Semiconductor Technology from A to Z

How does it work?

Fabrication Processes

How does Chemical-Mechanical Polishing (CMP) work?

CMP simultaneously attacks the wafer surface chemically with a reactive slurry and smooths it mechanically with a polishing pad. Raised features are removed faster than recessed areas, producing a globally planar surface. This is essential for multilevel interconnects, since each additional metal layer needs a flat base.

Learn more in the Metallization section →

How does lithography work in chip manufacturing?

A light-sensitive photoresist is applied to the wafer and exposed to UV light through a mask. Depending on the resist type, either the exposed or unexposed areas dissolve away during development, transferring the mask pattern into the resist. This pattern then serves as a template for subsequent etch or doping steps.

Learn more in the chapter Exposure and Resist Coating →

How does ion implantation work?

Dopant ions are accelerated in an electric field and shot into the wafer at high energy. Penetration depth depends on energy and ion mass, while the dose sets the dopant concentration. A subsequent anneal repairs crystal damage and electrically activates the dopant atoms.

Learn more in the chapter Doping Techniques →

How does Chemical Vapor Deposition (CVD) work?

Gaseous precursors are fed into a reaction chamber, where they react at or near the heated wafer surface to form a solid film. Depending on pressure and temperature, this yields APCVD, LPCVD, PECVD, or ALD processes. The technique produces highly uniform films that conformally coat even complex 3D structures.

Learn more in the chapter Chemical Vapor Deposition →

How does plasma etching work?

A process gas is converted by electrical energy into a plasma of ions, radicals, and electrons. Reactive radicals chemically attack the material being etched, while accelerated ions remove it physically as well. This combination enables anisotropic etching with steep, precise edge profiles.

Learn more in the chapter Dry Etch Processes →

How does epitaxy work?

In epitaxy, a crystalline layer grows atomic layer by layer on a single-crystal substrate and adopts its crystal orientation. This typically happens from the gas phase (e.g., silicon vapor-phase epitaxy) under precise temperature and gas-flow control. The result is ultra-pure layers with exactly defined doping.

Learn more in the chapter Chemical Vapor Deposition →

How does a cleanroom work?

A cleanroom keeps airborne particle concentration extremely low through continuous filtration (HEPA/ULPA) and positive-pressure ventilation. Staff wear special clothing and follow strict procedures to avoid contamination. At nanometer-scale feature sizes, even a single dust particle can destroy an entire chip.

Learn more in the Wet Chemistry section →

Devices & Digital Technology

How does a MOSFET work?

A gate electrode uses an electric field to control the conductivity of a channel between source and drain. Once the gate voltage is high enough, a conductive channel forms and current can flow; below that threshold, the transistor stays off. The gate is isolated from the channel by a thin oxide layer, so essentially no control current flows.

Learn more in the chapter Field-Effect Transistors →

How does a FinFET work?

In a FinFET, the conducting channel stands as a thin, vertical "fin" on the substrate, wrapped by the gate on three sides. This multi-sided wrap gives the gate much better control over the channel than in planar transistors. That reduces leakage currents and allows further shrinking of feature sizes.

Learn more in the chapter Construction of a FinFET →

How does a transistor work in general?

At its core, a transistor is an electronically controllable switch: a small control signal (voltage or current) turns a larger current path on or off, or amplifies it. The two main families are bipolar transistors (current-controlled) and field-effect transistors (voltage-controlled). Modern chips consist of billions of such switches.

Learn more in the chapter The p-n Junction →

How does flash memory work?

Inside a flash cell transistor, an additional, electrically isolated floating gate sits between the control gate and the channel. Electrons are moved onto or off this floating gate by quantum tunneling and stay trapped there even without power. Their presence measurably shifts the transistor's switching threshold — that's the stored "0" or "1".

Learn more in the chapter Why New Memory Technologies? →

How does MRAM work?

MRAM stores data through magnetization direction in a magnetic tunnel junction, not through electric charge. A reference layer has fixed magnetization, a free layer has switchable magnetization; whether the two are parallel or antiparallel changes the tunnel resistance measurably. In STT-MRAM, the free layer is switched directly by a spin-polarized write current.

Learn more in the chapter MRAM – Magnetoresistive RAM →

How does a microprocessor work at the chip level?

A microprocessor combines a control unit, arithmetic logic unit (ALU), registers, and cache memory on a single die, connected via internal buses. The control unit fetches instructions from memory, decodes them, and directs the ALU to execute them — typically using pipelining to process multiple instructions simultaneously. Physically, this is billions of transistors wired together across multiple metal layers.

Learn more in the chapter Processor Architecture →

Optoelectronics

How does an LED work?

At a p-n junction made of a direct-bandgap semiconductor, injected electrons and holes recombine and emit a photon. The light's wavelength is set by the material system's bandgap (e.g., InGaN for blue, AlGaInP for red). Quantum well structures in the active region significantly boost the efficiency of this recombination.

Learn more in the chapter Light-Emitting Diodes (LED) →

How does a laser diode work?

As with an LED, light is generated by recombination at a p-n junction, but here inside an optical resonator that reflects the light back and forth repeatedly. Above a threshold current, stimulated emission dominates over spontaneous emission — population inversion occurs and coherent laser light is produced. Depending on the resonator design, this yields Fabry-Pérot, DFB, or VCSEL lasers, among others.

Learn more in the chapter Laser Diodes →

How does a solar cell work?

When light hits a solar cell's p-n junction, photons with sufficient energy generate electron-hole pairs in the semiconductor. The junction's built-in electric field separates these charge carriers before they can recombine and drives them into the external circuit. This photovoltaic effect converts light energy directly into electrical energy.

Learn more in the chapter Fundamentals of Optoelectronics →

Packaging & Other Topics

How does 3D chip integration (advanced packaging) work?

Instead of fitting all functions onto one large die, multiple dies or wafers are stacked on top of each other and electrically connected using through-silicon vias (TSVs) or hybrid bonding. This drastically shortens signal paths and allows different technologies (logic, memory) to be combined in one package. HBM memory is the best-known example of this technique.

Learn more in the chapter 3D Stacking →

How does a MEMS sensor work?

Micro-electromechanical systems combine tiny mechanical structures — membranes, beams, comb electrodes — with electronic readout on a single chip. Physical quantities like acceleration or pressure minutely deform the mechanical structure, which is measured electrically as a change in capacitance or resistance. They're manufactured using adapted semiconductor processes such as surface or bulk micromachining.

Learn more in the chapter Introduction to MEMS →

How does wire bonding work?

In wire bonding, a thin gold or aluminum wire connects the chip's contact pads to the package leads. A combination of pressure, ultrasonic energy, and sometimes heat welds the wire to the pad without melting it. It remains the most widely used method for electrical interconnection in chip assembly today.

Learn more in the chapter Bonding →

Why do chips keep getting smaller (Moore's Law explained briefly)?

In 1965, Gordon Moore observed that the number of transistors on a chip roughly doubles every two years — driven by ever-finer lithography and denser devices like FinFETs. Smaller transistors switch faster, use less energy per switching event, and lower the cost per function. In recent years this scaling has been slowing down, which is why advanced packaging and chiplets are gaining importance as a complementary path.

Learn more in the chapter Applications & Outlook →