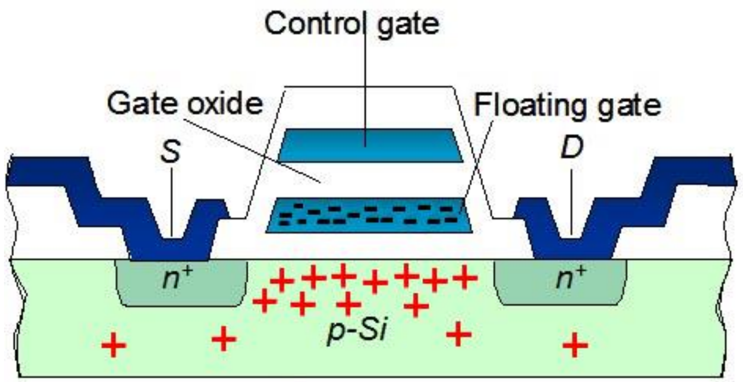


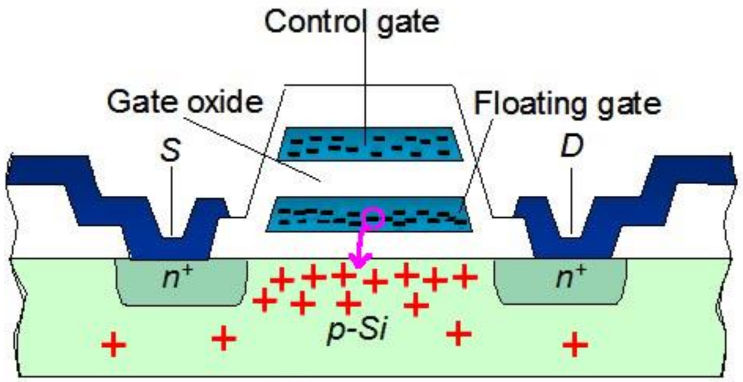


one can use the charge state of the floating gate as a bit:
the presence or absence of electrons on the floating gate is the binary bit

https://volga.eng.yale.edu/teaching-resources/flash-drives/methods-and-materials

Floating gate transistor (FGT)





Use the charge state of the floating gate as an information bit (a zero or one): presence or absence of electrons on the floating gate is the binary bit

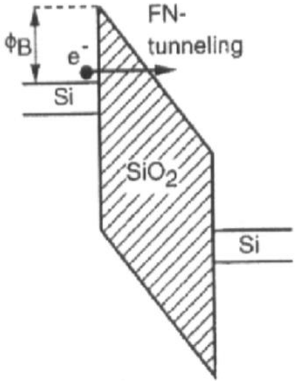
How to read it?

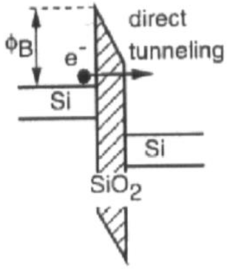
The electrons on the floating gate change the value of the gate voltage needed to achieve some desired channel conductivity:
When electrons are present on the floating gate, current can't flow through the transistor, and the bit state is 0. When electrons are removed from the floating gate, the current is enabled to flow, and the bit state is 1.

How to write it? (How do the electrons get on and off the floating gate?)

Quantum tunneling: the voltage on the control gate reduces the tunnelling barrier, and pushes the electrons away from the floating gate
Charging the floating gate is accomplished by either tunneling in the other direction, or via hot-electrons injection

Fowler-Nordheim tunneling regime: the barrier is lowered under high applied electric field and becomes “triangular”



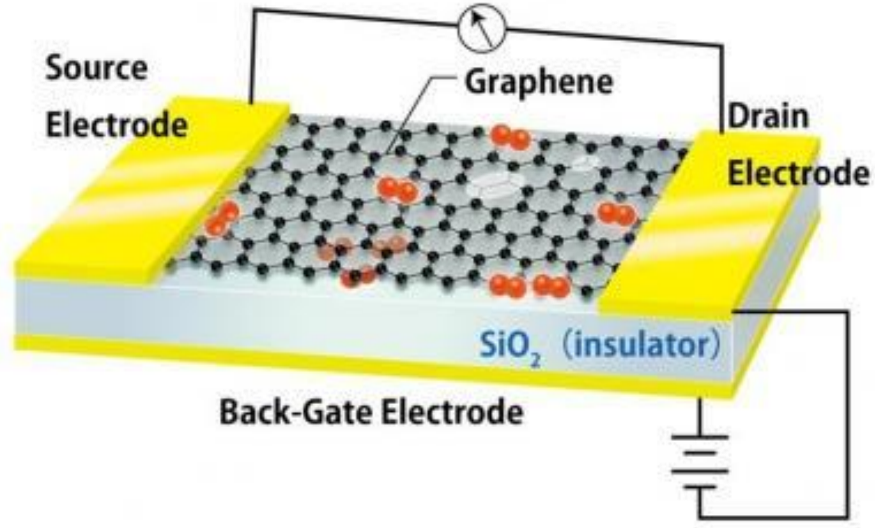


2D materials in new FETs

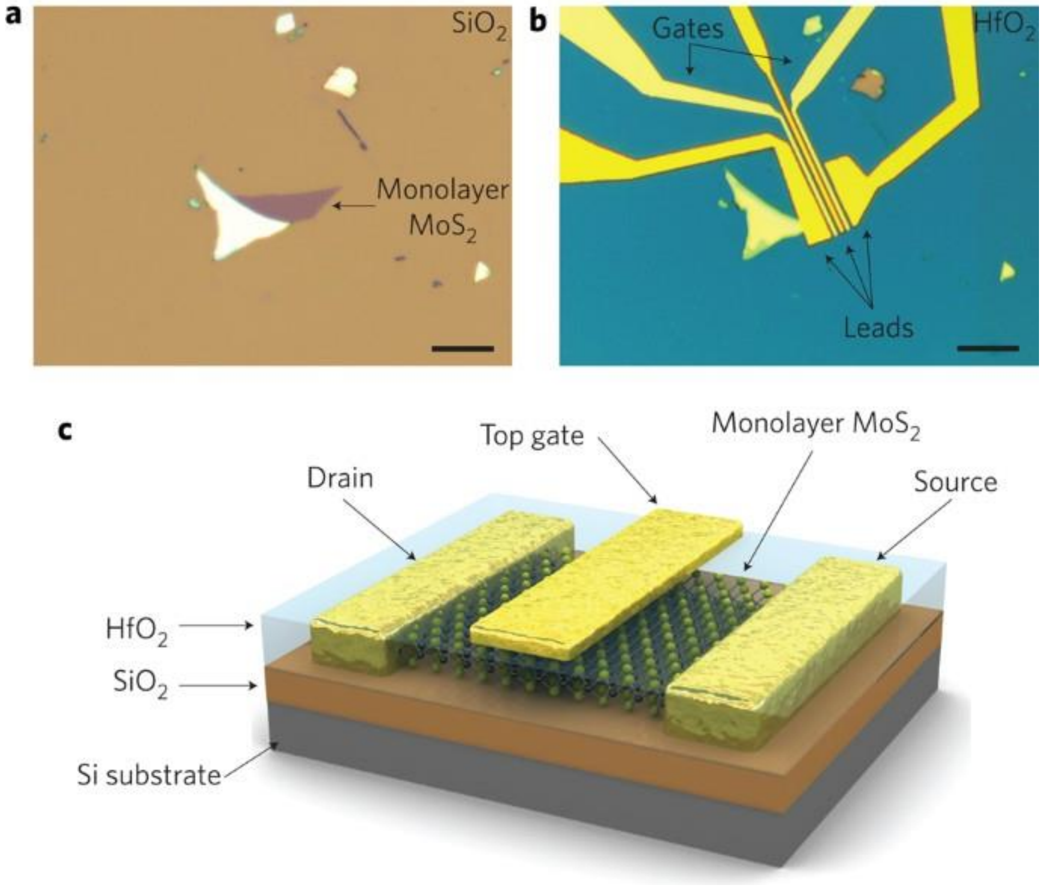
graphene-FET
the channel is made of a graphene layer

sensor applications: the channel is directly exposed to the material or environment under test (chemical-sensing applications)

example: O_2 sensor



MoS$_2$-FET
channel: single MoS$_2$ layer (0.65 nm thick)

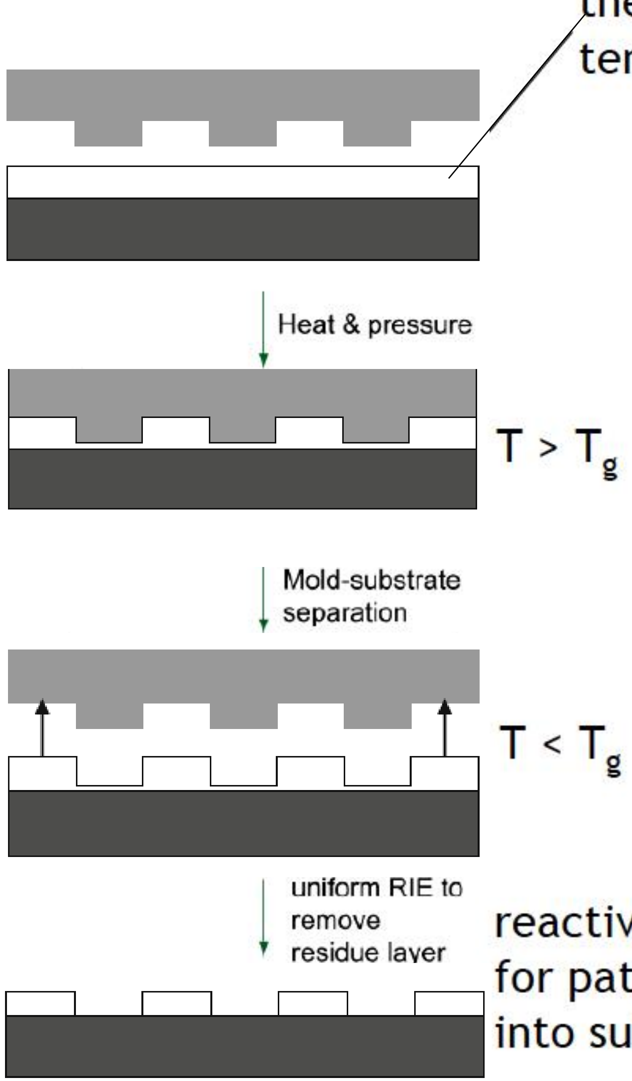


https://www.nature.com/articles/nnano.2010.279

Thermal nanoimprint lithography (T-NIL)

replication from a (soft) stamp
simple and cheap

Photonics insights **2**, R04 (2023) DOI: 10.3788/PI.2023.R04



thermoplastic resist, *e.g.*, PMMA (glass transition temperature $T_g \approx 100^\circ C$)

Heat & pressure

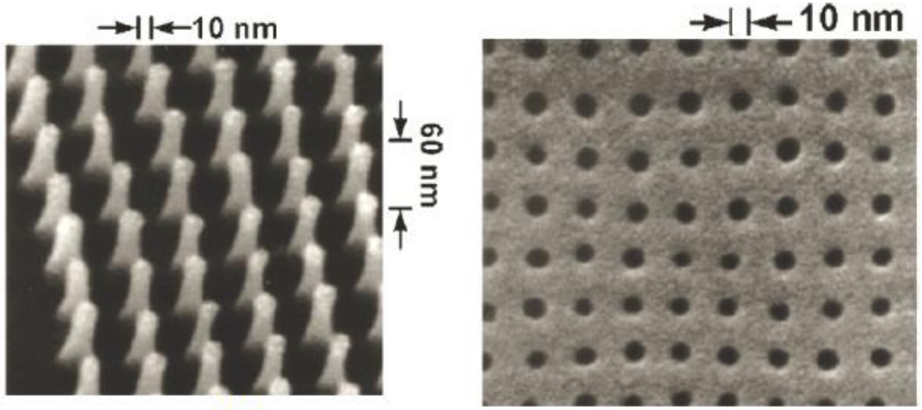
$T > T_g$

Mold-substrate separation

$T < T_g$

uniform RIE to remove residue layer

reactive ion etching for pattern transfer into substrate



mold

hole array in PMMA

Stephen Y. Chou et al., *J. Vac. Sci. Technol. B* **15**, 2897–2904 (1997)
https://doi.org/10.1116/1.589752

e-beam lithography

Based on scanning electron microscope (SEM) or transmission electrons microscope (TEM) technology

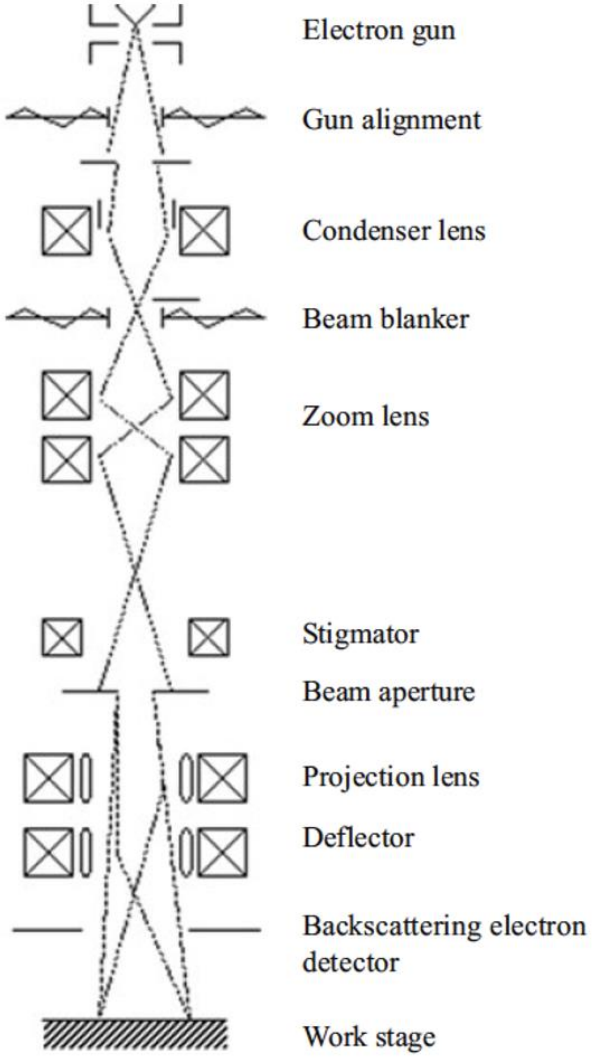
Electrons in place of photons $\lambda = \frac{h}{\sqrt{2\,m\,E}} \rightarrow \lambda = \frac{1.24}{\sqrt{eV}}$ nm

(non-relativistic approximation for simplicity, although here we should use the relativistic one)

typical energy: 10 – 200 keV

10 keV → $\lambda \approx 0.012$ nm

- the electron wavelength is not the resolution-limiting factor in e-beam lithography
- the resolution is limited by optic's aberration and electron scattering in the resist



Working principle:

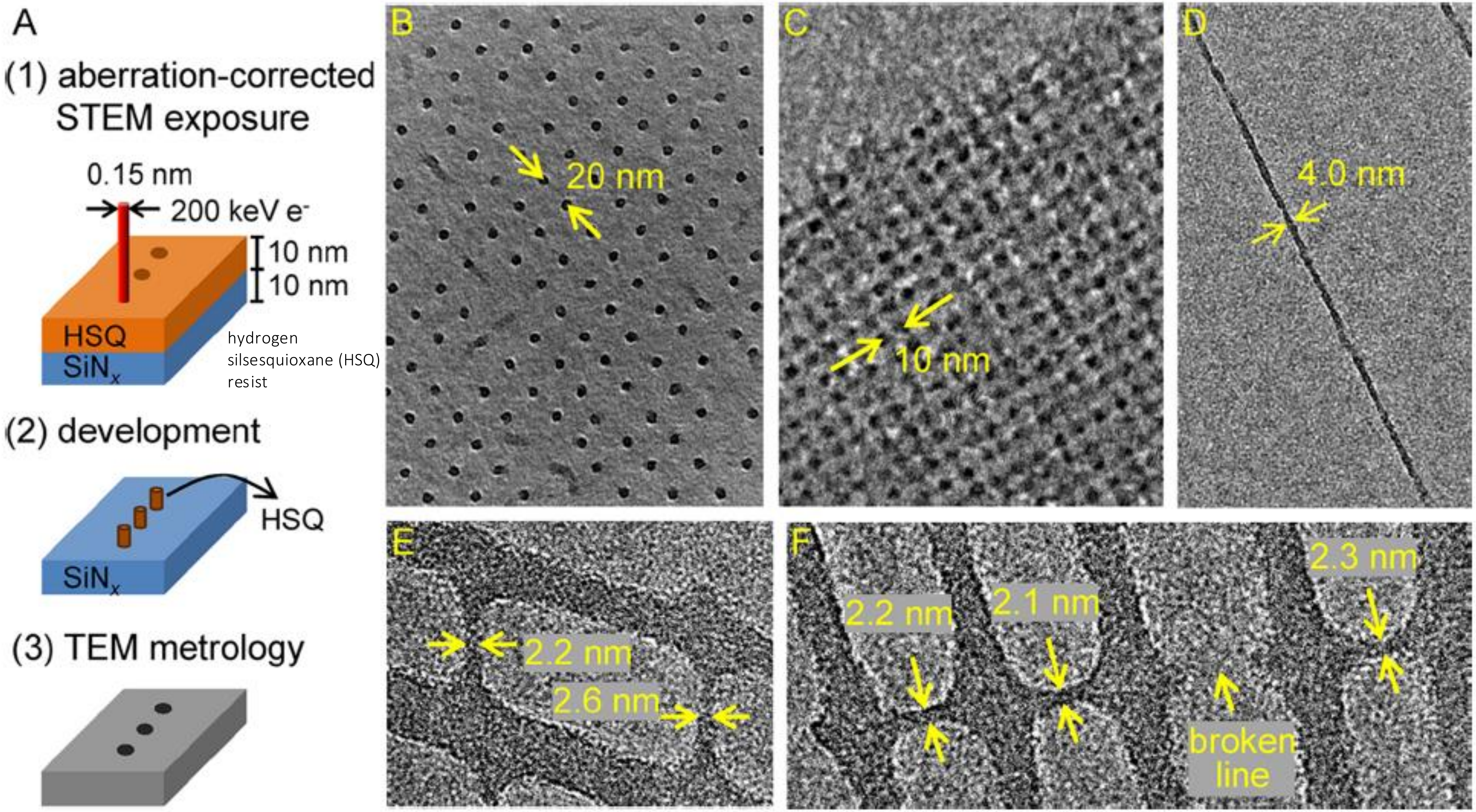
- scan a focused beam of electrons
- draw patterns on a surface covered with an electron-sensitive film (resist)
- development of the resist, etc

Characteristics and applications:

high resolution
flexible (research, prototypes)

low throughput (serial exposure)

e-beam lithography: resolution limit

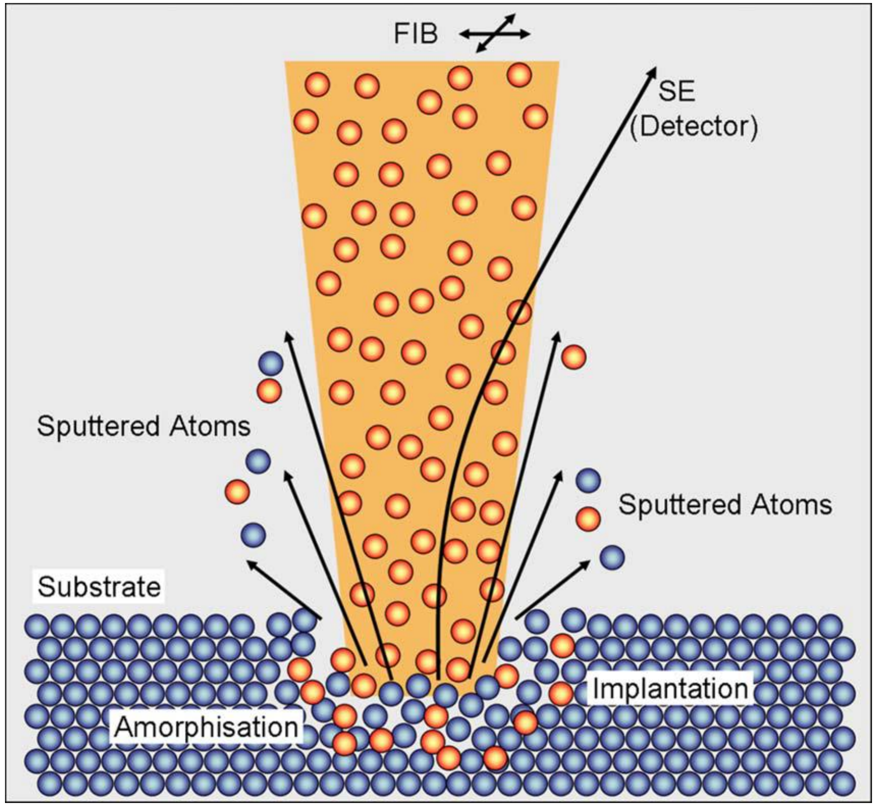


hydrogen silsesquioxane (HSQ) resist

STEM = Scanning Transmission Electron Microscope

DOI: 10.1021/nl304715p
Nano Lett. **13**, 1555 (2013)

Focused ion beam (FIB)



Ions impinging at 1 – 50 keV

ions in place of electrons, much larger mass
→
ion bombardment directly erodes the substrate to design the desired pattern

there are also versions in which some material is deposited

Spatial resolution of about 5 nm

It is a serial process (like e-beam lithography) → quite slow (typically used to shape the masks for photolithography)

ion sources:
- liquid metal ion source
- gas field ion source (He, Ne, Xe)

DOI: 10.1116/1.2955728